ADCHEM 2009

IFAC Symposium on

Advanced Control of Chemical Processes

July 12-15, 2009

Koç University, Istanbul, Turkey

Symposium Preprints

Part 1

Editors: Sebastian Engell

Yaman Arkun

International Program Committee

Sebastian Engell (Chair)

Technische Universität Dortmund, Germany

Jorge Mandler (Industrial Co-Chair)

Air Products, USA

IPC Coordination

Christian Sonntag, Thomas Tometzki

Technische Universität Dortmund, Germany

IPC Members

O. Abel (DE), F. Allgöwer (DE), J. Alvarez (MX), H. Arellano-Garcia (DE), T. Backx (NL), T. Badgwell (US), R. Berber (TR), L.Biegler (US), D. Bonvin (CH), R. Braatz (US), M. Çamurdan (TR), T. Chai (CN), M. Chiu (SG), P. Christofides (US), A. Çınar (US), I. Craig (ZA), P. Daoutides (US), C. de Prada (ES), D. Dochain (BE), F. J. Doyle, III (US), G. Dünnebier (DE), F. Forbes (CA), B. Foss (NO), F. Gao (HK), C. Georgakis (US), M. Guay (CA), S. Hasebe (JP), K. Hoo (US), M. Hovd (NO), B. Huang (CA), C. D. Immanuel (UK), S. B. Jorgensen (DK), M. V. Kothare (US), C. Kravaris (GR), J. Lee (US), D. Lewin (IL), P. Li (DE), J. Mandler (US), J. Marchetti (AR), W. Marquardt (DE), M. Nikolaou (US), D. Odloak (BR), C. Özgen (TR), A. Palazoğlu (US), S. Park (KR), M. Perrier (CA), H. Preisig (NO), Raisch (DE), J. Rawlings (US), J. Romagnoli (US), C. Scali (IT), A. R. Secchi (BR), S. L. Shah (CA), S. Skogestad (NO), M. Soroush (US), B. Srinivasan (CA), M. Tade (AU), N. Thornhill (UK), J. O. Trierweiler (BR), M. Türkay (TR), A. Vande Wouwer (BE), E. Ydstie (US), E. S. Yoon (KR), C.-C. Yu (TW), D. Zhou (CN)

National Organizing Committee

Yaman Arkun (Chair)

Koç University, Istanbul, Turkey

NOC Members

Erdoğan Alper, Tahsin Bahar, Ridvan Berber, Mehmet Çamurdan, Devrim Kaymak, Hitay Özbay, Canan Özgen, Metin Türkay, Mustafa Türker

NOC Coordination

Nur Diğdem Burak

Koç University, Istanbul, Turkey

COPYRIGHT CONDITIONS

Contents

Plenary Lectures

Challenges and problems with advanced control and optimization technologies

Campos, M.*, Teixeira, H.*, Liporace, F. and Gomes, M.****

** PETROBRAS / CENPES / Engenharia Básica Abastecimento e Gás&Energia/ Automação, Equipamentos Dinâmicos e Confiabilidade, Av. Horácio Macedo, 950 - Cidade Universitária, Ilha do Fundão, Rio de Janeiro, 21949-915, Brazil (Tel:55-21-3865-6347; e-mail: mariocampos@petrobras.com.br).*

*** PETROBRAS / CENPES / P&D de Gás, Energia e Desenvolvimento Sustentável / Gás Natural / Célula de Otimização e Eficiência Energética, Av. Horácio Macedo, 950 - Cidade Universitária, Ilha do Fundão, Rio de Janeiro, 21949-915, Brazil.*

Abstract: Oil & Gas companies continuously try to create and increase business value of their installations (platforms, refineries, etc). Particularly the increasing energy consumption on a worldwide basis and, as a result, the substantial increase in prices volatility is a major drive for better advanced control and optimization technologies. Advanced control and optimization system can play an important role to improve the profitability and stability of industrial plants. This paper discusses the problems and challenges of advanced control and optimization in petroleum industries nowadays. It emphasizes the importance of control performance assessment technology to maintain a good regulatory control and the difficulties in using these technologies. It also shows the importance of malfunction detection and diagnosis advisory system for critical equipment in order to increase the operational reliability. Model predictive control (MPC) has become a standard multivariable control solution in the continuous process industries, but there are still many open issues related to accelerate a new implementation and maintain the controller with a good performance along the years. Real time optimization tools also impose new challenges for Oil & Gas industries application, which are discussed in this paper.

Keywords: performance assessment, regulatory control, advanced control system, real time optimization

1. INTRODUCTION

The advanced control and optimization systems in oil & gas and petrochemical plants are an industrial reality (Qin and Badgwell, 2003). These advanced systems provide many advantages for the process units, as improved stability and safety, respect to constraints and higher profitability. PETROBRAS has been investing in the development of these systems for several years. Advanced control system is already a consolidated technology in its refineries with many model predictive controllers implemented (Zanin and Moro, 2004). However, the application of real time optimization (RTO) is recent, although this technology can bring great economical earnings, besides to increase the energy efficiency and minimization of emissions.

To install and maintain these advanced systems with good performance is a great challenge. Its performance is influenced by instrumentation problems, bad tuning of the regulatory and advanced control, unreliable process dynamic models (Ender, 1993; Kern, 2007), unmeasured disturbances, etc.

This article will discuss the problems and challenges of advanced control and optimization in petroleum industries nowadays. It discusses some tools for diagnosis and tuning of the regulatory and advanced control, and the challenge associated with the real time optimizers. In spite of the several tools in the market that deal with industrial control and optimization solutions, PETROBRAS has decided to invest on the development of its own tools and solutions in many situations, usually in association with some Brazilian universities. The goal of this paper is to show some challenges faced, solutions and results obtained in PETROBRAS facilities.

2. REGULATORY CONTROL LEVEL

Process control aims to maintain certain variables within their desirable operational limits and could be visualized as a pyramid. In the base of this pyramid, the first level is the regulatory control, that uses PID controllers (Campos and Teixeira, 2006; Ogata, 1982) and is configured in the digital systems (DCS - Distributed control system or PLC - Programmable logical controllers). In a second level, we have the advanced control systems that use for instance Model Predictive Control (MPC). This algorithm considers the interaction between control loops, and includes an optimization layer of the industrial plant. These algorithms are usually implemented in a process computer that communicates with DCS or PLC systems by the use of OPC protocol (OPC, 2008). The outputs of this advanced control are usually the set points of the PID controllers. The

architecture is conceived in such a way that if there is a failure in the advanced control level, the plant operation continues with the last PID set points in the DCS.

An advanced control system won't reach the expected benefits if is turned off constantly for the operators. Therefore, the instruments, valves and the regulatory control loops (PIDs) should operate appropriately. Hence, the performance of the regulatory control is fundamental for the success of the advanced control system. An industrial plant usually has hundreds of control loops, and less and less engineers to maintain the system. Therefore, the industries need tools to perform automatic analysis and diagnoses of the problems associated with the regulatory control. For example, these tools should be able to detect failures with the instrumentation (miscalibration, badly sizing, sensor noisy, out of scale, measurement resolution, etc.), non linear behavior in the process due to changes in the operational point, bad PID tuning (oscillation, stability, etc.) and control strategy problems (coupling between control loops, degrees of freedom, etc.).

There are several tools in the market that help engineers to maintain the regulatory control, but most of them require a well-trained engineers to interpret, analyze and define the correct actions, for instance: to change a control valve, tune PID controllers or to implement a new control strategy (decoupling, feedforward), etc. These engineers should also know very well the process in order to evaluate the better actions to be taken.

The great challenge for these tools will be to incorporate more "intelligence" to help engineers in the definition of the better actions. For instance, in certain case, only PID tuning could reach 80% of improvement in process variability reduction, and in some case, the process performance would improve only 10%. A lot of times in industries the engineer spends time and money with an action that won't bring great results. So, it is clear the importance of a tool that could perform the automatic diagnosis and assessment of the regulatory control (Farenzena et al., 2006). The most important features of this tool should be to have automatic ways to prioritize the actions for each process that might result in a better performance, and also to provide a standardized metric to compare different actions in different processes, even in different scales such as economical, environmental or safety (Harris, 1989; Kempf, 2003; Farenzena and Trierweiler, 2008). These features are a great development challenge for these tools.

Despite the several tools in the market, PETROBRAS and Federal University of Rio Grande do Sul (UFRGS) have developed their own tool, the software called “BR-PerfX”. Its main purpose is to compute some universal key performance indicators that reduce the subjectivity in the analysis and help engineers in their assessments and decisions about problems affecting the regulatory control.

In order to face the PID tuning problem, PETROBRAS and Federal University of Campina Grande (UFCG) developed the software "BR-Tuning" (Schmidt et al., 2008; Arruda and Barros, 2003), which is comprised by a group of techniques regarding open and close loop identification and the proposition of new tuning parameters. It communicates directly with the process automation system (DCS or PLC) using the OPC protocol.

As it was said previously, the challenge is to develop an "intelligent" layer that helps to make a diagnosis based on several indexes or indicators. The integration between different tools is also an important concern. The use of the OPC standard for the exchange of information could be an option. So, each tool could make available their indicators to others tools through OPC. This way, the engineers' work would be facilitated, avoiding losses of time and money.

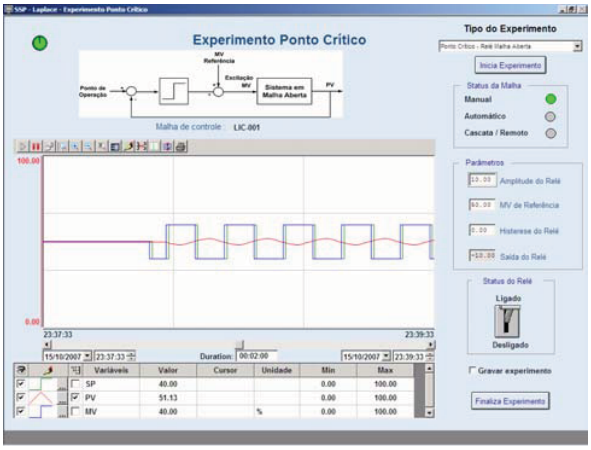


Fig. 1. BR-Tuning interface**.**

The challenges in relation to controllers' tuning are associated mainly with the identification of the models, the determination of the process non-linearities, interaction between control loops, as well as defining the desired performance for each control loop.

There are some processes where the disturbances’ pattern can change with the time, as in some off-shore petroleum platform. The slug flow can change its intensity for example due to changes in the gas-lift. So, we don't have a PID tuning parameters that are good for all these different situations. In this case, it was developed an "intelligent" system that supervises the process plant and changes the PID tuning automatically when necessary. This control strategy is equivalent a "gain-scheduling" where the control performance (deviation between the process variable and the setpoint) is evaluated during a time, and the system decides what is the best tuning for that moment. All the possible values for the PID tuning are chosen off-line. This system was installed in several PETROBRAS' platforms. The figure 2 shows the system changing the PID tuning parameters and the level performance. This project used a tool called MPA, which was developed by Catholic University of Rio de Janeiro (PUC-RJ) to PETROBRAS.

Another challenge is the development of non-linear controllers for some special cases, for example to pH control

in certain plants, although PID will continue to be the algorithm more used in this regulatory layer control for several years.

Researches and developments for the regulatory control level are still necessary, and they can bring great economical earnings. For example, an application of these tools (evaluation, tuning and changes in control strategy) allows an increased of about 9% in the production of LGN (Liquefied Natural Gas) in a natural gas plant (Campos et al., 2007).

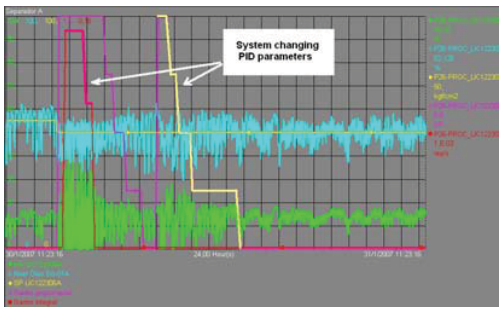


Fig. 2. Performance of this control strategy in production platform (1 day).

3. ADVANCED CONTROL SYSTEM

The multivariable predictive controllers (MPCs) are powerful tools for the process optimization and are available in many industrial plants. This system can increase feed and preferred product rates, reduce energy consumption and waste material. These benefits are more visible in complex processes where challenging dynamic responses (significant time delays, non-minimum phase responses, control loop interaction, etc.) due to disturbances (feed flow and composition, energy integration, usefulness, etc.) that must be dealt with while taking into account process constraints and trying to pursue the best economic performance. As an example of the benefits achieved, figure 3 shows an increase of about 16% in the LPG yield due to the implementation of an Advanced Process Control (APC) system in a natural gas plant.

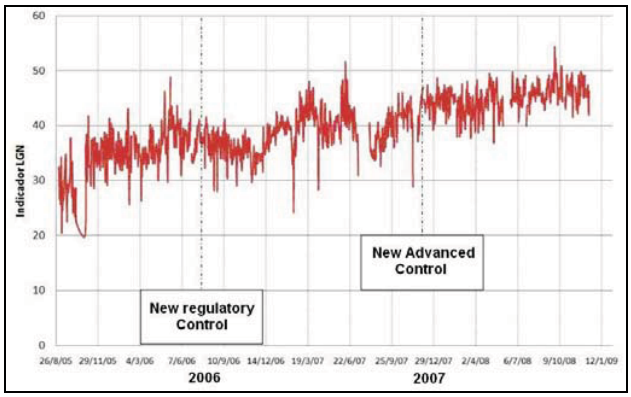


Fig. 3. LPG yield increase in a natural gas plant due to MPC.

However, even if MPC systems are nowadays seen as a commodity, there is still much to be done, due to the significant gap between the recent MPC technologies development in the academy and those effectively used on industrial plants. Most industrial MPC applications are based on the most traditional approaches: linear algorithms based on step-response models obtained through traditional step tests.

MPC maintenance

MPC performance decay throughout time is a well-known and widely reported fact (figure 4). If no maintenance work is done, the operators end up turning them off. There are many causes for this behaviour:

- Changes in the units operational objectives;
- Equipments efficiency losses (fouling);
- Changes in the feed quality;
- Problems in instruments and in the inferences;
- Lacks of qualified personnel for the controller's maintenance.

Therefore, the first great challenge associated with MPC control is to have reliable tools to keep performance and diagnose problems.

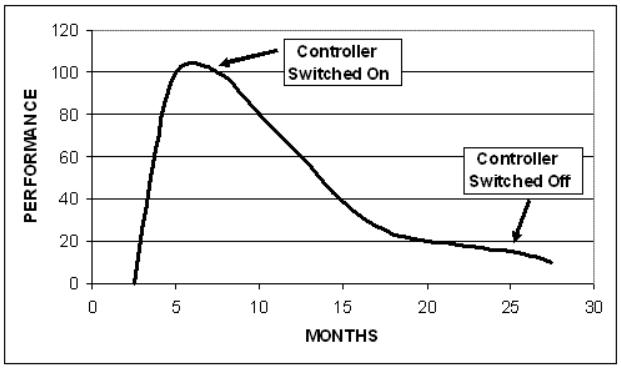


Fig. 4. Advanced Control Performance during the time.

Therefore, industry needs better tools to help maintenance personnel to answer the following questions:

- Is advanced control system accomplishing their objectives?
- What is its performance?
- Is the process optimized?
- What are the benefits?
- How is the level of disturbances?
- What is operational factor of the controller?
- How are the operators adjusting the limits of the manipulated variables?
- Are manipulated variables very limited?
- What is the variability of the main controlled variable?
- Is the process operating close to the constraints?

It is necessary a tool not only to answer these questions, but the system point out the causes of the bad performance: bad models, bad controller tuning, inference problems, non-linearities, frequent changes in the operation point, new constraints not considered in the design?

Nonlinear models, Identification and Model mismatch

Many different and even sophisticated approaches have been proposed in order to allow MPC algorithms to cope with process nonlinearity. Bequette (2007) presents a recent review on the subject. However, despite all this effort, industrial Nonlinear MPC (NMPC) applications are relatively few, and most of these are based on the simplest approaches.

One possible reason for that might be simply that the nonlinear behaviour is not known, and any lack of performance is seen as a typical model mismatch.

Another possibility might be that the nonlinear behaviour is known, but can not be easily determined with traditional plant tests. One way to overcome these problems might be the use of rigorous dynamic simulators, to improve the understanding of the process behaviour. Information obtained with dynamic simulation could be combined to the existing linear model in order to provide a reliable nonlinear one. Dynamic simulation might be useful also to find out the best way to characterize the observed nonlinearity. Once more, although there is availability of dynamic simulators, there is not much use of them in industrial applications.

Process identification of complex processes is still a hard task, where a significant part of the effort on MPC implementation is spent.

In order to address this problem, some commercial tools have been conceived in this decade for closed-loop identification. These tools are based on efficient ways to perform step tests allied to modelling strategies for minimization of the model order. While this approach has proved to be useful and promising, it is still a hard task to apply these techniques to complex processes, especially when dealing with noisy data. It seems to be a lot of space for development in this area.

Another interesting way to reduce implementation time can be the use of algorithms for automation of the plant test.

Tuning

MPC tuning is another interesting issue, where new technologies might help to reduce implementation time and also on the maintenance task.

Some interesting ideas have been proposed (Trierweiller and Farina, 2003) that try to combine desired and achievable performances. However, the controller tuning still consume time and is critical points for controller performance. Normally, all MPC tuning methods consider a square controlled variables x manipulated variables matrix, but, in fact all controller has a rectangular matrix that means different tuning scenarios depending on which constraints is active.

Another big challenge is to reduce the application time and maintenance time. For this, it is believed that the main critical points are:

- Tools for the development of inferences:
  - Use of rigorous dynamic simulators, or statistical methods for better inferences using less laboratory analysis data.
- Dynamic models identification:
  - Automation of the identification tests, minimizing problems and loss of data;
  - Efficient tools for closed loop identification;
  - Characterization and identification of the non linearities of the process.
- Better tools for tuning the predictive controller:
  - How to define the priorities in the several operating points of the controller and change automatically the tuning parameters. This activity is still done by trial and error in many industrial cases.

New advanced controllers that contemplate these aspects will help the users to implement and maintain these industrial systems.

4. REAL TIME OPTIMIZATION

Real Time Optimization (RTO) technology is a powerful tool for the continuous search of the most profitable way to run petroleum and petrochemical process units. Cutler and Perry (1983) state that despite being a hard and complex task, its potential benefits are relevant and might provide profit increases around 6 to 10% when allied to Advanced Process Control (APC).

The task of an RTO application is to make the best of an existing process unit, adjusting its process variables for every new change of external conditions, like operational variables, feed compositions and process constraints. The RTO benefits are usually associated with the maximization of products and minimization of the specific energy consumption and other resources, depending on the following factors:

- Market availability
- Products prices and feed costs
- Safety and environmental constraints
- Product specifications

The central figure of an optimization application is the mathematical model. It is expected to represent the process behaviour on a wide range of operating conditions with good accuracy. It should not only guarantee that the predicted potential profitability matches that of the real process, but also that when the optimal solution is implemented the process constraints must not be violated. Most RTO systems used nowadays are based on rigorous, steady-state, first-principles mathematical models.

The good performance of an RTO system depends on a reliable mathematical model and on reliable input data. In order to obtain that, many procedures must be executed before the economic optimization problem can be solved:

- Gross Error Detection
- Steady-state Detection
- Data Reconciliation
- Parameter estimation

Once that a reconciled data set and a fitted model have been obtained, the process optimization can be performed. The optimization problem usually consists of the maximization the operational profit (or minimization of operational costs) subject to a set of constraints. On most situations the optimization problem is posed as a non-linear programming problem (NLP). Most commercial applications are based on variations of the SQP (Successive Quadratic Programming) algorithm. This algorithm is also used to solve the previous Data Reconciliation and Parameter Estimation problems.

Real Time Optimization at PETROBRAS

Since 2004, RTO has been classified by PETROBRAS and its Strategic Downstream Committee as a "High Sustainable" technology. It means that RTO is seen as a key technology to improve PETROBRAS performance and profit, and therefore significant effort and resources will be spent on this subject.

PETROBRAS implementations on RTO covered a wide range of alternatives, focusing both on profitability and on the search of the best way to deliver the technology:

- Fluid Catalytic Cracking (FCC) and Crude Distillation Units (CDU);
- Proprietary and commercial process models and RTO systems;
- Sequential Modular (SM) and Equation Oriented (EO) approaches (Alkaya et al., 2003).

The first RTO initiatives were taken using PETROBRAS' in-house process simulator for FCC, with a small scope covering only the reactor/regenerator section. The proprietary process model used is based on a Sequential Modular (SM) approach. Though many difficulties were found (see next section), this initiative made possible to test the technology as well as to help our engineers to take a step further.

Distillation Unit / SM approach (2004)

This implementation took place at the Crude Distillation Unit (CDU) and the two Solvents Units of RECAP refinery (Gomes et al., 2008).

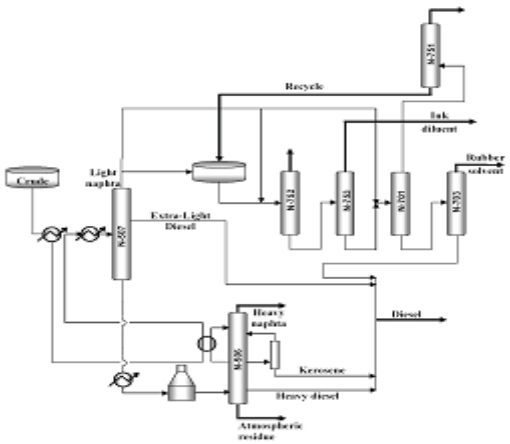


Fig. 5 - Scheme of the CDU and the Solvents Units of RECAP/PETROBRAS.

The process model was built using PETROX, a proprietary sequential-modular process simulator from PETROBRAS. The simulation comprises 53 components and pseudo-components and 64 unit operation modules, including the 7 distillation columns and a recycle stream. All modules are built with rigorous, first-principles models.

For optimization applications, PETROX was linked to NPSOL, an SQP optimization algorithm. Procedures for Steady-state and Gross error detection, Data Reconciliation, Parameter Estimation and Economic Optimization were implemented. The economic optimization problem consisted of the maximization of the operational profit, constrained by limits related to product specifications, safety constraints, feed rate and performance parameters. The whole optimization problem involves 19 decision variables and 21 constraints.

Most of the reported problems of optimization based on sequential-modular models were observed in this application:

- Low computational efficiency, due to slow recycle loops and the numerical derivatives that imply running the SM model several times. These derivatives are also inaccurate, which slows down the optimization process even more.
- Lack of reliability: the SM model is computed many times and must converge always. If a single failure happens during the optimization, all the effort is lost.

In order to minimize these problems, a lot of effort must be spent on the conception, customization and tuning of the SM model. However, that is no guarantee of success. When the Data Reconciliation and Parameter Estimation problems were implemented, the same problems were observed.

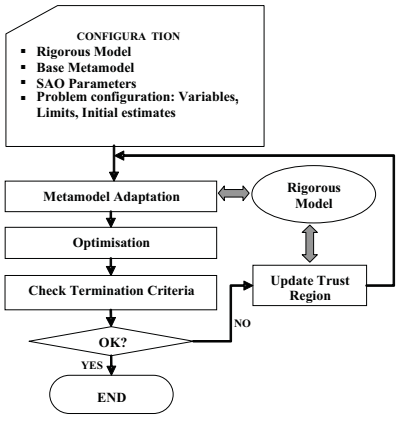


Fig. 6 - SAO strategy applied to the metamodel-based optimisation.

Metamodel approach

In order to overcome some of these shortcomings, a metamodel approach has been studied. Metamodels or surrogate models (Gomes et al., 2008) are reduced models whose parameters are obtained with data that is generated with rigorous, first principles models. In this work, an optimization procedure was developed, combining

metamodels and rigorous models with a Sequential approximate optimization (SAO) algorithm. The optimization problem is solved based on the metamodel that is updated with data obtained from the rigorous model throughout the optimization procedure. The RECAP optimization problem was addressed with this approach, with kriging models and neural nets used as metamodels. Accurate results have been obtained with considerable reduction of the computational effort on most of the studied cases.

Distillation Unit / EO (2005 to 2006)

This was the first EO RTO project PETROBRAS implemented. After an International Bid, where 3 well-known companies were invited to submit their proposals, AspenPlus Optimizer (Aspentech, Inc.) was selected. The project scope included all 3 preheat trains as well as Pre-flash, Naphtha Stabilizer, Atmospheric, Vacuum and Pre-vacuum distillations towers. The unit was fully modeled with the RTO software, which allowed for instance the understanding about the implications that changes on the preheat train, like feed distribution, have on the Atmospheric tower. Or to study the best pumparound heat removal distribution along this tower and its effects on the preheat train. In order to do that, all pumparounds were modeled as external streams from the tower and not as an internal model within its model (see Figure 7), as it is common on SM simulators.

The system is running on open loop since 2007. A few closed-loop tests were performed, but the unit had some operational problems which were solved on this last Oct/08 turnaround. PETROBRAS intends to close loop in 2009 after making model tuning adjustments in order to incorporate the new atmospheric trays and other unit improvements. Nevertheless, by keeping the system running open loop (around 9 runs / day), we were able to improve our knowledge of the system itself, how to overcome non convergence problems (feed reconciliation and optimization) and attaining expertise on how to maintain such a real time, strongly data and instrumentation dependent system as well as evaluate potential benefits (around 13 000,00 dollars / day).

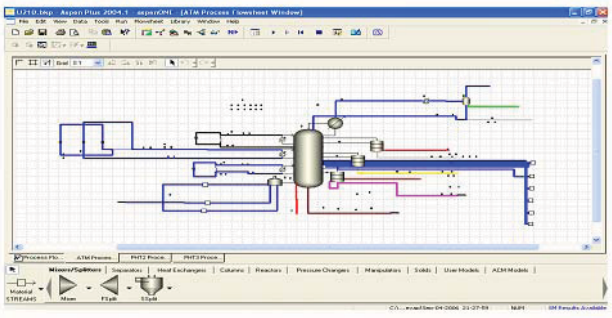


Fig. 7 - Aspen Plus Optimizer Screenshot - Atmospheric tower.

FCC Unit / EO (2007 to 2008)

Following the success on the distillation unit implementation, PETROBRAS moved forward to implement an RTO on another very important unit. Again, after an international bid, ROMeo (Invensys, Inc.) was selected.. The project scope included the Reactor / Regenerator section, Main Fractionator and Gas Recovery Plant. Again the unit was fully energy and mass integrated modeled.

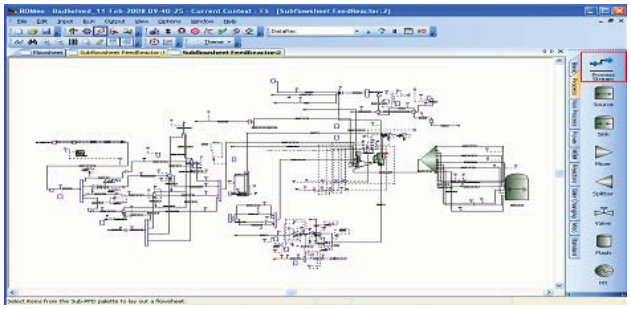


Fig. 8 - ROMeo screenshot - Reactor/Regenerator Section.

The system is running on closed loop (around 8 runs / day) since June/08 with most of the independent variables active. On average, around 60% of the successful runs are being accepted by Operations and targets are being sent to Advanced Control. PETROBRAS has evaluated an average gain of US$ 0.12 / bbl of FCC feed for this application, by comparing the unit performance with and without RTO.

A few comments on both projects:

- Lack of instrumentation on preheat train (FCC) – implied on simplifications, which has impacts on Main Fractionator heat balance and, thus, must be evaluated from time to time;
- Low feed lab analysis frequency – There is a need for a better way to estimate feed characterization;
- Non-convergence problems - Mainly, due to instrumentation faulty and/or out of service heat exchanger or other piece of equipment. Although there is a kind of standard procedure to deal with them, it is not possible to automate it. So each problem must be solved on a case to case, hands-on basis.

These facts enforce the need for a fully dedicated RTO engineer for each application, not only to assess its results and make sure they are being implemented, but to keep the system running despite of the many daily issues the application faces.

Modelling approach

PETROBRAS experiences showed that the Equation Oriented (EO) approach is more suitable for RTO, when compared to the Sequential-modular process models, especially when process unities of higher complexity are addressed.

Challenges associated with RTO

Non-convergence tracking

When the optimization process brakes down due to non-convergence, it is sometimes a hard task to find out the origin of the failure, especially when the cause of the problem is not

related to instrumentation or well-known process problems. Therefore, there is a need for better procedures or even an expert system that might identify the numerical failures and provide high-level analysis to support the user on the best actions to take.

The improvement of the initialization techniques (Fang et al., 2009) might also be useful to avoid convergence problems, especially for the data reconciliation problem.

Scaling

Scaling of variables is a subjective issue. Despite the available heuristic rules provided by the technology licensors, the users are sometimes required to define scaling factors or limits. However, it is possible that a numerical analysis of the system of equations to be solved might provide the best scaling factors.

Integrating multiple process unities

In order to take the most of process flexibilities, it might be important to expand the scope of the optimization problem to involve more than just one process unit. However, the increase of the problem size and the consequent shortcomings can be a challenge to be faced. In this case, the non-converge tracking procedures would become a key issue.

Steady-State detection

The steady-state detection procedures used nowadays in the commercial solutions require the definition of several parameters, which is a very subjective issue. This task demands from the user not only process experience, but also a long time of observation. It would be useful to have procedures that could drive a straightforward choice, especially when dealing with multiple-process optimization applications.

Multi-scale optimization

The integration and information exchange between different optimization levels is an issue that requires more attention.

Multi-level optimization concepts could be applied in order that procedures for model re-fitting or tuning and the redefinition of search spaces could be done automatically, while the different optimization problems are being solved.

Dynamic RTO

Dynamic Real Time Optimization (DRTO) is an open issue. The use of rigorous dynamic models for large-scale applications might allow the simultaneous solution of process optimization and control problems. Ideally it would also avoid the requirement of steady-state detection procedures. However, with the present resources, DRTO solutions would demand a significant computational effort and, possibly, many numerical issues should be addressed before this technology can be widely used in industrial applications.

5. CONCLUSIONS

This article has discussed some challenges associated with advanced process control and optimization in petroleum industries as well as how PETROBRAS is overcoming them. Our vision is that there is still plenty of space for further nd research and development on the improvement of those technologies. The best accomplishment of this task will come if Industry and Academy work together.

REFERENCES

Alkaya, D. et al., Generalization of a Tailored Approach for Process Optimization, Ind. Eng. Chem. Res., Vol. 39, pp. 1731-1742, 2003.

Arruda, G.H.M. and Barros, P.R., Relay based gain and phase margins PI controller design, IEEE Transactions on Instrumentation and Measurement Technology, Vol. 52, n. 05, pp. 1548-1553, 2003

Bequette, B. W, Non-linear Model Predictive Control: A Personal Retrospective, The Canadian Journal of Chemical Engineering, Vol. 85, pp. 408-415, 2007.

Campos, M. and Teixeira, H., Controles típicos de equipamentos e processos industriais, Ed. Edgard Blücher, São Paulo, 2006.

Campos, M. et al., Ganhos econômicos devidos à melhoria no controle de uma planta de processamento de gás natural, IV Congresso Rio Automação 2007, IBP, Rio de Janeiro, 2007.

Cutler, C. R. and Perry, R. T., Real-time optimization with multivariable control is required to maximize profits, Computers and Chemical Engineering, Vol. 7 (5), pp. 663-667, 1983.

Ender, D., Process Control Performance: Not as good as you think, Control Eng., pp. 180, 1993.

Fang, X. et al., Mnemonic Enhancement Optimization (MEO) for Real-Time Optimization of Industrial Processes, Ind. Eng. Chem. Res.,Vol. 48, pp. 499-509, 2009.

Farenzena, M. and Trierweiler, J., Fronteiras e desafios em gerenciamento de malhas de controle, In: COBEQ 2008 - Congresso Brasileiro de Engenharia Química, Recife, 2008.

Farenzena, M. et al., Using the Inference Model Approach to Quantify the Loop Performance and Robustness, SICOP 2006 - International Workshop on Solving Industrial Control and Optimization Problems, Gramado, 2006.

Gomes, M.V.C. et al.; Using kriging models for real-time process optimisation. Proceedings of the 18th European Symposium on Computer Aided Process Engineering, pp. 361-366, 2008.

Harris, Assessment of Control Loop Performance, The Can. J. of Chemical Engineering, Vol. 67, pp. 856-861, 1989.

Kempf, A., Avaliação de Desempenho de Malhas de Controle, Dissertação de Mestrado, Departamento Eng. Química, Universidade Federal do Rio Grande do Sul, UFRGS, 2003.

Kern, G., Summiting with multivariable predictive control, Hydrocarbon Processing, 2007

Ogata, K., Engenharia de Controle Moderno, Ed. Prentice/Hall do Brasil, 1982.

OPC Foundation, 2008, Site: <http://www.opcfoundation.org/>, Accessed in 03/17/08.

Qin, S. and Badgwell, T., A survey of industrial model predictive control technology, Control Engineering Practice 11, pp. 733-764, 2003.

Schmidt et al., BR-Tuning Ferramenta para sintonia de controladores PID, Primeiro CICAP – Congresso de Instrumentação, Controle e Automação da PETROBRAS, May, Rio de Janeiro, 2008.

Trierweiler, J. and Farina, L., RPN tuning strategy for model predictive control, Journal of Process Control, Oxford-Inglaterra - Elsevier, Vol. 13, pp. 591-598, 2003.

Zanin, A. and Moro, L., Gestão da Automação Industrial no Refino, Rio Oil&Gás 2004, IBP, Rio de Janeiro, 2004.

Real-time Embedded Convex Optimization

Stephen P. Boyd

Electrical Engineering Department,
Stanford University, 94305 Stanford, CA, USA
E-mail: boyd@stanford.edu.

Abstract: This talk concerns the use of convex optimization, embedded as part of a larger system that executes automatically with newly arriving data or changing conditions, in areas such as automatic control, signal processing, real-time estimation, real-time resource allocation and decision making, and fast automated trading. Such systems are already in use in applications such as model predictive control or supply chain optimization, with sample times measured in minutes (or longer); our focus is on systems with much faster dynamics, with execution times measured in milliseconds or microseconds for small and medium size problems. We describe a preliminary implementation of an automatic code generation system, which scans a description of the problem family and performs much of the analysis and optimization of the algorithm, such as choosing variable orderings used with sparse factorizations, at code generation time; compiling the generated source code yields an extremely efficient custom solver for the problem family.

The State of the Art in Advanced Chemical Process Control in Japan

Manabu Kano * **Morimasa Ogawa** **

* *Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan (e-mail: manabu@cheme.kyoto-u.ac.jp)*
** *Yamatake Corporation, Kurashiki, Okayama 712-8061, Japan (e-mail: ogawa-morimasa@jp.yamatake.com).*

Abstract: In this age of globalization, the realization of production innovation and highly stable operation is the chief objective of the process industry in Japan. Obviously, modern advanced control plays an important role to achieve this target; but it is emphasized here that a key to success is the maximum utilization of PID control and conventional advanced control. This paper surveys how the three central pillars of process control – PID control, conventional advanced control, and linear/nonlinear model predictive control – have been used and how they have contributed toward increasing productivity. In addition to introducing eminently practical methods, emerging methods, and their applications, the authors point out challenging problems. In Japan, industry and academia are working in close cooperation to share their important problems and develop new technologies for solving them. Several methods introduced in this paper are results of such industry-academia collaboration among engineers and researchers in various companies and universities. Furthermore, soft-sensor or virtual sensor design is treated with emphasis on its maintenance, because soft-sensors must cope with changes in process characteristics for their continuous utilization. Maintenance is a key issue not only for soft-sensors but also for controllers. Finally, we will expand our scope and briefly introduce recent activities in tracking simulation and alarm management. A part of the results of our recent questionnaire survey of process control are also introduced; the results are extremely helpful in clarifying the state of the art in process control in Japan.

Keywords: Advanced process control, Alarm management, Industrial application, Model-based control, Model predictive control, PID control, Process control, Production innovation, Soft-sensor, Tracking simulator.

1. INTRODUCTION

The Japanese chemical and petroleum refining industries has focused on *production innovation* and *highly stable operation*. The embodiment of these two concepts is believed to be indispensable. In fact, production innovation and highly stable operation have led to remarkably increased productivity at advanced chemical companies. Daicel Chemical Industries, for example, has tripled the productivity per plant employee since *Intellectual and Integrated Production System* was established in the Aboshi plant in 2000 (Daicel Chemical Industries, Ltd. (2008)). This reputable activity was motivated by the effort in Mitsubishi Chemical Corporation (MCC) in the 1990's (Shoda (1998)). MCC has developed *Super-stable Operation Technologies* (SSOTs) and *Super-stable Maintenance Technologies* (SSMTs) to maintain production stability and prevent facility accidents (Mitsubishi Chemical Corporation (2005)). SSOTs aim to keep stable plant operation by prevention and prediction of various troubles such as fouling, plugging, corrosion, and so on, and SSMTs are facility management technologies used to ensure high standards of stability.

In the 1990's, Japanese companies realized that many skilled operators were approaching retirement age. This social problem was called "year 2007 problem" in Japan. We are in the middle of this. Since the achievement of stable and efficient operation has largely depended on skilled operators in Japan, the year 2007 problem has heightened a sense of crisis and has motivated companies to initiate production innovation. Production innovation requires thorough review of personnel training, organizations, production methods as well as operation control systems.

To realize highly stable operation, process control plays an important role. In Japan, a task force was launched in 2007 to sift through problems regarding process control and investigate solutions. The task force, named "Workshop No.27 Process Control Technology," consists of 32 engineers from industry and 12 researchers from universities. It is supported by the 143rd committee on process systems engineering, the Japan Society for the Promotion of Science (JSPS). Currently, the following topics are being investigated by the members.

- Practical closed-loop system identification
- Practical tuning techniques of PID controllers
- Systematization of the control performance improvement activity based on control performance assessment

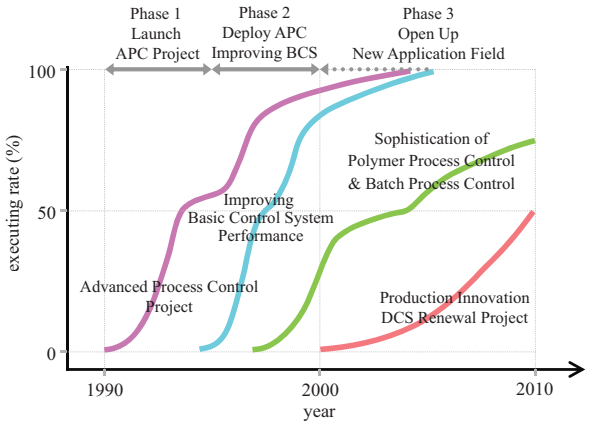


Fig. 1. Chronology of project execution in MCC

- Control system design from the viewpoint of plant-wide control
- Evaluation and maintenance of model predictive control
- Design and maintenance of soft-sensors

Most of these topics are also covered by the status report of the IFAC Coordinating Committee 6 (Dochain et al. (2008)). These are key issues not only in the Japanese chemical industry but also internationally.

This paper aims to reveal the state of the art in advanced chemical process control in Japan. First, the projects that process control sections of a general chemical corporation of Japan have executed in the last two decades are described in section 2. Then, several key technologies are investigated in more detail: PID control in section 3, conventional advanced control in section 4, model predictive control (MPC) in section 5, soft-sensor or virtual sensor in section 6, and other issues including an operation support system based on an on-line process simulator and alarm management in section 7. In each section, eminently practical techniques with successful application results are introduced, and challenges are clarified. Furthermore, this paper introduces results of a questionnaire to member companies of the JSPS 143rd committee on their process control applications including MPC and soft-sensors. The results will be extremely useful for grasping the state of the art in process control.

2. MILESTONE IN THE HISTORY OF PROCESS CONTROL APPLICATION

There are three phases in process control application projects in Mitsubishi Chemical Corporation (MCC), to which the second author had belonged for many years, as shown in Fig. 1: the advanced process control (APC) projects for large-scale continuous processes, the improvement activity of the control performance of basic control systems for small-to-medium-scale processes, and the advancement of polymer and batch process control.

2.1 Project Chronology

In the first phase in the early 1990's, multivariable MPC was applied to large-scale continuous processes such as olefin production units for generating a large profit. The APC project was conducted for 15 production units of 5 production sites by using DMCplus® as a standard tool, and satisfactory results were achieved. The key to success is nurturing process control engineers who can accomplish the projects independently on their own. They learned procedures and methods of planning, control system design, plant tests, tuning, and operation. In addition, they joined seminars on advanced control theory given by prominent researchers and professors. By accumulating experience on the projects, they grew into capable engineers who understood theory and had business acumen. These 15 process control engineers took a leading part and accomplished APC projects in MCC.

In the second phase, the performance of PID control systems was assessed and improved. All production units which APC projects did not cover were targeted. Both the operation section and the instrumentation section jointly carried out this project as a daily improvement activity in cooperation with the process control section. As a result, the operator workload was reduced through the improvement in service factors of PID control systems and a reduction in frequency of alarms and operator interventions. In addition, the improvement in control performance contributed toward the economic profit because it made operations energy-efficient through optimally changing set-points. It was also the perfect opportunity for finding applications of conventional advanced control such as override control and valve position control (VPC).

In the third phase, the advancement of polymer process control was investigated. It is important to achieve rapid grade transition while satisfying quality specification in polymer plants, because transitions among a wide variety of products are made frequently. Therefore, an original control algorithm that is based on precise first-principle models of polymerization reactions and quality models relating polymerization reaction conditions and product quality has been used since the 1980's. In this phase, process models such as catalyst activity were reviewed, and a new nonlinear MPC algorithm was developed and applied. As a result, the control performance was significantly improved, off-specification products were reduced, and quality was stabilized.

The focus of the process control section has shifted to problem-solving regarding process control of small-to-medium-scale processes and the maintenance of APC systems. The targets include 1) accumulating energy-saving effects by applying an in-house linear MPC algorithm to distillation, reforming furnace, and air separation processes, 2) developing soft-sensors, which are substituted for process gas chromatographs, for shortening the control interval and improving control performance, and 3) adapting APC systems for reinforcement of process units.

Since the 1990's, the movement to reform the whole production activity has started at advanced chemical companies as mentioned in the introduction. In addition to integration of control rooms, such production innovation requires the review of operation management, alarm management, emergency shutdown system, maintenance management, etc., and also it requires modernizing the control information system. Such an activity is triggered by the

Table 1. Classification of process control methodologies and the numbers of applications in the MCC Mizushima plant

classification	methodology	application
modern	linear MPC	54
advanced	nonlinear MPC	2
control	LQI with preview action	2
conventional advanced control	feed-forward control override control valve position control analyzer feedback control model-based control etc	500+
regulatory control	PID/I-PD control	5006

opportunity for DCS introduced in the 1980's to enter a renewal period as well as the year 2007 problem. Process control engineers are or will be involved in this movement.

2.2 Process Control Methodology

Control methodologies which bear the central role in process control systems can be classified into regulatory control such as PID control, conventional advanced control such as feedforward control and override control, and modern advanced control such as MPC. The number of applications of these control methodologies in the MCC Mizushima plant is summarized in Table 1. The ratio of applications of PID control, conventional advanced control, and MPC is 100:10:1. PID control is used in 5006 loops in 24 production units. The number of control loops repeatedly increases and decreases corresponding to new establishment, reinforcement, or stopping of production units. Conventional advanced control is effective in many cases, but the number of its applications is not as many as expected. MPC has become established as a standard technique for multivariable control which realizes economical operation of large-scale processes.

2.3 Survey Result of Control Methodology

A part of the questionnaire survey results of process control application is summarized in Table 2. This questionnaire asked control engineers to evaluate the level of their application of conventional advanced control, model-based control, adaptive control, modern-control-theory-based control, knowledge-based control, statistical process control, and soft-sensor in four grades.

This survey result clarifies the state of the art of process control application in Japan. As expected, linear MPC is the only methodology of modern advanced control that has been applied practically. Most companies have not used nonlinear MPC, adaptive control including self-tuning control, state feedback control, preview control, H_∞ control, or knowledge-based control including neural-network-based control. These control techniques have not been used because they are not available as a practical, easy-to-use tool and in-house development is troublesome. In particular, self-tuning control is a black box and has incurred a vague distrust of engineers and operators. In addition, it is not superior to gain scheduling control or robust PID tuning, which is more intuitive and understandable. On the other hand, the modern control theory has not been accepted in the chemical and petroleum refining industries. This situation is in stark contrast to that in the steel industry, for example, where there are many applications of modern control such as H_∞ control. This is because there have already been a number of successful MPC applications in the chemical and petroleum refining industries; thus control engineers are not motivated to use more theoretical control algorithms. Knowledge-based control is useful for complementing PID control and MPC, but it is difficult to generalize knowledge-based control so that it can be applied to a variety of processes.

Table 2. Level of control application (from the survey JSPS143 WS27 2009)

control methodology	level of application			
	A	B	C	D
conventional advanced control				
feedforward control	3	9	6	2
override control	2	6	5	7
valve position control	4	5	6	5
sampled-data control	1	5	9	5
dead-time compensation	0	2	11	7
gain-scheduled PID control	1	1	9	9
model-based control				
internal model control	2	5	3	9
linear model predictive control	4	6	6	3
nonlinear model predictive control	0	1	2	16
adaptive control				
self-tuning PID control	0	1	1	17
model reference adaptive control	0	0	1	18
modern-control-theory-based control				
state feedback control	0	0	4	15
preview control	0	0	1	18
H_∞ control	0	0	0	19
knowledge-based control				
fuzzy control	0	0	5	14
artificial-intelligence-based control	0	0	2	17
neural-network-based control	0	0	4	15
statistical process control	0	1	3	15
soft-sensor	3	7	4	5

Explanation of level of application:
A: standardized and always applied if necessary.
B: applied, but not standardized.
C: applied sometimes.
D: not applied.
The numbers in this table show the numbers of answers.

3. PID CONTROL

In Japanese chemical companies, KAIZEN activities aimed at safe and stable operation are actively continuing. One important activity is improvement in the control performance of PID control systems. The aims of this improvement activity, in which controllers are retuned appropriately, are 1) to realize stable operation by reducing the influence of disturbances, 2) to realize automatic rapid transition of operating conditions such as production rate, 3) to gain the ability to achieve economical operation, and 4) to allow operators to be released from taking care of PID controllers. Additional effects are to find out problems with sensors and actuators, and to clarify possible targets of advanced control application.

In the KAIZEN activities, improving the control performance with retuning should be stressed, rather than spending time and effort to strictly assess the control performance of PID control loops. The following simple indexes are sufficient to determine good or bad control performance: 1) Is the controller in auto mode at all

times? 2) Are PID parameters in the proper range? 3) Is fluctuation of the controlled variable and the manipulated variable sufficiently small? 4) Is the PID tuning agreeable to the control purpose such as flow-averaging level control? Other than these, it is necessary to check the range propriety of sensors and actuators, the necessity of filtering of measurement noise, the presence of stiction of control valves, and so on.

Experience leads us to believe that 80% of PID control loops can be successfully tuned with a method based on rule of thumb and trial and error. For example, initial settings for PID parameters should be "wide proportional band and fast reset time" for flow control and "narrow proportional band and slow reset time" for level control. After the initial PID setting, PID parameters are tuned gradually to strengthen control action while the control performance is verified.

The control performance improvement activity introduced in this section has attracted the attention of many enterprises in the chemical and petroleum refining industries in Japan, and the number of enterprises starting this activity has increased rapidly. Such a movement seems to be the result of the process control section not directly recognizing the reality that the operation section had an awareness of control performance issues and was dissatisfied with the control performance.

3.1 Actual Project Examples

The result of a project on a large-scale monomer plant, which has 190 PID control loops, is introduced here. In this plant, 90% of the PID controllers were in auto mode for 30 days. This value outperforms the average of 70% in the literatures (Desborough and Miller (2001); Ender (1993)). Operators had adjusted PID parameters to realize very loose control action. As a result, the process was easily affected by disturbances and a long time was required for production rate changes, thus the operators made frequent adjustments such as set-point changes and manual operation.

In all, 112 loops having a margin of improvement in control performance were retuned in 12 days. The standard deviations of controlled variables (CVs), σ_e, and manipulated variables (MVs), σ_u, were reduced by an average of 37% and 28%, respectively, as shown in Fig. 2. Here $\tilde{\sigma}$ denotes the standard deviation before the retuning. The reduction is almost the same as the value reported by Shah et al. (2004). A pronounced effect was achieved in tray temperature control loops of distillation columns. Temperature fluctuation was reduced to one-fourth up to one-seventh, and composition was also stabilized.

Figure 3 shows PID parameters for 29 level control loops before and after the retuning. Here PB and T_i denote proportional band and reset time, respectively. With the exception of a part such as six loops for a heat recovery boiler, the purpose of these control loops is flow-averaging level control (FALC). Operators made the proportional gain small (wide proportional band) in order for the manipulated variable not to change. However, the manipulated variable had been oscillatory due to small reset time. To solve this problem, Ogawa et al. (1998) developed

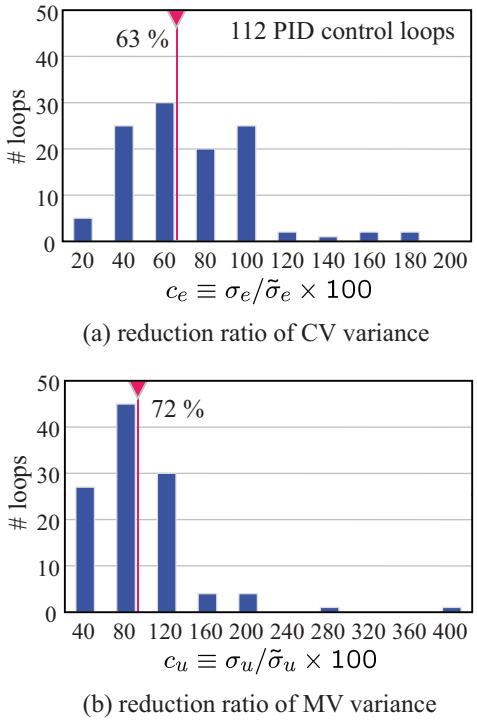


Fig. 2. Improvement in PID control performance: project on a large-scale monomer plant in MCC

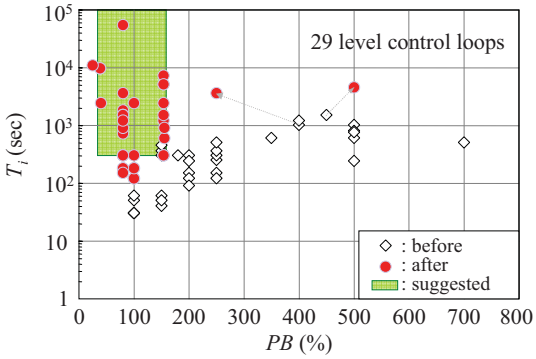


Fig. 3. PI parameters of 29 level controllers before and after retuning

a design method of flow-averaging level controllers and applied it to those loops. As a result, it became possible to suppress the oscillation of the manipulated variable by allowing the fluctuation of the level, utilizing the capacity of the drum, and absorbing flow disturbances. This FALC, explained in section 3.3, was very effective for decreasing changes in feed/product flow rate to distillation columns and lightening the burden of tray temperature control.

The above-mentioned example is the result for MCC. Generally, each company has its own in-house tool for assessing and improving PID control performance. In Sumitomo Chemical, for example, Kugemoto (2005) developed a control loop diagnostic tool "LoopDiag" that can execute control performance assessment, valve stiction detection, as well as time series data analysis. LoopDiag is a re-

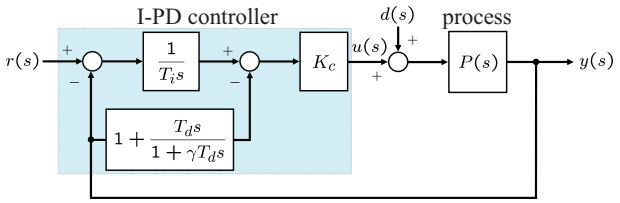


Fig. 4. I-PD control system

sult of industry-academia collaboration in the task force "Workshop No.25 Control Performance Monitoring" supported by the JSPS 143rd committee. In LoopDiag, control performance is evaluated on the basis of the minimum variance control benchmark concept (Harris (1989)), and valve stiction is detected by using the methods developed by Maruta et al. (2005) and Yamashita (2006). By the year 2005, control performance assessment was carried out for 300 PID control loops by using LoopDiag, and performance improvement was achieved. In addition, 12 valve failures were diagnosed in 118 control loops, and four of 12 valves had actually developed trouble.

Mitsui Chemicals has utilized "Plant Control Estimation & Tuning System (PCETS)" (Nishimura and Ootakara (2007)). The functions of PCETS include 1) control performance assessment based on operation data of controlled variables, set-points, and manipulated variables, 2) plant performance assessment, and 3) optimal PID tuning support. The function of control performance assessment has been applied to more than 5000 control loops, and more than 250 control loops whose performance was identified as poor were retuned by the function of optimal PID tuning support. The control performance was greatly improved in most control loops.

In Idemitsu Kosan, one-parameter tuning PID control has been used (Fujii and Yamamoto (2008)). This unique technique was developed to integrate control performance assessment and controller design and to make PID tuning easier and more intuitive for plant operators. It allows PID parameters to be tuned by adjusting just one user-specified parameter that corresponds to control strength or robustness. So far, one-parameter tuning PID control has been successfully applied to hundreds of control loops. This technique clarifies when controllers should be retuned and enables operators who do not have controller design experience to tune PID parameters effectively.

These examples would reveal the state of the art in PID control, which still plays a very important role in chemical process control. In the following part of this section, a few practical control techniques are introduced.

3.2 Robust I-PD Controller Tuning

Since most PID controllers have the I-PD algorithm at least in Japan, Ogawa and Katayama (2001) derived a robust model-based PID tuning method for the I-PD controller shown in Fig. 4. This method is suitable for specific control loops such as temperature and composition control, which are required a proper control performance in the presence of plant-model mismatch.

The advantage of I-PD control over conventional PID control is that I-PD control can realize milder response to set-point changes than PID control, while both control algorithms achieve the same performance against disturbances. When the set-point is changed stepwise in PID control systems, an abrupt change of the manipulated variable is unavoidable due to derivative and proportional actions. In practice, such an abrupt change is undesirable. On the other hand, in I-PD control systems, both derivative and proportional terms act only on the controlled variable; thus milder changes in the manipulated variable can be realized.

Here, the I-PD controller tuning method for a first-order plus time-delay (FOPTD) model is explained. The desired response $W_r(s)$ of the controlled variable y for the set-point r is specified by

$$W_r(s) \equiv \frac{y(s)}{r(s)} = \frac{1}{(1+T_F s)^n} e^{-T_L s} \tag{1}$$

where T_F denotes a tuning parameter and $n = r + 1 = 2$ for the relative order $r = 1$ of the process model. T_L, T_p, and K_p denote time-delay, time constant, and steady-state gain of the process model, respectively. By using the 1/1 Pade approximation and ignoring the derivative filter, the partial model matching method (Kitamori (1981)) provides the following PID setting rule.

$$K_c = \frac{p - 2q + 4}{K_p\,(p + 2q)} \tag{2}$$

$$T_i = \frac{(p+2q)\,(p-2q+4)}{2p+4} T_p \tag{3}$$

$$T_d = \frac{p\left(p + 4q - 2q^2\right)}{(p+2q)\,(p-2q+4)} T_p \tag{4}$$

where $p \equiv T_L/T_p$ represents the difficulty of control and $q \equiv T_F/T_p$ is a tuning parameter. Although the parameter q can be tuned so that ISE (Integral of Squared Error) is minimized, such tuning is not preferable in practice. To realize robust PID control that is intuitive and practical, a constraint on the maximum change of the manipulated variable $u(t)$ against a stepwise set-point change is introduced. Given $U_{\max}(\%)$, the parameter q is determined by solving the following equation.

$$\max_q \|u(t)/u(\infty)\|_\infty \le U_{\max}/100 \tag{5}$$

where $u(\infty)$ is the steady-state value of $u(t)$ after the set-point change. The relationship among q, p, and $U_{\max}$ is shown in Fig. 5.

This robust I-PD controller tuning method is derived not only for FOPTD models but for integral plus FOPTD models and second-order plus time-delay (SOPTD) models with/without an unstable pole.

3.3 Flow-Averaging Level Control

Consider a process described by

$$P(s) = \frac{y(s)}{u(s)} = \frac{1}{T_p s}, \quad T_p = \frac{K_m A}{K_u} \tag{6}$$

where T_p (h) denotes reset time constant, K_m (m/%) sensor gain, K_u (m^3/h/%) actuator gain, and A (m^2) sectional area.

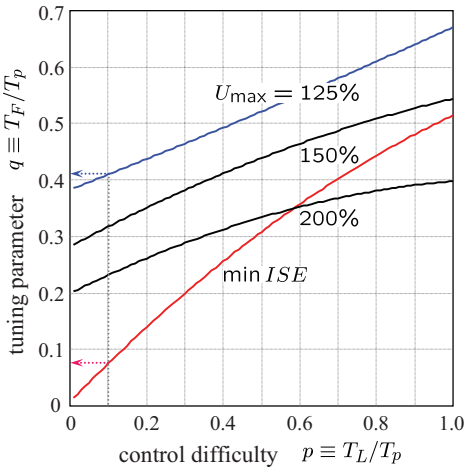


Fig. 5. Tuning of robust I-PD controllers

I-P control is used for FALC. Its block diagram is shown in Fig. 4 and derivative time T_d is set equal to 0. The control response to set-point r (%) and disturbance d (%) becomes the following second-order standard form.

$$y(s) = \frac{1}{1 + 2\zeta T_n s + T_n^2 s^2}\left(r(s) + \frac{T_i s}{K_c} d(s)\right) \tag{7}$$

The damping coefficient ζ and the natural frequency T_n are given by

$$\zeta = \sqrt{\frac{K_c T_i}{4T_p}}, \quad T_n = \sqrt{\frac{T_p T_i}{K_c}} \tag{8}$$

By defining the performance index of FALC under a step-wise disturbance as

$$\min \ J = \frac{1}{2}\int_0^\infty \left(q^2 y^2(t) + r^2 \dot{u}^2(t)\right) dt \tag{9}$$

and solving the optimization problem similar to the LQI problem, the control parameters can be related to the process parameter.

$$K_c T_i = 2T_p \tag{10}$$

As a result, the damping coefficient becomes $\zeta = 1/\sqrt{2}$ and the second-order standard form becomes Butterworth-type.

Given the size of the step-wise disturbance d_s and the maximum allowable level change y_s, the proportional gain and the reset time can be determined as follows:

$$K_c = \frac{\sqrt{2}e^{-\pi/4}}{y_s/d_s} \approx \frac{0.645}{\eta}, \quad T_i = \frac{2T_p}{K_c} \tag{11}$$

Here, $\eta \equiv y_s/d_s$ is the disturbance rejection ratio.

This tuning method has been widely used in industry to improve the performance of level control, in particular, to achieve FALC with the specified characteristics, because the calculation of PI parameters is very easy.

3.4 Direct PID Controller Tuning

Discussions with control engineers in the Japanese process industries confirm that PID controller tuning is still a key issue. A typical chemical plant has thousands of control loops whose maintenance is vital to efficient operation of the entire plant. The conventional approach to tackling this problem is to use an efficient open/closed-loop identification method and reduce the burden of modeling. However, any control system based on an identified model suffers from modeling errors and requires retuning of control parameters. In addition, identification is still one of the critical tasks in control system design. Control engineers and operators would prefer to avoid identification and manual tuning of PID controllers.

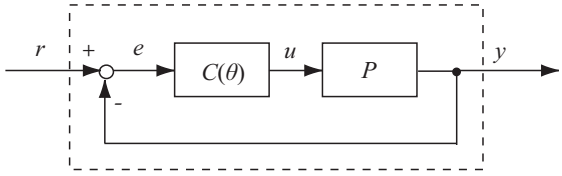


Fig. 6. Feedback control system

Extended fictitious reference iterative tuning (E-FRIT) is a new direct tuning method, which can optimize PID or I-PD control parameters directly from operation data without a process model (Tasaka et al. (2009); Kano et al. (2009b)). E-FRIT is a kind of extension of other direct tuning methods such as iterative feedback tuning (IFT) proposed by Hjalmarsson et al. (1998), virtual reference feedback tuning (VRFT) by Campi et al. (2002), and fictitious reference iterative tuning (FRIT) by Soma et al. (2004).

E-FRIT is briefly explained here. Figure 6 shows a block diagram of a feedback control system, where P denotes a process, $C(\boldsymbol{\theta})$ a controller with parameters $\boldsymbol{\theta}$, r set-point, and u and y are a manipulated variable and a controlled variable, respectively. When PID control is used,

$$C(\boldsymbol{\theta}) = K_{\mathrm{P}}\left(1 + \frac{1}{T_{\mathrm{I}} s} + T_{\mathrm{D}} s\right) \tag{12}$$

$$\boldsymbol{\theta} = (K_{\mathrm{P}}, T_{\mathrm{I}}, T_{\mathrm{D}}) \ . \tag{13}$$

In E-FRIT, a virtual output variable is formulated as a function of PID parameters by using input and output data together with a reference model. PID parameters are determined so that the difference between the real and virtual output variables is minimized. The following is the PID tuning procedure based on E-FRIT. Here, $G(s)x(t)$ or $Gx(t)$ is defined by $\mathcal{L}^{-1}\{G(s)\mathcal{L}\{x(t)\}\}$, which represents the discrete time series data collected at certain sampling intervals.

[Step 1] After the control system is stabilized with initial PID parameters $\boldsymbol{\theta}_0$, change the set-point and collect input and output data, $u_0(t)$ and $y_0(t)(t = 1, 2, \cdots, N)$.

[Step 2] Derive the fictitious reference (virtual set-point) $\tilde{r}(\boldsymbol{\theta}, t)$ that generates $u_0(t)$ and $y_0(t)$ even when $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$.

$$\tilde{r}(\boldsymbol{\theta}, t) = C(\boldsymbol{\theta})^{-1} u_0(t) + y_0(t) \tag{14}$$

[Step 3] Formulate the reference output $\tilde{y}(\boldsymbol{\theta}, t)$ by using a reference model M as shown in Fig. 7.

$$\tilde{y}(\boldsymbol{\theta}, t) = M\tilde{r}(\boldsymbol{\theta}, t) \tag{15}$$

The closed-loop system is close to the reference model when $\tilde{y}(\boldsymbol{\theta}, t)$ is close to $y_0(t)$.

[Step 4] Solve the following optimization problem and determine the optimal control parameters $\boldsymbol{\theta}^*$.

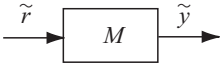


Fig. 7. Reference model of closed-loop system

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} J_{\text{ext}}(\boldsymbol{\theta}) \tag{16}$$

$$J_{\text{ext}}(\boldsymbol{\theta}) = \frac{1}{N}\sum_{t=1}^{N}\left\{(y_0(t)-\tilde{y}(\boldsymbol{\theta},t))^2+\lambda\Delta\tilde{u}(\boldsymbol{\theta},t)^2\right\} \tag{17}$$

$$\Delta\tilde{u}(\boldsymbol{\theta},t) = \tilde{u}(\boldsymbol{\theta},t)-\tilde{u}(\boldsymbol{\theta},t-1) \tag{18}$$

$$\tilde{u}(\boldsymbol{\theta},t) = C(\boldsymbol{\theta})\,(r_0(t)-Mr_0(t)) \tag{19}$$

where λ is a weighting coefficient.

A reference model plays an important role in defining the desirable control response. It is difficult, however, to determine an appropriate reference model in advance without information on the process. Therefore, parameters in the reference model are optimized together with the control parameters in E-FRIT. For example, when the reference model M is defined as the second-order binomial coefficient standard form given by

$$M = \frac{\omega_0^2}{s^2+2\omega_0 s+\omega_0^2}e^{-L_{\text{M}}s} \tag{20}$$

the optimization variables are

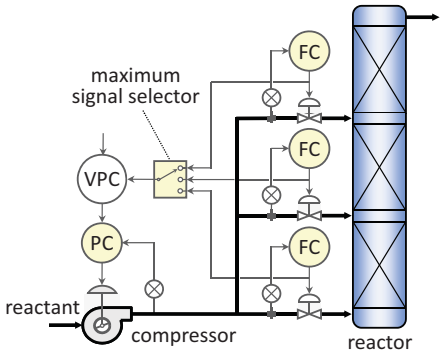
$$\boldsymbol{\phi} = (K_{\text{P}}, T_{\text{I}}, T_{\text{D}}, L_{\text{M}}) \tag{21}$$

instead of $\boldsymbol{\theta}$. This extension makes it possible to determine the reference model that is more suitable for the process.

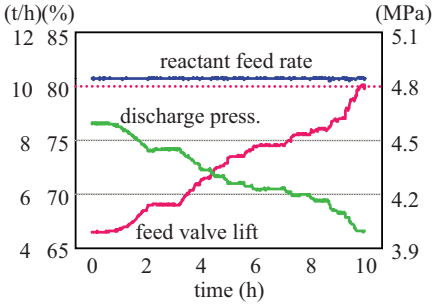
Kano et al. (2009b) proposed useful guidelines for applying E-FRIT to industrial processes: 1) use the fourth-order binomial coefficient standard form with dead time as a reference model, 2) set a parameter ω_0 of the reference model on the basis of the rise time of the closed-loop response, 3) optimize the dead time of the reference model together with control parameters, and 4) use a fixed value as a weighting coefficient λ for a penalty term for variation of the manipulated variable. A recommended value is $\lambda = 0.01$ for tight control and $\lambda = 1$ for mild control. E-FRIT with these guidelines was validated through industrial applications. The results have clearly shown the usefulness of E-FRIT for chemical process control. A software tool that can execute E-FRIT was developed as a result of industry-academia collaboration in the task force "Workshop No.27 Process Control Technology," and it has been used in industry.

4. CONVENTIONAL ADVANCED CONTROL

The status report of the IFAC Coordinating Committee 6 (Dochain et al. (2008)) stated that high performance multivariable control is key to achieving the desired high profits and that the technology for the design and realization of high performance model-based constrained control systems at reasonable engineering effort is one of the key challenges faced by industrial practice. In fact, MPC has contributed toward achieving high profitability for many years. However, the profit can also be realized by utilizing conventional advanced control such as valve position control and override control in particular. The following question arises here: do we make the most use of conventional advanced control? In this section, let us introduce one example showing the potential of conventional advanced control.



(a) control system based on VPC



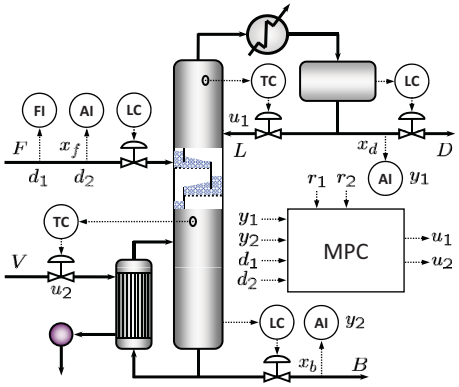
(b) control result

Fig. 8. Compressor-power-saving control

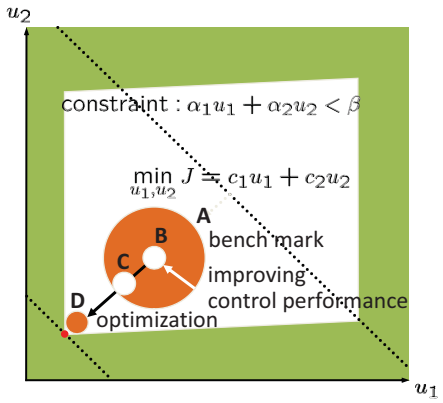
Conventional advanced control is effective for various processes and easy to implement on DCS. However, there has been a trend for control engineers to take little account of its application. This is the result that MPC became a standard tool for the advancement of process control. However, there is no doubt that production cost can be decreased by accumulating the effect of conventional advanced control.

The application of energy-saving control of the compressor with VPC is described here. As shown in Fig. 8, the feed gas is pressurized with the turbo compressor and supplied to three different stages of the reactor. Each flow rate of the feed gas is controlled. The discharge pressure of the compressor is controlled by using guide vane opening as the manipulated variable. To reduce compressor power, the discharge pressure is lowered gradually with VPC, until the largest valve opening among three feed flow control valves reaches the upper limit, while feed flow rate is kept constant. In this application, the discharge pressure was decreased from 4.6 MPa to 4.0 MPa by increasing the largest valve opening from 67% to 80%. As a result, motor electric power consumption was saved by 16%.

Shinsky (1977) listed the following objects in which there is an opportunity of the energy conservation by applying conventional advanced control: 1) excessive reflux of distillation column, 2) excessive combustion air of furnace, 3) high steam to oil ratio of reactor, and 4) fouled heat exchanger. In addition to these, excessive compression



(a) MPC for distillation column



(b) optimal operation using MPC

Fig. 9. MPC for distillation process

ratio provides an opportunity for the energy conservation as illustrated.

In the enterprise, it is important to find any loss that usually has been overlooked, to make the most use of conventional advance control, and to continue the effort at minimizing the loss. In comparison with the APC project of using MPC, profitable results can be obtained much more quickly without any further expense.

5. MODEL PREDICTIVE CONTROL

In this section, the present state of linear and nonlinear MPC is described through the typical applications and the survey results.

5.1 Linear MPC

The process that MPC is applied to most is distillation. A simple example of MPC for a distillation process is shown in Fig. 9(a). The controlled variables are the purity of products extracted from the column top and bottom, and the manipulated variables are the set-points of temperature PID control at the column top and bottom. The disturbance variables are flow rate and composition of feed. The constraints are upper and lower limits of the manipulated variables and the controlled variables and upper limits of changes in the manipulated variables.

The economic benefit that MPC brings is illustrated in Fig. 9(b). Since the achievable performance of PID control is limited due to interaction, which is a feature of multivariable processes, it is assumed that the current operating region corresponds to region A in the figure. In such a situation, the operating condition bound has to be set far from the real constraints to ensure a sufficient margin of safety. Using MPC can improve control performance and reduce variation. As a result, the operating region becomes small from A to B. This improvement makes it possible to move the operating region from B to C, which is close to the bound of operating conditions. Furthermore, more economical operation D can be realized by optimizing set-points to minimize operational costs. MPC takes on the responsibility of this set of functions. The benefit is not only the improvement of the control performance by using model-based control, but also the realization of stable operation close to the optimal point under disturbances by using optimization.

Implementation of MPC releases operators from most of the adjustment work they had to do in the past because the optimal operating condition is automatically determined and maintained under disturbances. In addition, MPC makes it possible to maximize production rate by making the most use of the capability of the process and to minimize cost through energy conservation by moving the operating condition toward the control limit. Both the energy conservation and the productive capacity were improved by an average of 3 to 5% as the result of APC projects centered on MPC at MCC.

The control performance of MPC depends on the accuracy of the process model and the appropriateness of tuning, but MPC has outstanding robustness. For example, stable operation is realized by MPC in spite of large model parameter errors of about 50%. However, it is difficult to assess the control performance of MPC due to a large number of variables. A plant test for modeling sometimes requires two weeks. The engineers who have experienced it can readily understand that the implementation of MPC including modeling and tuning is a demanding job.

MPC is highly effective, but it has several weak points (Hugo (2000)). First, it is not good at level control when the process has an integrator. For such a case, PI control is easy to design and superior to MPC in control performance. Second, the control performance of MPC deteriorates against ramp-wise disturbances because the MPC algorithm is developed by assuming step-wise disturbances (Lundstrom et al. (1995); Hugo (2000)). In addition, linear programming (LP) is usually used for optimizing set-points under constraints, and the optimal point is located at one of the extreme points of a polyhedron consisting of linear constraints. When the gradient of the objective function and that of constraints are similar to each other, the optimal point jumps from one extreme point to another and the set-points change suddenly (Forbes and Marlin (1994); Hugo (2000)). Research and development are continuing to solve these problems.

Ohshima et al. (1995), who wrote about the state of MPC application in the petroleum and chemical enterprises in Japan, reported that 154 MPC controllers were in operation and 43 under implementation. The total number

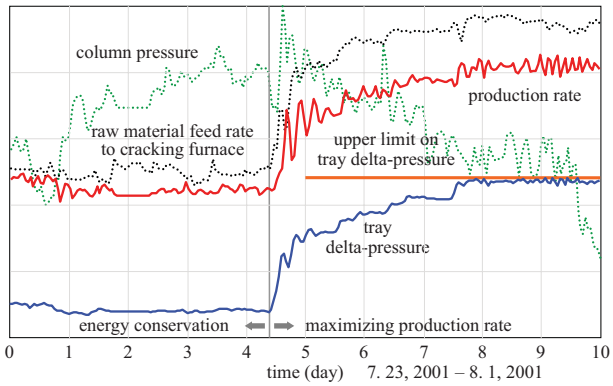


Fig. 10. Actual performance of the large scale MPC in the olefins unit

of 197 was 2.5 times as much as the number of 75 in 1990. At present, the number of MPC controllers is 169 only at MCC (Ogawa (2006)).

At the very end of this subsection, the MPC application for energy conservation and production maximization of the olefins unit at MCC Mizushima plant is briefly explained (Emoto et al. (1994)). Qin and Badgwell (2003) reported that this application was the largest MPC application in the world, consisting of 283 manipulated variables and 603 controlled variables. The process was operated in energy conservation mode for the first four days in Fig. 10. Since the productive capacity was beyond the demand, the temperature difference between vapor and coolant in the overhead condenser was increased by making the column pressure higher. As a result, an amount of heat exchanged was increased, and the amount of coolant used was decreased. This operation made it possible to reduce the refrigerator power. On the other hand, the process was operated in production maximization mode for the last five days. To maximize the production rate for fulfilling the demand, the separation performance was improved by decreasing the column pressure and increasing the relative volatility. The feed flow rate to the cracking furnace was increased until the tray delta-pressure reached its upper limit, that is, the flooding limit. In this production maximization mode, the MPC system is large because MPC controllers for many cracking furnaces and distillation columns function in cooperation.

A skilled operator made the following comment on this MPC application: "We had operated the Ethylene fractionator in constant pressure mode for more than 20 years. I was speechless with surprise that we had made an enormous loss for many years, when I watched the MPC decreased the column pressure, improved the distillation efficiency, and maximized the production rate." Another process control engineer said "I had misunderstood that set-points were determined by operation section and process control section took the responsibility only for control. I realized MPC for the first time; it makes the most use of the capability of equipments, determines set-points for economical operation, and maintains both controlled variables and manipulated variables close to the set-points."

Table 3. Statistics of MPC applications (from the survey JSPS143 WS27 2009)

in-house vs vendor			
in-house development			6 %
introduction from vendor			94 %
targeted process			
distillation			32 %
reaction			23 %
others			45 %
product			
DMCplus®			46 %
RMPCT®			36 %
Connoisseur®			5 %
SMOC®			4 %
others			9 %
number of MV, CV, and DV	MV	DV	CV
0	0	28	0
1	40	45	24
2	57	50	33
3-5	83	103	58
6-9	47	40	59
10-19	59	27	48
20-29	12	5	25
30-39	1	3	29
40-49	1	3	16
50 or more	5	1	13

MV: manipulated variable
CV: controlled variable
DV: disturbance variable

5.2 Nonlinear MPC

Nonlinear MPC has attracted attention in recent years (Qin and Badgwell (2003)). It is suitable for control of a nonlinear process operated in a wide range, e.g. polymerization reaction processes. In MCC, an independently developed nonlinear MPC has been applied to polymerization reactors at the polyolefin production units, and it has been put successfully to practical use (Seki et al. (2001)).

However, application of nonlinear MPC has not spread as well as was expected. It is difficult to build a nonlinear model of a process, or process control engineers have slackened their efforts at modeling nonlinear processes. On the other hand, most polymer production processes are operated without any quality problem by existing control systems supported with operators' suitable manual intervention. Therefore, it is difficult to justify any benefit of using nonlinear MPC. These obstacles should be overcome to expand nonlinear MPC application.

5.3 Survey Result of MPC

A part of the questionnaire survey results, related to MPC, is introduced here. The total number of MPC applications answered is 305, which is 1.5 times as much as the number of 197 in 1995. The statistics of 305 MPC applications are summarized in Table 3. Most of them are introduced from vendors; DMCplus® and RMPCT® are dominant tools. Distillation and reaction processes cover half the applications.

Table 4 clarifies objectives and effects of MPC. In addition to disturbance rejection and set-point tracking, the time to achieve the optimal condition and the realization of

Table 4. Effects of MPC applications (from the survey JSPS143 WS27 2009)

objective of tuning	
disturbance rejection	56 %
set-point tracking	38 %
time to optimal condition	6 %
major effect on control performance	
disturbance rejection	43 %
automatic operation	36 %
set-point tracking	18 %
others	3 %
major effect on productivity	
saving resources and energy	38 %
increasing production capacity	31 %
reducing operators' load	17 %
improving product quality	10 %
increasing flexibility toward changes	4 %
major key to success	
careful modeling	37 %
suitability for objective	33 %
education of operators and engineers	15 %
suitability for process characteristics	11 %
hardware/software environment	4 %

automatic operation are important. Saving resources and energy, increasing production capacity, reducing operators' load, and improving product quality are major effects achieved by MPC. Furthermore, process control engineers have identified the following major keys to success: 1) a process model should be developed with care, 2) MPC should be suitable for objectives, 3) operators and engineers should be adequately educated, and 4) MPC should be suitable for process characteristics.

Although MPC has been widely and successfully applied in the chemical and petroleum refining industries, problems still remain to be solved as summarized in Table 5. The major problem would be described as follows. To achieve desirable performance, it is necessary to build an accurate model and to tune control parameters appropriately. However, both of them are difficult in practice due to process nonlinearity and changes in process characteristics. To keep sufficient control performance and to prevent or at least cope with performance deterioration, the maintenance of MPC is crucial. Control engineers need to know the reason of performance deterioration and the effective countermeasure. In addition, they would like to know the relationship between model accuracy and achievable control performance. Modeling of a multivariable process is an exceedingly laborious engineering task; thus it needs to be clarified how accurate a model should be to achieve the goal. Of course, not only clarifying the relationship but also improving modeling and tuning methods is necessary. In addition, the implementation of MPC should be easier. As for the maintenance of MPC, very recently, Badwe et al. (2008) proposed a model-plant mismatch detection method by using partial correlation analysis, and Huang (2008) proposed the used of Bayesian methods. Another problem is how to transfer engineering technology from skilled engineers to others. Unfortunately, a lack of process control engineers aggravates the situation. Furthermore, it is also crucial in practice to answer the question: how can we estimate the economical benefit of installing MPC to justify the project? Most APC suppliers and users are required to report the benefit to management. Bauer and

Table 5. Problems of MPC applications (from the survey JSPS143 WS27 2009)

problem: general	
low robustness against model error	26 %
difficulty in tuning	23 %
inability to cope with specific objective	15 %
difficulty in modeling	12 %
others	24 %
problem: maintenance	
transfer of engineering technology	44 %
response to performance deterioration	33 %
education of operators	7 %
difficulty in tuning	7 %
others	9 %
need for improvement: general	
to improve modeling technology	28 %
to clarify method of estimating effect	25 %
to simplify implementation	22 %
to increase process control engineers	14 %
others	11 %
need for improvement: theory	
to cope with changes in process characteristics	26 %
to clarify relations between model accuracy and control performance	24 %
to cope with unsteady operation (SU/SD)	16 %
to incorporate know-how in control system	16 %
to cope with nonlinearity	13 %
others	5 %
need for improvement: response to changes/nonlinearity	
to switch multiple linear models	28 %
to improve robustness of linear MPC	25 %
to use time-varying/nonlinear model	18 %
to add adaptive function to linear MPC	18 %
to integrate other technique with MPC (e.g. knowledge-based control)	11 %

Craig (2008) reported that benefit estimation methods based on variance reduction are still carried out, but they are sometimes rudimentary and based on experience.

6. SOFT-SENSOR

A soft-sensor, or a virtual sensor, is a key technology for estimating product quality or other important variables when on-line analyzers are not available. In chemical processes, for example, soft-sensors have been widely used to estimate product quality of distillation columns, reactors, and so on. Artificial neural network (ANN) has been dominant in the literature since the middle 1990's, while partial least squares (PLS) is popular in industry (Kano and Nakagawa (2008)). ANN is a useful tool for building nonlinear models and supposed to be suitable for industrial processes. However, linear models have produced satisfactory results in many cases because industrial processes are operated within certain range to produce the required products and linear approximation functions well. In addition, collinearity has to be taken into account for developing reliable soft-sensors. Thus, PLS has been very popular as a tool for soft-sensor design (Mejdell and Skogestad (1991); Kresta et al. (1994); Kano et al. (2000)). In recent years, support vector machine (SVM), support vector regression (SVR), and other kernel-based methods have emerged (Boser et al. (1992); Cortes and Vapnik (1995)). These methods have attracted researchers' and engineers' attention and have been used for soft-sensor design (Yan et al. (2004); Desai et al. (2006)). Another method for

developing soft-sensors is subspace identification (SSID), which can build a state space model from input and output data (Verhaegen and Dewilde (1992); Overschee and Moor (1994)). SSID is a useful tool to build a dynamic inferential model of a multivariable process, and it is suitable for soft-sensor design because the performance of soft-sensors can be greatly improved by taking process dynamics into account (Kano et al. (2000)). Amirthalingam and Lee (1999) used SSID for inferential control of a continuous pulp digester. Amirthalingam et al. (2000) developed a two-step procedure to build SSID-based inferential control models, in which the stochastic part was idetified from historical data and the deterministic part was identified from plant test data. Kano et al. (2009a) proposed two-stage SSID to develop highly accurate soft-sensors that can estimate unmeasured disturbances without assumptions that the conventional Kalman filtering technique must make. Thus it can outperform the Kalman filtering technique when innovations are not Gaussian white noises or the properties of disturbances do not stay constant with time. The superiority of the two-stage SSID over conventional methods was demonstrated through their application to an industrial ethylene fractionator.

6.1 Reliability of Soft-sensor

A great deal of research has been conducted to develop data-based soft-sensors for various processes. A data-based soft-sensor, however, does not always function well, because a black-box model is not valid when a process is operated outside certain conditions where operation data used for modeling were obtained. The product quality and process performance will deteriorate if estimates of the soft-sensor are blindly believed by operators and used in a control system. On-line monitoring of the validity of the soft-sensor will avoid such a dangerous situation. The simplest approach is to check whether an estimation error exceeds its control limit when a measurement becomes available. This approach enables us to detect the inconsistency between the analyzer and the soft-sensor, but the cause of the inconsistency cannot be identified. In industry practice, it is assumed that an estimation error is caused by inaccurate estimation; however, this assumption is not always true because analyzers are not always reliable. For example, when blockage occurs within a sampling line, a hardware sensor cannot provide accurate measurements. To address such practical problems, Kamohara et al. (2004) proposed a PLS-based framework for developing a soft-sensor and monitoring its validity on-line. The on-line monitoring system was based on the multivariate statistical process control (MSPC) technique (Jackson and Mudholkar (1979); Kresta et al. (1991)) in which the dynamic PLS model designed for estimating the product quality is used. In addition, simple rules were established for checking the performance of a process gas chromatograph by combining the soft-sensor and the statistical monitoring system. The effectiveness of the developed system was demonstrated through its application to an ethylene production plant.

6.2 Changes in Process Characteristics

Generally, building a high performance soft-sensor is very laborious, since input variables and samples for model construction have to be selected carefully and parameters have to be tuned appropriately. Even if a good soft-sensor is developed successfully, its estimation performance deteriorates when process characteristics change. In chemical processes, for example, equipment characteristics are changed by catalyst deactivation or scale adhesion. Such a situation may lead to a decline of product quality. Therefore, from the practical viewpoint, maintenance of soft-sensors is very important to keep their estimation performance. Soft-sensors should be updated as the process characteristics change, and manual and repeating construction of them should be avoided due to its heavy workload.

To cope with changes in process characteristics and to update statistical models automatically, recursive methods such as recursive PLS were developed (Qin (1998)). These methods can adapt models to new operating conditions recursively. However, the prediction performance would deteriorate if the model is updated with an abnormal sample. Kaneko et al. (2009) used independent component analysis (ICA) to detect abnormal situations and improve the prediction accuracy. Recently, ICA is recognized as a useful technique for fault detection and diagnosis (Kano et al. (2003, 2004); Lee et al. (2004)). The combination between soft-sensors and fault detection is effective to a certain extent. But, as far as a recursive method is used, the model will adapt excessively and will not function in a sufficiently wide range of operating condition when a process is operated within a narrow range for a certain period of time. In addition, recursive methods cannot cope with abrupt changes in process characteristics.

Just-In-Time (JIT) modeling or lazy learning was proposed to cope with changes in process characteristics as well as nonlinearity, and it has been used for nonlinear process monitoring as well as soft-sensing (Atkeson et al. (1997); Bontempi et al. (1999)). In JIT modeling, a local model is built from past data around a query point only when an estimated value is requested. JIT modeling is useful when global modeling does not function well. However, its estimation performance is not always high because the samples used for local modeling are selected on the basis of the distance from the query point and the correlation among variables is not taken into account. A good model cannot be developed when correlation among input and output variables is weak even if the distance between samples is small. Conversely, a very accurate model can be developed when the correlation is strong even if the distance is large. On the basis of this idea, recently, correlation-based JIT (CoJIT) modeling was proposed by Fujiwara et al. (2009). In this technique, the samples used for local modeling are selected on the basis of correlation together with distance, and the Q statistic is used as an index of the correlation dissimilarity. The Q statistic is derived from principal component analysis (PCA), and it is a measure of dissimilarity between the sample and the modeling data from the viewpoint of the correlation among variables (Jackson and Mudholkar (1979)). CoJIT can cope with abrupt changes of process characteristics and also achieve high estimation performance. It can also cope with process nonlinearity.

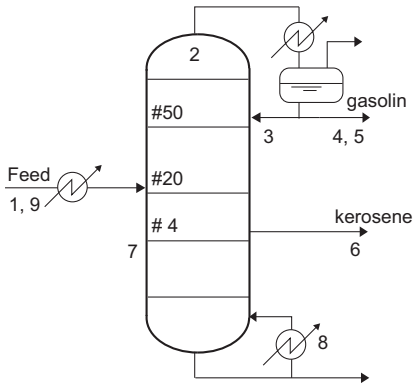


Fig. 11. Schematic diagram of the cracked gasoline fractionator of the ethylene production process at the Showa Denko K.K. (SDK) Oita plant

6.3 Industrial Case Study of CoJIT

Here, an application of CoJIT to an industrial chemical process is introduced (Fujiwara et al. (2009)). A soft-sensor for estimating the aroma concentration was constructed to realize highly efficient operation of the cracked gasoline fractionator of the ethylene production process at the Showa Denko K.K. (SDK) Oita plant in Japan. A schematic diagram of the cracked gasoline (CGL) fractionator of the ethylene production process is shown in Fig. 11. The CGL fractionator is controlled by applying multivariable MPC with an optimizer, and the aroma concentration in the CGL (aroma denotes the generic name for benzene, toluene, xylene and styrene, etc.) is used as one of the constraints in the optimizer. Although the operation data of the CGL fractionator are stored in the database every hour, the aroma concentration is analyzed in a laboratory usually once a day because of its long analysis time. For safety, the process must be operated in a condition that has a wide margin and is far from constraints. Therefore, a soft-sensor that can estimate the aroma concentration accurately in real time needs to be developed for realizing efficient operation.

In addition to eight variables measured in the CGL fractionator, the coil outlet temperature of the cracking furnace, measured four hours before, was used as an input variable, since the product composition is affected by the operating condition of the cracking furnace which is located in the upstream of the ethylene production process, and it takes about four hours for materials to reach the CGL fractionator from the cracking furnace. The selected input variables of the soft-sensor are listed in Table 6 and Fig. 11.

First, the aroma concentration was estimated with recursive PLS. The model was updated every 24 hours when the aroma concentration was analyzed in the laboratory. The estimation result is shown in Fig. 12(top). There is a bias between the measurements and the estimates after the 100th day when the pressure of the compressor was changed.

Next, the aroma concentration was estimated with CoJIT. In the initial state, the operation data obtained from April 30, 2006 to February 23, 2007 were stored in the database.

Table 6. Input variables of the soft-sensor for the CGL fractionator

No.	variable
1	Feed flow rate
2	Tower top temperature
3	Reflux volume
4	Outlet cracked gasoline temperature
5	Outlet cracked gasoline flow rate
6	Outlet cracked kerosene flow rate
7	Tray #4 differential pressure
8	Reboiler flow rate
9	Cracked furnace coil outlet temperature

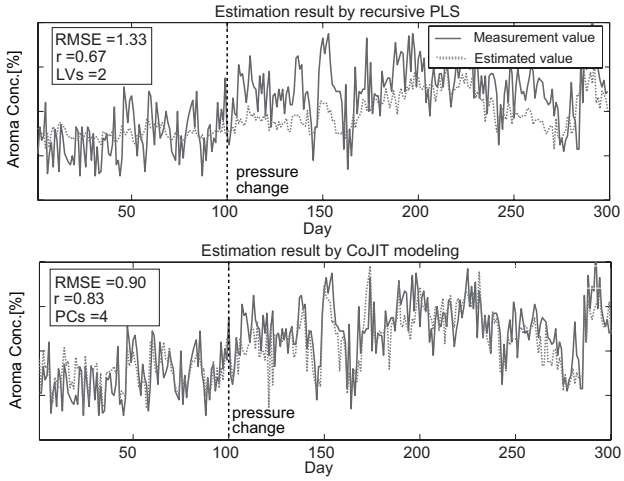


Fig. 12. Prediction results of aroma concentration: recursive PLS (top) and CoJIT modeling (bottom)

Then, the soft-sensor was updated and the aroma concentration was estimated for the next 300 days, February 24, 2007 to December 25, 2007. The estimation result is shown in Fig. 12(bottom). The estimation performance of CoJIT is high and RMSE (root mean squared error) is improved by about 28% in comparison with recursive PLS. CoJIT would have a potential for realizing efficient maintenance of soft-sensors in the real world.

6.4 Survey Result of Soft-sensor

A part of the questionnaire survey results related to soft-sensors is introduced here. This questionnaire asked control engineers the number of soft-sensor applications, targeted processes, methods for designing soft-sensors, and problems to be solved. The total number of soft-sensor applications answered was 439. The number is rapidly increasing. The survey result is summarized in Table 7.

This survey result clarifies the state of the art of soft-sensor application in Japan. First, a major targeted process is distillation (331/439), followed by reaction (86/439) and polymerization (20/439). Second, a major modeling method is multiple regression analysis (MRA) (293/439), followed by PLS (93/439). Nonlinear modeling methods are rarely used in the Japanese chemical and petroleum refining industries. It is confirmed that linear regression such as MRA and PLS can achieve sufficient estimation accuracy for most distillation and reaction processes. On the other hand, polymerization reaction processes are more difficult to model by linear regression than distillation and

Table 7. Statistics of soft-sensor applications (from the survey JSPS143 WS27 2009)

	methodology							
process	Phys	MRA	PLS	O.L.	ANN	JIT	Gray	total
distillation	20	256	41	6	0	5	3	331
reaction	5	32	43	0	0	5	1	86
polymerization	0	4	8	0	3	0	5	20
others	0	1	1	0	0	0	0	2
total	25	293	93	6	3	10	9	439

Phys: physical model
MRA: multiple regression analysis
PLS: partial least squares regression
O.L.: other linear regression
ANN: artificial neural network
JIT: just-in-time model
Gray: gray-box model or hybrid model between physical model and statistical model

Table 8. Problems of soft-sensor applications (from the survey JSPS143 WS27 2009)

accuracy deterioration due to changes in process characteristics	29 %
burden (time/cost) of data acquisition	22 %
burden of modeling itself	14 %
burden of data preprocessing	7 %
inadequate accuracy since installation	7 %
inadequate accuracy due to changes in operating conditions	7 %
difficulty in evaluating reliability	7 %
unjustifiable cost performance	7 %

other reaction processes. Thus, some companies have used gray-box models (5/20) or ANN models (3/20).

In addition, we have asked engineers what are problems related with applications of soft-sensors. The answers are summarized in Table 8. This result confirms that the maintenance of models is the most important issue concerning soft-sensors.

7. RELATED ISSUES

In this section, other important issues related to process operation are described: tracking simulator and alarm management.

7.1 Tracking Simulator

As process engineers, we have a dream that one day a plant simulator based on a rigorous first-principle model is realized and it provides functions such as 1) estimation and visualization of all states and parameters, 2) prediction of plant behavior in the future, 3) optimization of operating conditions, and 4) detection and diagnosis of abnormal situations. This plant simulation technology will become the core of future operation support system and lead to production innovation.

As mentioned before, the achievement of stable and efficient operation has largely depended on skilled operators in Japan, and many skilled operators are approaching retirement age. Thus, an advanced operation support system and an efficient operator training system are required. A training simulator for teaching operators to cope with start-up, shut-down, and other operations under abnormal situation has been developed and widely used in the process industry. The training simulator aims at faithful reproduction of real plant behavior. On the basis of the training simulator, a *tracking simulator* is now under development to realize the above-mentioned functions (Fukano et al. (2007)). The tracking simulator works simultaneously with an actual plant, adjusts parameters, estimates states, analyzes the plant, and optimizes operation by using plant models and measurements. The tracking simulator consists of a mirror model for visualizing plant states, an identification model for parameter estimation, and an analysis model for realizing the other necessary functions.

Such a tracking simulator has been developed by a few companies and introduced and tested in real plants in Japan. Further development is required to realize our dream, and various challenging problems confront us.

7.2 Alarm Management

Recently, alarm management has attracted considerable attention to achieve highly stable operation in the process industry. General recognition for current alarm systems in Japan is as follows (Higuchi et al. (2009)). With the advance of distributed control systems (DCS) in the chemical industry, it has become possible to install many alarms cheaply and easily. While most alarms help operators detect and identify faults, some are unnecessary. A poor alarm system may cause alarm floods and nuisance alarms, which reduce the ability of operators to cope with plant abnormalities because critical alarms are buried in many unnecessary alarms.

If an alarm system does not work as designed, the effects can be very serious. The explosion and fires at the Texaco Milford Haven refinery in 1994 injured 26 people and caused around £48 million of damage and a significant loss in production. The Health and Safety Executive's (HSE) investigation (1997) mentions that there were too many alarms and these were poorly prioritized and the control room displays did not help operators understand what was happening.

To improve the quality and safety of industrial plants, and to reduce cost of the design and maintenance of plant alarm systems, the Engineering Equipment and Materials Users Association (EEMUA) provided the general design and evaluation principles of plant alarm systems (The Enginnering and Equipment Materials Users' Association (EEMUA) (2007)). While this guide gathered many valuable plant engineers' experiences, it is only a general guide, and some of the design methods are only conceptual, such as the selection of alarm source signals and the decisions on alarm limits (Yan et al. (2007)). In addition, the role of operators in Japan is far different from that in other countries; thus, it is recognized that direct application of the EEMUA 191 Guide is not appropriate in Japan. In fact, a bottom-up approach has succeeded in reducing the number of alarms, average alarm frequency standards proposed by EEMUA are achieved in some plants, and further improvement is required. Generally, Japanese companies are excellent at such a bottom-up approach as TPM (total productive maintenance), which combines preven-

tive maintenance with Japanese concepts of total quality control (TQC) and total employee involvement (TEI). It is true, however, the alarm management in Japan has been short of a viewpoint of such a top-down approach as EEMUA suggested. In Japan, the industry-academia collaboration task force "Workshop No.28 Alarm Management" supported by the JSPS 143rd committee was established in 2007. This task force aims at developing new methodologies and standardizing alarm management by emphasizing distinctive culture in Japanese industries.

8. CONCLUSIONS

The state of the art in process control in Japan was described in this paper on the basis of the authors' experience and the questionnaire survey results. The realization of *production innovation* and *highly stable operation* is the chief objective of the process industry in Japan. To achieve this objective and solve the year 2007 problem, i.e., retirement of skilled operators, process control and operation need to be further improved. This improvement does not necessarily mean the adoption of novel advanced technologies. Rather, it is important to reform the whole production activity through reviewing it as leading chemical companies have done and consequently have increased productivity remarkably.

In Japan, several industry-academia collaboration task forces have been organized to sift through problems related to process operation and solve them. Such task forces include Workshop No.25 Control Performance Monitoring, Workshop No.27 Process Control Technology, Workshop No.28 Alarm Management, and so on; they are supported by the JSPS 143rd committee. More than a few methods and tools have been developed by task forces and utilized in various companies. Several examples were introduced in this paper together with practical methods developed outside task forces. The topics discussed here include PID control, advanced conventional control, model predictive control, soft-sensor, tracking simulator, and alarm management. The current situation and the problems were clarified.

In recent years, there has been a strong trend to produce polymer products having special functions in a small amount in a batch process. At the forefront of production, the necessity of practical technological development is being recognized: for example, precise control of reaction temperature, estimation of reaction state, and batch-to-batch control. Process control engineers have been committed to continuous process control so far. In the future, however, they need to open their eyes to batch process control and to meeting the challenges to its advancement.

This paper has surveyed what process control engineers have done in the last two decades and what they might do in the future, especially focusing on the projects at a Japanese chemical company. The authors expect that engineers share practical methods and best practice and also that they spare no effort in developing their own methods to solve their own problems.

ACKNOWLEDGEMENTS

The authors express their appreciation to the process control group of MCC for permitting the disclosure of many data and to the task force members of JSPS PSE 143rd committee for their constructive comments.

REFERENCES

Amirthalingam, R. and Lee, J. (1999). Subspace identification based inferential control applied to a continuous pulp digester. *Journal of Process Control*, 9, 397–406.

Amirthalingam, R., Sung, S., and Lee, J. (2000). Two-step procedure for data-based modeling for inferential control applications. *AIChE Journal*, 46, 1974–1988.

Atkeson, C., Moore, A., and Schaal, S. (1997). Locally weighted learning. *Artificial Intelligence Review*, 11, 11–73.

Badwe, A., Shah, S., Patwardhan, S., and Patwardhan, R. (2008). Model-plant mismatch detection in MPC applications using partial correlation analysis. In *IFAC World Congress*, 14926–14933. Seoul, Korea.

Bauer, M. and Craig, I. (2008). Economic assessment of advanced process control – a survey and framework. *Journal of Process Control*, 18, 2–18.

Bontempi, G., Birattari, M., and Bersini, H. (1999). Lazy learing for local modelling and control design. *International Journal of Control*, 72, 643–658.

Boser, B., Guyon, I., and Vapnik, V. (1992). A training algorithm for optimal margin classifiers. In *Annual Workshop on Computational Learning Theory*, 144–152. Pittsburgh, US.

Campi, M., Lecchini, A., and Savaresi, S. (2002). Virtual reference feedback tuning: a direct method for the design of feedback controllers. *Automatica*, 38, 1337–1346.

Cortes, C. and Vapnik, V. (1995). Support vector networks. *Machine Learning*, 20, 273–297.

Daicel Chemical Industries, Ltd. (2008). *Environmental, Safety and Social Report*.

Desai, K., Badhe, Y., Tambe, S., and Kulkarni, B. (2006). Soft-sensor development for fed-batch bioreactors using support vector regression. *Biochemical Engineering Journal*, 27(3), 225–239.

Desborough, L. and Miller, R. (2001). Increasing customer value of industrial control performance monitoring – Honeywell's experience. In *Chemical Process Control*, 153–186. Tucson, US.

Dochain, D., Marquardt, W., Won, S., Malik, O., Kinnaert, M., and Lunze, J. (2008). Monitoring and control of process and power systems: adapting to environmental challenges, increasing competitivity and changing customer and consumer demands. In *IFAC World Congress*, 7160–7171. Seoul, Korea.

Emoto, G., Ohta, Y., Matsuo, H., Ogawa, M., Raven, D., Preston, R., and Ayala, J. (1994). Integrated advanced control and closed-loop real-time optimization of an olefins plant. In *IFAC Symposium on Advanced Control of Chemical Processes (ADCHEM)*, 97–102. Kyoto, Japan.

Ender, D. (1993). Process control performance: Not good as you think. *Control Engineering*, 180–190.

Forbes, J. and Marlin, T. (1994). Model accuracy for economic optimizing controllers: the bias update case.

Industrial & Engineering Chemistry Research, 33, 1919–1929.

Fujii, K. and Yamamoto, T. (2008). One-parameter tuning pid control of chemical processes. In *International Symposium on Advanced Control of Industrial Processes (AdCONIP)*, paper MoC2. Jasper, Canada.

Fujiwara, K., Kano, M., Hasebe, S., and Takinami, A. (2009). Soft-sensor development using correlation-based just-in-time modeling. *AIChE Journal*, accepted.

Fukano, G., Onoe, Y., Nakaya, M., and Ohtani, T. (2007). Application of tracking simulator to steam reforming process. *Yokogawa Technical Report English Edition*, 43, 13–16.

Harris, T. (1989). Assessment of closed loop performance. *Canadian Journal of Chemical Engineering*, 67, 856–861.

Higuchi, F., Yamamoto, I., Takai, T., Noda, M., and Nishitani, H. (2009). Use of event correlation analysis to reduce number of alarms. In *International Symposium on Process Systems Engineering (PSE)*. Salvador, Brazil.

Hjalmarsson, H., Gevers, M., Gunnarsson, S., and Lequin, O. (1998). Iterative feedback tuning: theory and applications. *IEEE Control Systems Magazine*, 18, 26–41.

Huang, B. (2008). Bayesian methods for control loop monitoring and diagnosis. *Journal of Process Control*, 18(9), 829–838.

Hugo, A. (2000). Limitations of model predictive controllers. *Hydrocarbon Processing*, 79, 83–88.

Jackson, J. and Mudholkar, G. (1979). Control procedures for residuals associated with principal component analysis. *Technometrics*, 21, 341–349.

Kamohara, H., Takinami, A., Takeda, M., Kano, M., Hasebe, S., and Hashimoto, I. (2004). Product quality estimation and operating condition monitoring for industrial ethylene fractionator. *Journal of Chemical Engineering of Japan*, 37, 422–428.

Kaneko, H., Arakawa, M., and Funatsu, K. (2009). Development of a new soft sensor method using independent component analysis and partial least squares. *AIChE Journal*, 55(1), 87–98.

Kano, M., Hasebe, S., Hashimoto, I., and Ohno, H. (2004). Evolution of multivariate statistical process control: application of independent component analysis and external analysis. *Computers & Chemical Engineering*, 28(6-7), 1157–1166.

Kano, M., Lee, S., and Hasebe, S. (2009a). Two-stage subspace identification for softsensor design and disturbance estimation. *Journal of Process Control*, 19, 179–186.

Kano, M., Miyazaki, K., Hasebe, S., and Hashimoto, I. (2000). Inferential control system of distillation compositions using dynamic partial least squares regression. *Journal of Process Control*, 10, 157–166.

Kano, M. and Nakagawa, Y. (2008). Data-based process monitoring, process control, and quality improvement: recent developments and applications in steel industry. *Computers & Chemical Engineering*, 32, 12–24.

Kano, M., Tanaka, S., Hasebe, S., Hashimoto, I., and Ohno, H. (2003). Monitoring independent components for fault detection. *AIChE Journal*, 49(4), 969–976.

Kano, M., Tasaka, K., Hasebe, S., Masuda, S., Ogawa, M., Ootakara, S., and Shigemasa, T. (2009b). Practical direct pid controller tuning for process control. *AIChE Journal*, submitted.

Kitamori, T. (1981). A design method for nonlinear control systems based upon partial knowledge about control objects. In *IFAC World Congress*, IV25–IV30. Kyoto, Japan.

Kresta, J., MacGregor, J., and Marlin, T. (1991). Multivariate statistical monitoring of process operating performance. *Canadian Journal of Chemical Engineering*, 69, 35–47.

Kresta, J., Marlin, T., and MacGregor, J. (1994). Development of inferential process models using PLS. *Computers & Chemical Engineering*, 18, 597–611.

Kugemoto, H. (2005). Technology of controller performance monitoring and diagnosis in chemical plants. *Sumitomo Kagaku*, 2005-II, 1–10.

Lee, J., Yoo, C., and Lee, I. (2004). Statistical process monitoring with independent component analysis. *Journal of Process Control*, 14(5), 467–485.

Lundstrom, P., Lee, J., Morari, M., and Skogestad, S. (1995). Limitations of dynamic matrix control. *Computers & Chemical Engineering*, 19, 409–421.

Maruta, H., Kano, M., Kugemoto, H., and Shimizu, K. (2005). Modeling and detection of stiction in pneumatic control valves. *Trans. of the Society of Instrument and Control Engineers*, E-4(1), 18–26.

Mejdell, T. and Skogestad, S. (1991). Estimation of distillation compositions from multiple temperature measurements using partial-least-squares regression. *Industrial & Engineering Chemistry Research*, 30, 2543–2555.

Mitsubishi Chemical Corporation (2005). *RC Report 2005*.

Nishimura, Y. and Ootakara, S. (2007). The introduction of software for performance improvement of control loops. In *SICE Annual Conference on Control Systems (CCS)*, paper 72–2–1. Chofu, Japan.

Ogawa, M. (2006). Process control practice: the key to stable operation for maximizing the profit. In *KIChE Annual Meeting*. Korea.

Ogawa, M., Kano, M., and Hashimoto, I. (1998). Design of flow averaging level control system using i-p controller with gap. *Kagaku Kogaku Ronbunshu*, 24(2), 259–265.

Ogawa, M. and Katayama, T. (2001). A robust tuning method for i-pd controller incorporating a constraint on manipulated variable. *Transactions of the Society of Instrument and Control Engineers, Japan*, E-1(1), 265–273.

Ohshima, M., Ohno, H., and Hashimoto, I. (1995). Model predictive control experiences in the university-industry joint projects and statistics on mpc applications in japan. In *International Workshop on Predictive and Receding Horizon Control*, paper 72–2–1. Korea.

Overschee, P. and Moor, B. (1994). N4sid: subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, 30, 75–93.

Qin, S. (1998). Recursive pls algorithms for adaptive data modeling. *Computers & Chemical Engineering*, 22, 503–514.

Qin, S. and Badgwell, T. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11, 733–764.

Seki, H., Ogawa, M., Ooyama, S., Akamatsu, K., Ohshima, M., and Yang, W. (2001). Industrial application of a nonlinear model predictive control to polymerization reactors. *Control Engineering Practice*, 9, 819–828.

Shah, S., Mitchell, W., and Shoock, D. (2004). Challenges in the detection, diagnosis and visualization of controller performance data. In *Control Systems*. Quebec City, Canada.

Shinsky, F. (1977). *Distillation control for productivity and energy conservation*. McGraw-Hill, Inc.

Shoda, Y. (1998). Industrial benefits and unsolved problems of process control. In *IFAC Symposium on Dynamics and Control of Process Systems (DYCOPS)*, 223–225. Corfu, Greece.

Soma, S., Kaneko, O., and Fujii, T. (2004). A new approach to parameter tuning of controllers by using one-shot experimental data – a proposal of fictitious reference iterative tuning. *Transactions of Institute of Systems, Control and Information Engineers*, 17, 528–536.

Tasaka, K., Kano, M., and Ogawa, M. (2009). Direct pid/i-pd controller tuning and its application to unstable polymerization process. *Journal of Process Control*, submitted.

The Enginnering and Equipment Materials Users' Association (EEMUA) (2007). *Alarm systems, a guide to design, management and procurement*. EEMUA Publication, No. 191.

Verhaegen, M. and Dewilde, P. (1992). Subspace model identification. I - the output-error state-space model identification class of algorithms. *International Journal of Control*, 56(5), 1187–1210.

Yamashita, Y. (2006). An automatic method for detection of valve stiction in process control loops. *Control Engineering Practice*, 14, 503–510.

Yan, L., Xiwei, L., Noda, M., and Nishitani, H. (2007). Systematic design approach for plant alarm systems. *Journal of Chemical Engineering of Japan*, 40, 765–772.

Yan, W., Shao, H., and Wang, X. (2004). Soft sensing modeling based on support vector machine and Bayesian model selection. *Computers & Chemical Engineering*, 28(8), 1489–1498.

Keynote Lectures

Approximate dynamic programming approach for process control

Jay H. Lee * **Wee Chin Wong** *

* *School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: jay.lee,wwong@chbe.gatech.edu).*

Abstract: In this talk, we assess the potentials of the approximate dynamic programming (ADP) approach for process control, especially as a method to complement the model predictive control (MPC) approach. In the Artificial Intelligence (AI) and Operations Research (OR) research communities, ADP has recently seen significant activities as an effective method for solving Markov Decision Process (MDP), which represents a type of multi-stage decision problems under uncertainty. Process control problems are similar to MDPs with the key difference being the *continuous* state and action spaces as opposed to discrete ones. In addition, unlike in other popular ADP application areas like robotics or games, in process control applications first and foremost concern should be on the safety and economics of the on-going operation rather than on efficient learning. We explore different options within ADP design, such as the pre-decision state vs. post-decision state value function, parametric vs. nonparametric value function approximator, batch-mode vs. continuous-mode learning, exploration vs. robustness, etc. We argue that ADP possesses great potentials, especially for obtaining effective control policies for stochastic constrained nonlinear or linear systems and continually improving them towards optimality.

Keywords: Stochastic optimal control, constraints

1. INTRODUCTION

Model predictive control (MPC) is a technique in which the current control action is obtained by minimizing on-line, a cost criterion defined on a finite time interval. Nominal deterministic trajectories of future disturbance signals and uncertainties are necessarily assumed in order to obtain an optimization problem amenable to on-line solution via math programming. The solution generates a control sequence from which the first element is extracted and implemented. The procedure is repeated at the next time instant. Owing to its ability to handle constrained, multi-variable control problems in an optimal manner, MPC has become the de-facto advanced process control solution for the process industries today.

MPC is by now considered to be a mature technology owing to the plethora of research and industrial experiences during the past three decades. Despite this, it has some fundamental limitations, which prevents it from being a panacea for all process control problems. One well-known limitation is the potentially exorbitant on-line computation required for solving a large-scale, and potentially non-convex math program that scales with the dimension of the state as well as the length of prediction horizon. Recent developments (Laird and Biegler (2008)) have made some headway in tackling this problem although nontrivial computational challenges still exist.

⋆ JHL gratefully acknowledges financial support from Owens Corning and Weyerhaeuser.

The second limitation arises from the fact that the deterministic formulation adopted by MPC is inherently incapable of addressing uncertainty in a closed-loop optimal fashion. Its open-loop optimal control formulation used to find the control moves at each sample time means the fact that information about future uncertainty will be revealed, this being generally beneficial for control performance, is not considered. Most of the past attempts at ameliorating the impact of uncertainty has been reflected in robust MPCs formulations based on the objective of minimizing the worst-case scenarios (Scokaert and Mayne (1998)) at the expense of overly conservative policies. Multi-scenario formulations (Laird and Biegler (2008)) have also been developed but the number of scenarios is limited and they do not give closed-loop optimal policies in general. Stochastic programming based methodologies (Pena et al. (2005)) allow for recourse actions at the computational expense of enumerating an exponentially growing number of scenarios.

In this paper, we examine the possibility of lessening or removing the above-mentioned limitations by combining MPC with an approach called "approximate dynamic programming (ADP)." ADP is a technique that surfaced from the research on reinforcement learning in the Artificial Intelligence (AI) community (Sutton and Barto (1998); Bertsekas and Tsitsiklis (1996)). It has its theoretical foundations in the traditional dynamic programming by Richard Bellman (Bellman (1957)) but its computational bottlenecks, termed as "the curse of dimensionality" by Bellman himself, are relieved through ideas such as intelli-

gent sampling of the state space through simulations and function approximation. ADP, due to its root in AI, has mainly been studied in the context of Markov Decision Processes (MDPs), which involve discrete finite state/ action spaces and probabilistic transitions. Hence, its application to process control problems, which typically involve continuous state/ action spaces, is not straightforward. In addition, the characteristics of process control problems are somewhat different from those of robotics, games, and resource allocation problems. For example, in process control applications, the idea of "learning by mistakes" for the sake of efficient learning, may not be tolerated as mistakes often bring unacceptable consequences in terms of safety and economics. Hence, extension of ADP to process control may require significant care and possibly some new tools.

Design of an ADP algorithm involves a variety of choices, including type of function approximator, pre-decision vs. post-decision formulation, batch vs. continuous updating of the value table, and exploration vs. robustness tradeoff. We will visit these issues, carefully examining the implications of these choices in the context of designing a learning algorithm for process control applications. In addition, we will also consider the complementary nature or synergies between ADP and MPC.

The rest of the paper is organized as follows. In Section 2, we will briefly review the basics of MDP, ADP and also present a mathematical representation of the system we consider for control. In Section 3, we will examine the various options and choices and their implications for process control applications. In Section 4, we will present a few examples, including those involving both linear and nonlinear stochastic systems. In Section 5, we conclude the paper and discuss other control-related areas where ADP can potentially be useful in the process industries.

2. BACKGROUND

2.1 Markov Decision Processes and Approximate Dynamic Programming

Markov Decision Processes (MDPs) provide a framework for modeling real world processes that have a stage-wise structure. The stage can denote a time epoch or other quantities like location, processing step, etc. At any stage, the system is recognized as being in a state (designated as s), which is a set of attributes that aid decision-making. The set of all possible states is called state space (designated as S). Starting in state s belonging to S, there is a set of actions from which the decision-maker must choose. The set of all possible actions is called action space (A) and an element of the action space is denoted by a. When action a is taken in state s, and the system transitions to the next stage, it ends up in a unique next state $s' \in S$ in the absence of any uncertainty. However, for stochastic problems, there is a set of possible next states for each state-action pair. The probability of transition to a particular next state in this case is governed by a state transition probability function, P. In the process, reward $r(s, a, s')$ is received, which is determined by the reward function r. The dependence of r on s' is often suppressed by taking a weighted average over all possible states at the next stage. At each stage, actions are taken so that the sum of stage-wise rewards is maximized. In the presence of uncertainty, the expected sum of rewards is maximized. When infinite stages are present, i.e., extremely large time horizon, the future rewards are often discounted using a discount factor γ. When the number of stages is infinite, the problem is called an infinite horizon MDP as opposed to a finite horizon MDP for finite number of stages. In most applications, a stage symbolizes a time epoch. Therefore, the term time epoch or time step is often used synonymously with 'stage'.

More formally, MDP is defined by a tuple (S, A, P, R, γ) where S is a set of states, A is a set of actions, $P : S \times A \times S \rightarrow [0, 1]$ is a set of transition probabilities that describe the dynamic behavior of the modeled environment, $R : S \times A \times S \rightarrow \mathbb{R}$ denotes a reward model that determines the stage-wise reward when action a is taken in state s leading to next state s' and $\gamma \in [0, 1)$ is the discount factor used to discount future rewards. A γ value close to 0, places very little weightage on future rewards, while γ close to 1 results in very little discounting.

One of the fundamental properties of the MDPs is that the transition and reward functions associated with the stage-wise transition of state are independent of the past states and actions. Referred to as Markov property, this memory-less feature enables the decomposition of the overall optimization problem into separate stage-wise problems. This is accomplished by using a recursive relationship between the value of being in a state at any stage.

An important notion in this regard is the so called value function denoted by $V(s)$, which is defined as the (often discounted) sum of rewards over a time horizon which can be either finite or infinite (shown below) and discussed hereafter:

$$V^{\pi}(s) = \mathbb{E}\left[\sum_{t=0}^{\infty} \gamma^t r(s_t, \mu(s_t)) | s_0 = s\right] \tag{1}$$

where t denotes the time epoch, s_t is the state at time t and $\pi : S \rightarrow A$, is the policy that dictates the choice of action for a given state at time t.

The goal is to find an optimal policy that maximizes the value function for all $s \in S$. This is achieved by solving the Bellman equation (Bellman (1957)) for finite or infinite horizon problems. The optimal policy can be derived via dynamic programming. Let $a^*(s)$ be the optimal action to be taken when the system is in state s, independent of time t. $V^*(s)$ is called the optimal value function and is obtained as the solution to the (Bellman equation) (2), which must be solved for all $s \in S$:

$$V^*(s) = \max_{a \in A} \left\{ r(s, a) + \gamma \sum_{s' \in S} p(s'|s, a) V^*(s') \right\} \quad \forall s \tag{2}$$

$$a^*(s) = \pi^*(s) \triangleq \arg\max_{a \in A} \left\{ r(s, a) + \gamma \sum_{s' \in S} p(s'|s, a) V^*(s') \right\} \tag{3}$$

where symbol $p(.)$ denotes the probability of a quantity. It is well-known (Putterman (1994)) that for infinite horizon problems, a stationary optimal policy of the form in (3) exists, where $V^*(s)$ is the average discounted infinite horizon reward obtained when the optimal policy is followed starting from s until infinity (Putterman (1994)). This implies

that the state to action mapping in the form of optimal policy is independent of the time epoch. The existence of stationary optimal policy is conditioned on the properties of model elements. One of the sufficient conditions is that there be a finite action space A_s corresponding to each state $s \in S$, maximum attainable stage-wise reward is finite and discount factor $\gamma \in [0, 1)$. The alternative sets of sufficient conditions for existence of a stationary optimal policy for discounted infinite horizon MDPs can be found in (Putterman (1994)).

It must be noted that the set of Bellman equations also called optimality equations are difficult to solve analytically because of the presence of the max operator. One of the popular solution methods is called value iteration (Putterman (1994)): Starting with an arbitrary value function $V_0(s)$ for each state $s \in S$, the value function is iteratively improved by successive substitution into (2) until ϵ-convergence is reached. The operator for one iteration, that is the maximization in (2), can be denoted as H such that $V_{n+1} = HV_n$. The sequence of estimates of value function $V(s)$, $\forall s \in S$, converges to a fixed point solution. This is a consequence of Banach's theorem for contraction mappings (Putterman (1994)). Since H is a proven contraction map, the convergence properties hold.

Due to ease of implementation, value iteration is perhaps the most widely used algorithm in dynamic programming. Certain other methods like policy iteration (Bertsekas (2005)), a hybrid between value iteration and policy iteration (Powell (2007)) and linear programming method for dynamic programs (Farias and Roy (2003)) are also used depending on the problem structure. The complexity of the value-iteration algorithm grows as a function of $o(|S|^2 \times |A|)$. This is attributed to the following three aspects of the value iteration:

(1) Equation (2) needs to be solved for all s belong to S, so the solution time is directly proportional to $|S|$.
(2) The complexity of max operation depends on the size of the action space $|A|$.
(3) The calculation of expectation within the max operator depends on the number of possible next states, i.e., $|S|$.

In the presence of very large state and (or) action spaces, the value iteration algorithm cannot be implemented in its exact form. Several approximation methods have been developed to circumvent this difficulty, including: approximate dynamic programming methods using value function approximations (Powell (2007)), Q-learning, temporal difference learning (Sutton and Barto (1998)) functions (Farias and Roy (2003)) and dynamic programming methods using post decisions state (Powell (2007)).

All the above methods assume that the system state is completely known or observed at all times. When this assumption does not hold, the equivalent framework is called a Partially Observed Markov Decision Process (POMDP) (Cassandra et al. (1994)), for which a significant but less body of literature exists.

2.2 System Definition: Process Control Problems

Consider the optimal control of the following discrete-time stochastic system:

$$x_{t+1} = f(x_t, u_t, \omega_t) \tag{4}$$

where $x_t \in \mathcal{X} \subseteq \mathbb{R}^{n_x}$ refers to the system state at discrete time index t, $u_t \in \mathcal{U} \subseteq \mathbb{R}^{n_u}$ a control or action vector, and ω_t an exogenous, unmeasured, stochastic signal. x may contain physically meaningful states as well as measured disturbances, and parameters subject to uncertainty. f refers to the single-stage transition function. For problems where the system's dynamics are represented by ordinary differential equations, f is then the result of numerical integration across a single sample-time, with vectors u and ω held constant. Throughout this paper, it is assumed that full state feedback is available. In the event that only output feedback is available, x is interpreted as an information vector that contains the sufficient statistics of the state estimate's probability density function. Such lifting is possible as the information vector is governed by another related set of differential equations. (*i.e.*, the filter dynamics).

Let $\mu \in \Gamma$ be a 'state-feedback policy' that maps the state vector to the action vector, where Γ represents the set of all admissible (stationary) such policies. To distinguish from the earlier value function $V(s)$, $J^\mu(x)$ will be used to denote the 'cost-to-go' function, which is defined as the infinite horizon, discounted sum of the stage-wise costs under the policy μ starting from an arbitrary state x:

$$J^\mu(x) = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k \phi(x_k, u_k = \mu(x_k)) | x_0 = x\right] \tag{5}$$

where ϕ represents a pre-specified stage-wise cost (*e.g.* $\phi(x, u) := ||x||_Q^2 + ||x||_R^2$) and $\gamma \in [0, 1)$ is a discount factor. The goal then is to find the optimal (stationary) policy $\mu^* : \mathcal{X} \to \mathcal{U}$, that yields the minimum cost-to-go function as below:

$$J^{\mu^*}(x) = \min_{\mu \in \Gamma} \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k \phi(x_{t+k}, u_{t+k} = \mu(x_{t+k})) | x_t = x\right] \tag{6}$$

$J^{\mu^*} : \mathcal{X} \to \mathbb{R}^{0+}$ is the *optimal* 'cost-to-go' function and is an indication of the attractiveness of a given state in terms of future rewards. By definition, $J^{\mu^*}(x) \leq J^\mu(x), \forall x$ and $\forall \mu \in \Gamma$.

The main difference between the above and the previously introduced MDPs is that the state and action spaces are continuous. However, the fundamental concepts of DP still apply here. Based on the principle of optimality (Bellman (1957)), one is able to re-write (6), thereby obtaining Bellman's optimality equations:

$$\begin{aligned} J^{\mu^*}(x) &= \min_{u \in \mathcal{U}} \left\{ \phi(x, u) + \gamma \mathbb{E}_{(\omega|x)}[J^{\mu^*}(f(x, u, \omega))] \right\} \\ &= \left(T J^{\mu^*}\right)(x) \end{aligned} \tag{7}$$

T above represents the single-pass DP operator represented by the minimization operation. The optimal policy is implicitly obtained through the solution of the associated single-stage optimization:

$$\mu^*(x) = \arg\min_{u \in \mathcal{U}} \left\{ \phi(x, u) + \gamma \mathbb{E}_{(\omega|x)}[J^{\mu^*}(f(x, u, \omega))] \right\} \tag{8}$$

In principle, the optimal control problem is solved once J^{μ^*} is known. It is noted that for deterministic problems (where the expectation operator is dropped), the DP formalism provides a convenient way of solving multi-stage problems through an equivalent single-stage optimization.

Unfortunately, analytical solutions to Bellman's optimality equations are available for only a small class of problems, of which the celebrated Linear Quadratic Gaussian (LQG) problem is one. For situations of practical interest, numerical techniques are required. Similar to the case of discrete state and action space, the repeated application of T on an arbitrarily initialized cost-to-go leads to convergence and underpins the idea behind Value Iteration (VI).

$$J^{\mu^*}(x) = TJ^{\mu^*}(x) = \lim_{i\to\infty}(T)^i J^{\mu}(x),\ \forall \mu, x \qquad (9)$$

In process control problems, due to the continuous nature of the state and action spaces which must be discretized, numerical solutions become quickly bottle-necked as the problem dimensions grow. In fact, the growth would be exponential as the number of discretized points grows with the dimension as such. Hence, a naive application of VI in this case is computationally prohibitive and the 'curse-of-dimensionality' is even more apparent in continuous problems. For problems with continuous state and action space, one needs to resort to approximations that involve an intelligent state-sampling/ discretization scheme and/ or an efficient representation of the cost-to-go (Lee and Lee (2006); Powell (2007)).

2.3 Approximate dynamic programming for problems with continuous state and action space

Value iteration or policy iteration in general can work with only finite state space. For systems with continuous state and action space, one must then work with discretized state state, either through gridding, or more preferably, sampling. It is often the case that only a small portion of $\mathcal{X}$ and $\mathcal{U}$ will ever be visited under optimal and/ or high-quality sub-optimal policies. This is especially true when the dimension of the state space is large compared to that of the input. Let us denote the subset of the state space that is 'relevant', *i.e.*, visited with non-trivial probability under the optimal control, as $\mathcal{X}^*_{REL}$. Such a set would be continuous but much smaller-sized than $\mathcal{X}$ in general. The key notion is that if one could identify $\mathcal{X}^*_{REL}$ or a parsimonious superset of it, one can sample the set with sufficient density to perform the dynamic programming at significantly reduced computation. Of course, the difficulty is that it is not easy to obtain such a set ahead of time without knowing the optimal controller itself.

The ADP approach proposed by the authors of (Lee and Lee (2004, 2006); Tosukhowong and Lee (2009)) for process control applications, the skeleton of which is described in this subsection, employs carefully designed simulation schemes for the sampling of the state space and function approximation (for the purpose of cost-to-go interpolation) to this end. For the VI-variant, we have the following off-line computations:

(1) Identify a finite-sized, 'relevant' state-space, $X_{sam} \subset \mathcal{X}$, $|X_{sam}| = N$. This is achieved, for instance, by simulating all possible combinations of sub-optimal policies (potentially with dithering) and operating conditions. The latter are defined as all starting states of interest (for servo problems) as well as potential values of measured disturbance values. Dithering may also be introduced for the purpose of exploration.
(2) Assign a cost-to-go for all elements of X_{sam}, using the simulation data according to (5). The initial, finite-sized 'cost-to-go' table, denoted by $\mathcal{T}_{[0]} \triangleq \left\{x, \hat{J}^{\mu^*}_{[0]}(x)\,\middle|\, x \in X_{sam}\right\}$, is obtained. The symbol $\hat{(\cdot)}$ is used to emphasize the approximate nature of the cost-to-go sequence, even at its limit. Exact initialization is not critical per se since the fixed point derived from the following step is unique.
(3) Obtain converged cost-to-go values for X_{sam} through VI, yielding the sequence of value tables $\{\mathcal{T}_{[0]}, \mathcal{T}_{[1]}, \ldots\}$. Since the VI requires the evaluation of the cost-to-go function for states $(f(x,u,\omega))$ not necessarily in X_{sam}, a well-designed function approximator is needed to interpolate among the stored values (see discussion in Section 3.1). A certain choice of function approximator ensures that each pass of the iteration is a *contraction-map* with a unique fixed point (see Section 3.1). In other words, each step of the modified VI involves:
$$\hat{J}^{\mu^*}_{[i+1]}(x) = \left(TF(\hat{J}^{\mu^*}_{[i]})\right)(x),\ \forall x \in X_{sam} \qquad (10)$$
Here $F(\hat{J}^{\mu^*})$ denotes the cost-to-go function approximator based on the stored values $\{\hat{J}^{\mu^*}_{[i]}(x), x \in X_{sam}\}$. Termination occurs when $\|\hat{J}^{\mu^*}_{[i+1]} - \hat{J}^{\mu^*}_{[i]}\|_\infty$ is less than a pre-defined tolerance.
(4) Return to step 1, since the relevant domain of the state-space may not be properly ascertained a-priori. Otherwise, use the converged values for online control.

The authors of Ma and Powell (2009) used an approximate policy iteration scheme where $J^{\mu}(x), \forall x \in \mathcal{X}$ is assumed to be linear in a set of basis functions (known or otherwise assumed to be orthogonal polynomials of sufficiently large degree). The coefficients are learnt through a least-squares procedure once the system of interest is allowed to evolve according to the current policy, which is similar to step (1) where relevant states are collected. The limitation is that suitable basis functions are difficult to ascertain in general.

3. ISSUES AND CHOICES

3.1 Function approximation and stable learning

The need for function approximation for the purpose of generalization has been discussed. Given a training set $\mathcal{T} \triangleq \{x_i, \hat{J}(x_i)\}_{i=1}^N$, a value table composed of a finite number (N) of input $(x_i \in \mathcal{X})$ and target values $(\hat{J}(x_i) \in \mathbb{R})$, a function approximator, F, whose domain is $\mathcal{X}$, maps a query point $x_q \in \mathcal{X}$ to a subset of the real line.

The dominant and natural choice for function approximators has typically involved parametric global approximators such as neural networks or the use of basis functions

such as high order orthogonal polynomials or Fourier series (Tsitsiklis and Roy (1996); Konidaris and Osentoski (2008)). While this approach has met with some success in certain applications (*e.g.* in Backgammon (Tesauro (1992)), it is not immune from divergent behavior (Lee and Lee (2004, 2006)) when employed in the context of ADP. In certain cases, the off-line iteration would fail to converge, with the cost-to-go approximation showing non-monotonic behavior or instability with respect to iterations. Thrun and Schwartz (1993) were the first to attribute the failure with function approximation to an 'over-estimation' effect. Sabes (1993) demonstrated that sub-optimality can be severe when a global approximator with a linear combination of basis functions is employed. Boyan and Moore (1995) provide insightful illustrations showing the failure of popular function approximators during off-line learning.

There are considerably fewer papers that address function approximation schemes for problems with continuous state and action spaces (Ma and Powell (2009), Lee et al. (2006)). The problem of linear quadratic regulation, for which the value function is known to be quadratic in structure, is a noted exception Bradtke (1993). Ormoneit and Sen (2002) proposed a kernel-based approach for problems with continuous states but finite actions and demonstrate convergence to the optimal cost-to-go value function with an increasing number of samples and decreasing kernel bandwidth under a model-free scheme. Ma and Powell (2009) proposed a provably convergent approximate policy iteration under the assumption of known basis functions and other technical conditions.

Stable learning during the off-linear value iteration step of the proposed ADP strategy is highly desirable as it can be frustrating to run a large number of iterations only to have the result "blow up" all of sudden due to some complicated coupling between the function approximation error and value iteration. To have provable convergence of the approximate value iteration (not necessarily to the optimal value function, however), one needs to use a function approximator with a certain property called "non-expansion" . Gordon (1995) discussed the viability of using such a class of function approximators. With such a choice, the overall operator composed of value-iteration and then function approximation results can be shown to be a contraction map therefore ensuring convergence.

Definition 1. A γ-contraction mapping m defined on a normed vector space (mapping elements from this space, $\mathcal{V}$, to itself) is defined as such:

$\forall v_1, v_2 \in \mathcal{V}$, $||m(v_1) - m(v_2)|| \leq \gamma ||v_1 - v_2||$, $\gamma \in [0, 1)$

where v_1, v_2 are arbitrarily chosen elements of $\mathcal{V}$.

Definition 2. When $\gamma = 1$, $m : \mathcal{V} \rightarrow \mathcal{V}$ is termed a non-expansion (Gordon (1995)).

From Banach's fixed-point theorem, it can be easily shown that every the iterated sequence $\{v, m(v), m^2(v), \ldots\}$ converges to a unique fixed point. As explained earlier, the proposed ADP method starts with initial estimates $\hat{J}^{\mu^*}_{[0]}(x)$, $\forall x \in X_{sam}$. This is followed by function approximation (recall that this mapping is denoted by F), and an application of the DP operator, T to yield $\hat{J}^{\mu^*}_{[1]}$. The process is repeated again. Our experience with the ADP approach (Lee and Lee (2004, 2005)) has been that stability of learning and the quality of a learned control policy are critically dependent on the structure of the function approximator. A sufficient condition for convergence is to demonstrate that the overall operator T with function approximator F is a contraction map. This, in turn, holds true if F is a non-expansion map.

Proposition 1. T is a γ-contraction map if F is non-expansive.

Proof. Given arbitrary vectors $\hat{J}_1, \hat{J}_2 \in \mathbb{R}^N$,

$$||TF(\hat{J}_1) - T^{\mu^*}F(\hat{J}_2)||_\infty \leq \gamma ||F(\hat{J}_1) - F(\hat{J}_2)||_\infty \quad (11)$$

$$\leq \gamma ||\hat{J}_1 - \hat{J}_2||_\infty \quad (12)$$

The first line is true since T is a γ-contraction map defined on the space of value functions. The second inequality follows if one employs a function approximator with a non-expansion property.

Function approximators that employ averaging, as defined below, can be shown to possesses a non-expansion property.

Definition 3. F is an averager if every fitted valued is the weighted average of of target values, potentially with the addition of a bias term. Specifically,

$$F(\hat{J})(x_q) = \beta_0(x_q) + \sum_{i=1}^{N} \beta_i(x_q)\hat{J}(x_i) \quad (13)$$

Here, $\{\beta_i\}_{i=0}^N \geq 0$, and $\sum_{i=1}^{N} \beta_i \leq 1$. Note that the weights β are allowed to depend on the query point (x_q) and input values $(\{x_i\}_{i=1}^N)$ but not the target values.

That such an averager is a non-expansion (*i.e.* (12) is true) is easily demonstrated.

One such type of approximator we have experimented with extensively is instance-based (Lee et al. (2006)) local averagers, such as k-Nearest Neighbors-based (kNN) predictors. Instance-based algorithms are non-parametric representations using stored points 'close' to a query point for making predictions. Closeness is usually defined according to some distance metric (such as Euclidean distance). Predictions of the weighted kNN are given by:

$$F(\hat{J})(x_q) = \beta_0(x_q) + \sum_{x_i \in \mathcal{N}_k(x_q)} \beta_i(x_q)\hat{J}(x_i) \quad (14)$$

where $\mathcal{N}_k(x_q)$ refers to the set containing the k points closest to x_q. The weights (normalized by constant c) are defined as: $\beta_i = c((x_q - x_i)^T W (x_q - x_i))^{-0.5}$, $i \geq 1$. W is a feature weighting matrix use to scale and also to emphasize dimensions that are more important.

3.2 Cautious learning for robustness

It has been demonstrated (Smart and Kaelbling (2000); Lee et al. (2006)) that simply using a local averager (with $\beta_0 = 0$), though guaranteeing convergence, does not necessarily give a converged function leading to a stable closed-loop behavior. This is because function approximation

error can be significant, particularly when the training data is insufficient. Safeguards against 'over-extrapolation' during value iteration is often needed for the successful implementation of the proposed ADP method. For a query point located in regions with little data present, distance-weighted averaging may fail to provide meaningful generalizations of the cost-to-go. Prevention of taking such a query point may be achieved by including in the cost-to-go term β_0, a penalty that is imposed whenever the minimization step encounters a query point (x_q) far away from X_{sam}:

$$\beta_0(x_q) = A.U\left(\frac{1}{f_\Omega(x_q)} - \rho\right) \cdot \left(\frac{\frac{1}{f_\Omega(x_q)} - \rho}{\rho}\right)^2 \quad (15)$$

Here, ρ is a data-density threshold value, A a scaling parameter, and U, the Heaviside step function that returns a zero value whenever its argument is non-positive and unity, otherwise. $f_\Omega(x_q)$ is a measure of data density as ascertained by fitting a Kernel density estimator over training set Ω:

$$f_\Omega(x_q) = \frac{1}{N_\Omega}\sum_{i}^{N_\Omega} K\left(\frac{x_q - x_i}{\sigma}\right) \quad (16)$$

where kernel $K(\cdot)$ refers to a zero-mean Gaussian with variance $\sigma^2 I_{n_x}$. For generality, Ω is allowed to differ from X_{sam}. Furthermore, a bound is imposed on β_0 whenever it exceeds a threshold value. Tuning rules for ρ, A, σ can be found in (Lee et al. (2006)) and are not reproduced here in the interest of brevity.

3.3 Pre-decision vs. post-decision state formulation

For stochastic control problems, the single-stage optimization required during off-line value-iteration (see (7)) and on-line implementation of the optimal policy (see (8)) requires the generally cumbersome evaluation of an expectation.

The use of an intermediate post-decision (x^p) state, first introduced by Roy et al. (1997) and employed extensively by Powell (2007) in solving operations research problems, oftentimes allows for more computationally effective strategies. x^p refers to the the system state immediately after the control vector is introduced but before the uncertainty is realized. As a result, f is decomposed into the following sub-transitions:

$$x_t^p = f_1(x_t, u_t)$$
$$x_{t+1} = f_2(x_t^p, \omega_t)$$

where the composition of f_1 and f_2 is equivalent, in effect, to f, in (4). Note that f_1 describes a deterministic transition between the pre-decision state variable (x) and x^p. f_2 involves the transition due to uncertainty after the control action is implemented. Consequently, the value function of x^p, $J^{\mu,p}(x^p)$, may be expressed in terms of the value function of x, as such:

$$J^{\mu,p}(x_t^p) = \mathbb{E}_{(\omega|x_t^p)}\left[J^\mu(x_{t+1})\right], \forall \mu \quad (17)$$

By considering the optimal policy μ^*, and substituting (7) into (17), the min and $\mathbb{E}$ operators are interchanged, yielding:

$$J^{\mu^*,p}(x_t^p) = \mathbb{E}_{(\omega|x_t^p)}\left[\min_{u_{t+1}\in\mathcal{U}}\left\{\phi(x_{t+1}, u_{t+1}) + \gamma J^{\mu^*,p}(x_{t+1}^p)\right\}\right] \quad (18)$$

The single-stage on-line optimization is also streamlined:

$$\mu^*(x) = \arg\min_{u\in\mathcal{U}}\left\{\phi(x,u) + \gamma J^{\mu^*,p}(f_1(x,u))\right\} \quad (19)$$

The introduction of the post-decision state allows the generally non-commutative min and $\mathbb{E}$ operators to be interchanged. (18), used off-line during value iteration, consists of an independent collection of deterministic optimization problems, which may be run *in parallel* using off-the-shelf solvers. It is noted that the latter have been cornerstone of MPC technology. In this case, differentiable local averagers such as Kernel regression (Hastie et al. (2008)) may be employed. In this case, given the training set $\{x_i, \hat{J}(x_i)\}_{i=1}^N$, we have:

$$F\hat{J}(x_q) = \beta_0(x_q) + \frac{1}{\sum_{j=1} K\left(\frac{x_q - x_j}{\sigma}\right)} \cdot \sum_{i=1}^{N} K\left(\frac{x_q - x_i}{\sigma}\right)\hat{J}(x_i) \quad (20)$$

$\beta_0(x_q)$ is defined as in (15), where the Heaviside step function is replaced by a smooth approximator.

In addition to the off-line iteration step, the on-line calculation of the optimal input (in (19)) based on a precomputed post-decision state cost-to-go function, is a deterministic optimization that does not involve an expectation operator, much like the one solved for MPC. Again, an off-the-shelf NLP solver may be employed. Another benefit, this time, compared to MPC, is that it involves only a single stage optimization as the cost-to-go function contains the precomputed optimal cost information for the rest of the horizon.

As with the pre-decision case, we can define a γ-contraction H^p so as to simplify (18). The aforementioned discussion on value-function approximation still holds for this case.

3.4 Adding new state samples: batch mode vs. continuous mode learning

In the standard ADP algorithm presented (that is, without Step (4)), we fix the set of sampled states, X_{sam}, in the beginning and do not introduce any more samples as the learning proceeds. One potential problem with this is that it may not contain sufficient samples in all the important regions of the state space.

Additional samples can be introduced as the learning proceeds in two different ways. First is to perform the simulation and the value iteration simultaneously, resulting in an X_{sam} that varies with simulation time. This approach is seen in the methods known as real-time dynamic programming (RTDP) (Barto et al. (1995)) and RTADP (Pratikakis et al. (2009)). In these approaches, one typically starts with an empty value table and introduces entries one by one as simulations proceed. Whenever a "new" state, a state that is not already recorded and does

not have a "sufficiently close" neighbor recorded in the value table, is visited during the simulation, it is entered into the value table and its cost-to-go value is assigned by evaluating the Bellman equation. The optimal action suggested by the current value table is implemented and next state is sampled according to the transition equation. If a state is revisited or there are sufficient neighbors close by, the value update for that particular state (or the sufficiently close neighbor) is performed without adding a new entry. This goes on until "new" entries are no longer added to the value table. Just how fast the convergence happens depends on the level of exploration, which will be discussed later.

The second way is to alternate between the modes of simulation and value iteration. A value iteration gives a converged cost-to-go function, which corresponds to a new policy. This new policy can be simulated to find a set of new state samples to be added to the current X_{sam}^i to yield X_{sam}^{i+1}. This continues until the simulation no longer yields different state samples.

In the case that an accurate simulation model is not available, one may have to replace the simulation with an actual on-line implementation. The continuous mode learning behaves much like adaptive control as the value table, and therefore the control policy, gets updated at every sample time. In addition, the performance during the initial phase of learning, when the value table has very few entries, may be highly unpredictable and poor. Hence, it may not be suitable in an industrial setting, whereas the other option in which the control policy remains fixed until the next off-line value iteration. Of course, given the typical constraints of industrial processes, one still has to exercise caution in implementing only a half-learned cost-to-go function on-line. The trade-off between exploration and robustness in this context is discussed next.

3.5 Exploration vs. robustness

The trade-off between exploration and robustness becomes one of the central issues when one chooses to expand X_{sam} as a part of the learning. In general, exploration gives new information to improve the eventual closed-loop performance (by expanding X_{sam}) but at the expense of slower convergence and decreased robustness. Exploration can be performed in two ways. First, one can add dither signals to the input to encourage more randomness in the state trajectories. Second, one can use an optimistic initialization of the cost-to-go value for previously unseen states, which will encourage the optimizer to choose actions leading the state trajectory to those states. If the learning is to be done directly in closed loop, the latter practice may be unacceptable for industrial practices as excursion to previously unseen states could jeopardize the safety and economics of the on-going operation. In other words, unlike in applications like robotics, "learning by mistake" is not a permissible practice for most industrial process control applications. One can in fact actively prevent such potentially harmful excursions through "pessimistic" valuation of unseen states. The previously discussed penalty approach in Section 3.2 is one way to achieve this. On the other hand, carefully chosen dither signals may be able to generate sufficiently new trajectories without imposing unacceptable risks or unduely slowing down the convergence.

3.6 Model-based vs. Model-less approach

For many industrial processes, sufficiently accurate models may not be available. In such a case, one can resort to empirical models derived from input-output data. In such as case, the state vector may simply be composed of the past input and output samples. The model can be learned separately from the ADP or it can be done as a part of it. In the latter, one learns instead of the cost-to-go function a function called Q function, which has the argument of state-input pair and assigns the cost-to-go to the pair. In other words, Q function already has the model embedded in it, which is learned together with the cost-to-go function. The two approaches are tried and compared in a recent paper by Lee and Lee (2005).

3.7 Integration with MPC

ADP can be integrated with model predictive control at several fronts. Some obvious ways include: (1)using MPC in the initial simulation to sample relevant states, (2)using the learned cost-to-go function in order to reduce the horizon size, and (3)use of the nonlinear programming solver of MPC in the post-decision-state formulation. Other methods may include the dual mode implementation, where MPC replaces the ADP controller whenever one encounters a state that is sufficiently new and the information in the learned value function cannot be trusted. Such states can be collected separately and added to X_{sam} in the next phase of value iteration.

4. EXAMPLES

Here, we demonstrate the proposed ADP algorithm on a variety of stochastic optimal control problems.

4.1 Example 1: Constrained linear stochastic system-double integrator problem

We consider the following constrained double integrator problem studied by (Batina (2004)) in the context of MPC for stochastic systems:

$$x_{t+1} = Ax_t + Bu_t + \Upsilon w_t$$

where matrix $A = [1\,0; 1\,1]$ [1], $B = [1; 0]$, $\Upsilon = [1; 0]$ and ω_t is zero-mean, white Gaussian noise with its second moment, $\mathbb{E}[\omega_t \omega_t'] = 0.2$, $\forall t$. The nominal stage-wise cost is $\tilde{\phi}(x_t, u_t) \triangleq 0.7||x_t||_2^2 + 0.33||u_t||_2^2$. The second dimension of the state vector is constrained, as is the input vector: $u_t \in [-0.5, 0.5]$, $x_2 \geq 0$.

The goal is to bring the system optimally from an arbitrary initial state ($[0; 14]$ in the following simulations) to the origin, whilst respecting the imposed constraints. We compare the performance of a Linear Model Predictive Controller (LMPC) (with prediction and control (p) horizon set to 15) against the proposed ADP approach based on the post-decision state variable. The post-decision state

[1] in Matlab notation

is defined as the quantity obtained after an action is taken but before the uncertainty is realized. That is, $x_t^p \triangleq Ax_t + Bu_t$. Since ω is an unbounded signal, we employ a soft-constraint approach for both LMPC (to avoid running into infeasibility issues) and the proposed ADP strategy. As is typically done, LMPC is implemented assuming ω remains at its nominal value of 0 over the prediction horizon. Namely, for LMPC, we solve at each time step, t:

$$\min \sum_{k=0}^{p} \tilde{\phi}(x_{t+k}, u_{t+k}) + 100||\epsilon_{t+k}||_2^2 \tag{21}$$

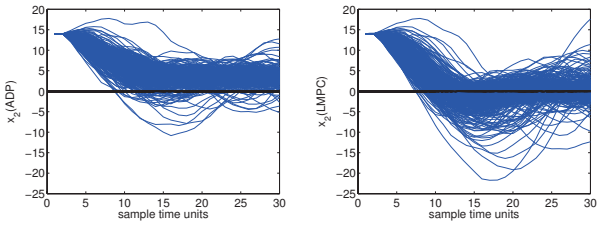
where $\epsilon_{t+k} \geq 0$ are non-negative auxiliary decision variables representing the least amount of slack required to make the LMPC problem feasible. That is, $[0\ 1]\, x_{t+k} + \epsilon_{t+k} \geq 0$. These inequalities are easily incorporated into the math program defined by (21). Also, the input vector is constrained to satisfy the aforementioned bounds of ± 0.5.

For the proposed ADP approach, we set the discount factor to a value close to unity, that is, $\gamma = 0.98$ and modify the stage wise cost to penalize deviations from the state constraints. Namely, $\phi(x_t, u_t) \triangleq \tilde{\phi}(x_t, u_t) + 100 \max(0, -[0,\ 1]x_t)^2$. Hard constraints on u are imposed during the off-line value iteration process and on-line implementation of the ADP-based controller.

To construct X_{sam}, we used an LMPC controller (with horizon length 5) and conducted closed-loop experiments bringing the system from 40 different initial post-decision states to the origin. Note that the initial state used for on-line testing is excluded from these 40 initial states. Namely, we consider various combinations of the sets $\{-2, 0, -1, 1, 2\}$ and $\{-4, -2, 0, 2, 4, 6, 8, 10\}$ to create various values for the first and second dimension of the initial state respectively. Consequently, a total of 3587 training points, whose initial cost-to-go values were initialized by computing the cost for LMPC over a sufficiently long horizon, was obtained as a result of the initialization scheme. For the purpose of function approximation, we employed kernel regression with the bandwidth, σ, set to 0.16. To avoid over-extrapolation, we selected $A = 1220$, and $\rho = 0.2652$. Value-iteration converged within 50 iterations, where the relative error termination criterion is set to $\left|\left| \frac{\hat{J}^{\mu^*,p}_{[i+1]} - \hat{J}^{\mu^*,p}_{[i]}}{\hat{J}^{\mu^*,p}_{[i]}} \right|\right|_\infty \leq 0.1$.

Results from 500 stochastic realizations presented as follows. As can be seen from Table 1, the proposed ADP controller has an average [2] finite horizon score an order of magnitude lower than a deterministic approach typified by LMPC. In particular, LMPC suffers from excessively high variance in terms of closed-loop performance. A look at the time series plots of the second dimension of x for both methods (see Fig. 1) reveals that LMPC results in significant constraint violation. On the other hand, the majority of the realizations based on the ADP approach do not violate the lower bound constraint.

[2] based on sample averaging



(a) ADP algorithm. (b) LMPC.

Fig. 1. Double integrator example: x_2 vs. t for 500 realizations. Lower bound for x_2 is 0.

4.2 Example 2: Constrained nonlinear stochastic system - chemostat problem

Consider the governing equations of an archetypal chemostat.

$$\dot{x}_1 = x_1 \tfrac{\mu_{max} x_2}{\kappa + x_2} - x_1 u$$
$$\dot{x}_2 = u[x_{2,f} - x_2] - \tfrac{\mu_{max}}{Y} \tfrac{x_1 x_2}{\kappa + x_2}$$

where $x = [x_1; x_2] \in \mathbb{R}^2$ is the state vector composed of the instantaneous concentration of the product (x_1) and substrate (x_2) respectively. $0 \leq u \in \mathbb{R}$, the dilution rate, is the non-negative manipulated variable. $x_{2,f}$ refers to the instantaneous concentration of the substrate feed. The maximum specific growth rate μ_{max} is set to 1, the yield coefficient to 1 and κ to 0.02. For the following simulations, the sampling rate is set to 0.5.

For the purpose of simulation, we assume that the feed concentration ($x_{2,f}$) fluctuates around a mean value of 1, and is perturbed by zero-mean, white Gaussian noise (ω):

$$x_{2,f,t} = 1 + \omega_t,\ \mathbb{E}[\omega_t \omega_t'] = 10^{-3} \tag{22}$$

It is desirable to maximize the productivity of the product, $\mathcal{P}_t \triangleq x_{1,t} u_t$, whilst ensuring that the conversion of the substrate, $f_{x_2} \triangleq 1 - \frac{x_2}{x_{2,f}}$, does not go lower than a relatively high value of 95%. Such an economically motivated constraint is common in several key process industries, such as bioethanol production. There is a tradeoff between productivity and conversion. Productivity increases with dilution rate and then decreases as the system approaches washout. Conversion, on the other hand, is a decreasing function of space-velocity or equivalently the dilution rate. Maximum productivity ($\mathcal{P}^* = 0.7543$) occurs at a dilution rate that corresponds to conversion levels significantly below the required 95% threshold.

We compare the performance of Non-linear MPC (NMPC) against the proposed ADP strategy. Instead of full-fledged NMPC, we employ successive-linearization based MPC (slMPC), a computationally efficient alternative proposed by Lee and Ricker (1994). For this example, we have found the closed-loop performance of slMPC to be similar to that of NMPC. For slMPC, we employed a prediction and

Table 1. Example 1: comparing performance

Score	ADP	LMPC
$\mathbb{E}\left[\sum_{t=0}^{30} \phi(x_t, u_t)\right]$	1600	10000

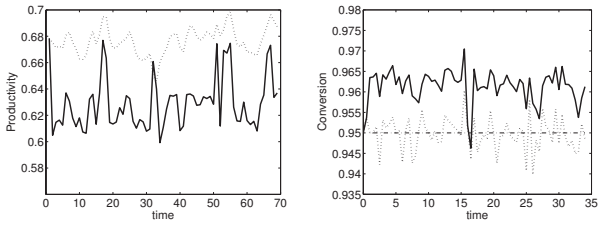
control horizon, p, of 10 sample units. The following math program is solved at each time instant:

$$\min \sum_{k=0}^{p} \left(||\mathcal{P}_{t+k} - \mathcal{P}^*||_2^2 + 100||\epsilon_{t+k}||_2^2 \right) \qquad (23)$$

where as in the previous example, $\epsilon \geq 0$, is a non-negative variable representing the least amount of slack required for conversion to be greater than 95%. That is, $\epsilon_{t+k} + f_{x_2,t+k} \geq 0.95$. The idea is to regulate the system at an equilibrium point that corresponds to the largest possible value of the dilution rate without exceeding the conversion bound so that productivity is maximized. The dynamics of the system are assumed to be governed by matrices obtained through linearization of the governing ordinary differential equations about the current state and past input vector. This results in a convex quadratic program.

For the proposed ADP approach, γ is set to 0.98 and the stage-wise cost defined as such: $\phi(x_t, u_t) \triangleq ||\mathcal{P}_t - \mathcal{P}^*||_2^2 + 100 \max(0, 0.95 - f_{x_2,t})^2$. To determine X_{sam}, we used an slMPC controller (with horizon length of 10 time units) and conducted closed-loop experiments regulating the system at an initial state corresponding to a conversion of 0.95. A total of 300 training points was obtained from the initialization scheme. We used kernel regression for function approximation with the bandwidth, σ, set to 0.15, A to 1.93 and ρ to 0.087 in order to prevent over-extrapolation. Value iteration terminated within 50 iterations with a relative error tolerance of 0.1.

Results from a typical realization are depicted in Fig. 2. It is apparent that the ADP-based approach, compared to slMPC, results in minimal constraint violation at the expense of slightly lower productivity. It is noted that the steady-state productivity corresponding to 95% conversion is 0.68.



(a) Productivity vs time. (b) Conversion vs time.

Fig. 2. Example 2. Closed-loop performance of a typical realization. ADP: solid line (-); slMPC: dotted line(..); lower bound on conversion: dash-dot (-.)

4.3 Other examples in the literature

There are a number of other applications to process process control problems in the published literature. Interested readers may look at the following references for applications to more complex examples. These include: integrated reactor-separator system control (Tosukhowong and Lee (2009)), dual adaptive control (Lee and Lee (2009)), fed-batch reactor control (Peroni et al. (2005)), and microbial reactor (Kaisare et al. (2003)).

5. CONCLUSIONS

We have examined the potentials of ADP for process control and found that it can complement MPC to reduce the on-line computational load and also address stochastic system uncertainties. ADP offers a number of design options and one must think carefully through them to choose the right options for a given application. We have argued that, for process control problems, post-decision-state formulation offers the ability to use deterministic math programming solvers to be utilized, both off-line and on-line and therefore may be more convenient than the more conventional pre-decision-state formulation. In addition, the use of function approximators with nonexpansion properties offer stable learning. Robustness against over-extrapolation can be achieved through the use of a tailor-made penalty function. Finally, to achieve performance close to optimal ones, we recommend alternation between the value function update and simulation (or on-line implementation) to increase the sample set as the learning proceeds.

Though not discussed in this paper, there are a number of other application areas within process industries where ADP can prove to be a valuable tool, including resource allocation and inventory management (Pratikakis et al. (2008, 2009); Choi et al. (2004, 2006)), design and planning under uncertainty (Cheng et al. (2003)), scheduling of multiple controllers (Lee and Lee (2008)), and equipment / product inspection (Agrawal (2009)). Raised awareness of the ADP technique within the process systems engineering research community will undoubtedly bring forth additional applications that can benefit from it.

REFERENCES

Agrawal, R. (2009). *Planning and scheduling problems in manufacturing systems with high degree of resource degradation*. Ph.D. thesis, Georgia Institute of Technology.

Barto, A., Bradtke, S., and Singh, S. (1995). Learning to act using real-time dynamic programming. *Artificial Intelligence*, 72(1), 81–138.

Batina, I. (2004). *Model predictive control for stochastic systems by randomized algorithms*. Ph.D. thesis, Technische Universiteit Eindhoven.

Bellman, R. (1957). *Dynamic Programming*. Princeton University Press, New Jersey.

Bertsekas, D.P. (2005). *Dynamic Programming and Optimal Control, Vol. I*. Athena-Scientific, Belmont, MA, 3 edition.

Bertsekas, D. and Tsitsiklis, J. (1996). *Neuro-dynamic programming*. Athena Scientific, 1 edition.

Boyan, J.A. and Moore, A.W. (1995). Generalization in reinforcement learning: Safely approximating the value function. In *Advances in Neural Information Processing Systems 7*, 369–376. MIT Press.

Bradtke, S. (1993). Reinforcement learning applied to linear quadratic regulation. In *Adavances in neural information processing systems*, 295–302.

Cassandra, A., Kaelbling, L.P., and Littman, M.L. (1994). Acting optimally in partially observable stochastic domains. In *Twelfth National Conference on Artificial Intelligence*. Seattle, WA.

Cheng, L., Subrahmaniam, E., and Westerberg, A. (2003). Design and planning under uncertainty: issues on problem formulation and solution. *Computers & Chemical Engineering*, 27, 781–801.

Choi, J., Realff, M.J., and Lee, J.H. (2004). Dynamic programming in a heuristically confined state space: a stochastic resource constrained project scheduling application. *Computers & Chemical Engineering*, 8, 1039–1058.

Choi, J., Realff, M.J., and Lee, J.H. (2006). Dynamic programming in a heuristically restricted state space: stochastic supply chain management application. *AIChE Journal*, 52(7), 2473–2485.

De Farias, D.P. and Van Roy, B. (2003). The linear programming approach to approximate dynamic progrmaming. *Operations research*, 51(6), 850–865.

Gordon, G.J. (1995). Stable function approximation in dynamic programming. Technical report, School of Computer Science, Carnegie Mellon University.

Hastie, T., Tibshirani, R., and Friedman, J. (2008). *The Elements of Statisical Learning: Data Mining, Inference and Prediction*. Springer-Verlag, 2 edition.

Kaisare, N., Lee, J.M., and Lee, J.H. (2003). Simulation based strategy for nonlinear optimal control: application to a microbial cell reactor. *International journal of nonlinear and robust control*, 13, 347–363.

Konidaris, G. and Osentoski, S. (2008). Value function approximation in reinforcement learning using the fourier basis. Technical report, University of Massachusetts Amherst.

Laird, C.D. and Biegler, L.T. (2008). Large-scale nonlinear programming for multi-scenario optimization. In H.G. Bock, E. Kostina, H.X. Phu, and R. Rannacher (eds.), *Modeling, Simulation and Optimization of Complex Processes*, 323–336. Springer Berlin Heidelberg.

Lee, J.H. and Lee, J.M. (2006). Approximate dynamic programming based approach to process control and scheduling. *Computers & Chemical Engineering*, 30(10-12), 1603–1618.

Lee, J. and Ricker, N. (1994). Extended kalman filter based nonlinear model predictive control. *Industrial & Engineering Chemistry Research*, 33(6), 1530–1541.

Lee, J.M., Kaisare, N.S., and Lee, J.H. (2006). Choice of approximator and design of penalty function for an approximate dynamic programming based control approach. *Journal of process control*, 16, 135–156.

Lee, J.M. and Lee, J.H. (2004). Simulation-based learning of cost-to-go for control of nonlinear processes. *Korean Journal of Chemical Engineering*, 21(2), 338–344.

Lee, J.M. and Lee, J.H. (2005). Approximate dynamic programming-based approaches for input-ouptut data-driven control of nonlinear processes. *Automatica*, 41, 1281–1288.

Lee, J.M. and Lee, J.H. (2008). Value function-based approach to the scheduling of multiple controllers. *Journal of process control*, 18, 533–542.

Lee, J.M. and Lee, J.H. (2009). An approximate dynamic programming approach based approach to dual adaptive control. *Journal of process control*, 19, 859–864.

Ma, J. and Powell, W. (2009). A convergent recursive least squares policy iteration algorithm for multi-dimensional markov decision process with continuous state and action spaces. In *IEEE Conference on Approximate Dynamic Programming and Reinforcement Learning (part of IEEE Symposium on Computational Intelligence)*. Nashville, TN.

Ormoneit, D. and Sen, S. (2002). Kernel-based reinforcement learning. *Machine Learning Journal*, 49, 161–178.

de la Penad, D.M., Bemporad, A., and Alamo, T. (2005). Stochastic programming applied to model predictive control. In *44th IEEE Conference on Decision and Control, and the European Control Conference*, 1361–1366. Seville, Spain.

Peroni, C., Kaisare, N., and Lee, J.H. (2005). Optimal control of a fed batch bioreactor using simulation-based approximate dynamic programming. *IEEE Transactions in Control Systems Technology*, 13, 786–790.

Powell, W.B. (2007). *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley Series in Probability and Statistics. Wiley-Interscience.

Pratikakis, N., Realff, M.J., and Lee, J.H. (2008). A controlled exploration of the state space via an off-line adp approach: applications on stochastic shortest path and manufacturing systems. *Computers & Chemical Engineering*, in press.

Pratikakis, N., Realff, M.J., and Lee, J.H. (2009). Strategic capacity decisions in manufacturing using real-time adaptive dynamic programming. *Naval Research Logistics*, to be published.

Putterman, M.L. (1994). *Markov decision processes: discrete stochastic dynamic programming*. John Wiley and Sons, New York.

Roy, B.V., Bertsekas, D.P., Lee, Y., and Tsitsiklis, J. (1997). A neuro-dynamic programming approach to retailer inventory management. In *IEEE Conference on Decision and Control*, volume 4, 4052–4057.

Sabes, P. (1993). Approximating q-values with basis function representations. In *Fourth Connectionist Models Summer School*. Hillsdale, NJ.

Scokaert, P. and Mayne, D. (1998). Min-max feedback model predictive control for constrained linear systems. *IEEE Transactions of Automatic Control*, 43(8), 1136–1142.

Smart, W. and Kaelbling, L. (2000). Practical reinforcement learning in continuous spaces. In *17th international conference on machine learning*, 903–910.

Sutton, R.S. and Barto, A.G. (1998). *Reinforcement learning: an introduction*. MIT Press, Cambridge, MA.

Tesauro, G. (1992). Practical issues in tempora-difference learning. *Machine Learning Journal*, 8, 257–277.

Thrun, S. and Schwartz, A. (1993). Issues in using function approximation for reinforcement learning. In *Fourth Connectionist Models Summer School*. Hillsdale, NJ.

Tosukhowong, T. and Lee, J.H. (2009). Approximate dynamic programming based optimal control applied to an integrated plant with a reactor and a distillation column with recycle. *AIChE Journal*, 55(4), 919–930.

Tsitsiklis, J. and Roy, B.V. (1996). Feature-based methods for large scale dynamic programming. *Machine Learning Journal*, 22(1), 59–94.

Quality by Design in the Pharmaceutical Industry: Process Modelling, Monitoring and Control using Latent Variable Methods

Theodora Kourti

GlaxoSmithKline & McMaster University

E-mail: kourtit@ mcmaster.ca

Abstract: The pharmaceutical industry has entered a new era. Attention is now being paid to real time process monitoring, real time process control, continuous improvement of processes, and quick product technology transfer. Terms like Quality by Design, Design Space, Control Strategy, Process Analytical technology, Process Signature reflect the current state. Multivariate Statistical Analysis has played an integral part in several industries, enabling process understanding, process monitoring, utilization of real time analysers and real time product release. It is therefore appropriate to see it as an integral part of the pharmaceutical industry effort to address issues like Design Space, Control Strategy, real time process signature monitoring, process understanding and correct technology transfer. In this work it is demonstrated that multivariate, data based statistical methods play a critical role in providing solutions to these issues. From determining the acceptability of raw material entering the plant to ensuring quality of the product that leaves the plant, the multivariate analysis philosophy should govern all the operations that take that raw material and convert it to a final product in a cost efficient way, while meeting safety and environmental constraints, from development to manufacturing to site transfer.

Keywords: Multivariate Process Monitoring, Design Space, process analytical technology, multivariate statistical process control, scale – up, latent variables, process understanding

1. INTRODUCTION

"The pharmaceutical industry has a little secret: Even as it invents futuristic new drugs, its manufacturing techniques lag far behind those of potato-chip and laundry-soap makers" proclaimed the Wall Street Journal in 2003 (Abboud and Hensley, 2003). The article went on to explain that "in other industries, manufacturers constantly fiddle with their production lines to find improvements. But FDA regulations leave drug- manufacturing processes virtually frozen in time. As part of the drug- approval process, a company's detailed manufacturing plan -- and even the factory itself -- must pass FDA muster. After approval, even a tiny change to how a drug is made requires another round of FDA review and authorization, requiring time and paperwork. The process discourages updating by the companies, which worry they will face a production delay that could cost them heavily". The article mentioned FDA as a regulatory agency because it was published in the USA, but similar were the situations with Pharmaceutical companies and other regulatory bodies around the world.

A lot of changes have happened since that article was published and the pharmaceutical industry has entered a new era. The FDA guidance on Process Analytical Technology (PAT) was introduced in 2004 which aims to improve product quality and process performance (manufacturing efficiency) in the pharmaceutical industry; it describes PAT as: systems for the analysis and control of manufacturing processes based on timely measurements during processes of critical quality parameters and performance attributes of raw and in-process materials and processes to assure acceptable end product quality at the completion of the process. (Guidance for Industry: PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. FDA. September 2004)

The introduction of concepts like Quality by Design, Design Space and Control Strategy are also examples of this change. These terms are defined as follows by the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH is a unique project that brings together the regulatory authorities of Europe, Japan and the United States and experts from the pharmaceutical industry in the three regions to discuss scientific and technical aspects of product registration):

Quality-by-Design (QbD) is defined as a systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management (ICH, 2008a).

Design Space is the multidimensional combination and interaction of input variables (e.g., material attributes) and process parameters that have been demonstrated to provide assurance of quality (ICH, 2008b).

Control Strategy: a planned set of controls, derived from current product and process understanding that ensures (good) process performance and product quality. The controls can include parameters and attributes related to drug substance and drug product materials and components, facility and equipment operating conditions, in-process controls, finished product specifications, and the associated methods and frequency of monitoring or control (ICH 2008a, 2008b).

The above definitions and actions indicate that the regulatory framework for the Pharmaceutical industry is changing.

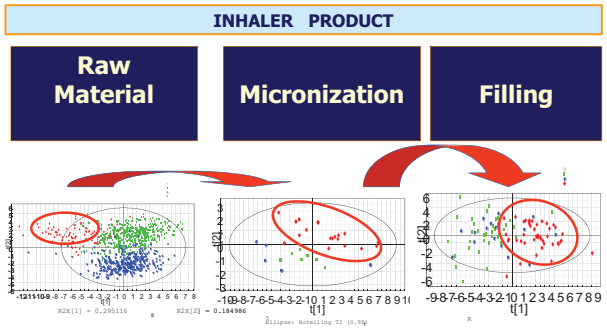


Figure 1. Projection space. Raw material properties, micronization properties and filling performance are projected on a latent variable space. Product batches produced from similar raw material (red) have similar filling performance.

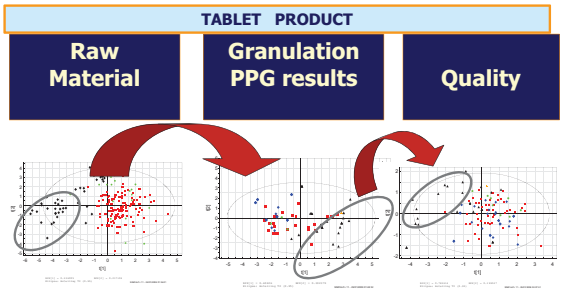


Figure 2. Tablet Product. Batches produced from raw material with similar characteristics, have similar final quality.

The ICH definition of the design space reflects a well known concept, namely that variability in the input of a process will be transferred to the quality of the final product (output) if the process is not controlled to compensate for such variability. Despite the fact that the concept is well known, it requires new ways of thinking in the pharmaceutical industry that was used to dealing with "fixed" processes, as described in the above mentioned Wall Street Journal article.

In this work the role of multivariare statistical methods in modelling, process control and monitoring under this new regulatory framework will be discussed. Multivariate latent variable methods are shown to be most suitable for process understanding, modelling for Design Space, multivariate statistical process control (MSPC), process control and product transfer. The use of these methods for the development of the Design Space for multi-unit operations will be illustrated in a case where the Tablet Quality is related to API, Excipients, Granulation, Drying and Compression parameters. Examples of how the Control Strategy can be derived from such models will also be shown.

Other topics like Process Signature and MSPC, application of soft sensors, relation of design space to clinical relevance as well as quality by design for analytical methods will be discussed.

2. LATENT VARIABLE METHODS

Latent variables exploit the main characteristic of process databases, namely that although they consist of measurements on a large number of variables (hundreds), these variables are highly correlated and the effective dimension of the space in which they move is very small (usually less than 10 and often as low as 2). Typically only a few process disturbances or independent process changes routinely occur, and the hundreds of measurements on the process variables are only different reflections of these few underlying events. For a historical process dataset consisting of a $(n \times k)$ matrix of process variable measurements $\mathbf{X}$ and a corresponding $(n \times m)$ matrix of product quality data $\mathbf{Y}$, for linear spaces, latent variable models have the following common framework :

$$\mathbf{X} = \mathbf{T}\,\mathbf{P}^{\mathrm{T}} + \mathbf{E} \qquad (1)$$

$$\mathbf{Y} = \mathbf{T}\,\mathbf{Q}^{\mathrm{T}} + \mathbf{F} \qquad (2)$$

where $\mathbf{E}$ and $\mathbf{F}$ are error terms, $\mathbf{T}$ is an $(n \times A)$ matrix of latent variable scores, and $\mathbf{P}$ $(k \times A)$ and $\mathbf{Q}$ $(m \times A)$ are loading matrices that show how the latent variables are related to the original $\mathbf{X}$ and $\mathbf{Y}$ variables. The dimension A of the latent variable space if often quite small and determined by cross-validation or some other procedure.

Latent variable models assume that the data spaces ($\mathbf{X}$, $\mathbf{Y}$) are effectively of very low dimension (i.e., non-full rank) and are observed with error. The dimension of the problem is reduced by these models through a projection of the high-dimensional $\mathbf{X}$ and $\mathbf{Y}$ spaces onto the low-dimensional latent variable space $\mathbf{T}$, which contains most of the important information. By working in this low-dimensional space of the latent variables (t_1, t_2, ... t_A), the problems of process analysis, monitoring, and optimization are greatly simplified.

Multivariate Statistical Process Control is possible utilizing latent variable methods. The following charts are used:

The Hotelling's T^2 for scores (derived either from PCA or PLS models on typical production) is calculated as:

$$T_A^2 = \sum_{i=1}^{A} \frac{t_i^2}{\lambda_i} = \sum_{i=1}^{A} \frac{t_i^2}{s_{t_i}^2} \qquad (3)$$

where $s_{t_i}^2$ is the estimated variance of the corresponding latent variable t_i. This chart essentially checks if a new observation vector of measurements on k process variables projects on the hyper-plane within the limits determined by the reference data.

As mentioned above the A principal components explain the main variability of the system. The variability that cannot be explained forms the residuals (Squared Prediction Error, SPE). This residual variability is also monitored and a control limit for typical operation is being established. By monitoring the residuals we test that the unexplained disturbances of the system remain similar to the ones observed when we derived the model. When the residual variability is out of limit, it is usually an indication that a new set of disturbances have entered the system; it is necessary to identify the reason for the deviation and it may become necessary to change the model.

SPE_X is calculated as :

$$SPE_X = \sum_{i=1}^{k} (x_{new,i} - \hat{x}_{new,i})^2 \qquad (4)$$

where $\hat{\mathbf{x}}_{new}$ is computed from the reference PLS or PCA model. Notice that SPE_x is the sum over the squared elements of a row in matrix **E** in equation (1). This latter plot will detect the occurrence of any new events that cause the process to move away from the hyperplane defined by the refence model. The calculation of the limits for the charts is discussed in Kourti (2009).

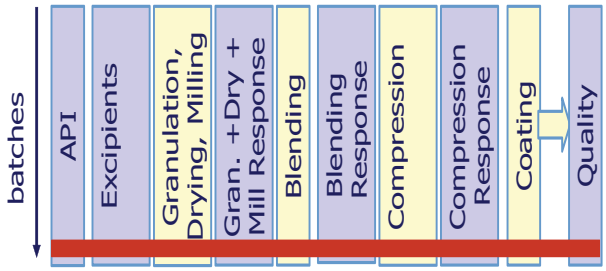


Figure 3. The quality can be modelled as a function of input material and process parameters

These two charts (T^2 and SPE) are two complementary indices; together they can give a picture of the state of the system at a glance. With this methodology, the hundreds of measurements collected from the process variables at each instant in real time are translated into one point for the T^2 chart and one point for the SPE chart (these two points summarize the process at that instant). As long as the points are within their respective limits everything is in order. Once a point is detected out of limit, then the so called *contribution plots* can be utilized that give us a list of all the *process* variables that mainly contribute to the out of limit point, and hence allow us to diagnose the process problem immediately. Contribution plots can be derived for out of limit points in both charts.

A detailed discussion on latent variable methodology for modelling and process monitoring can be found in Kourti (2002, 2005, 2009). Experiences from industrial practitioners can be found in Miletic et al (2004, 2008).

3. PROCESS UNDERSTANDING - EFFECT OF RAW MATERIAL ON FINAL QUALITY

The effect of raw material characteristics in the process performance, if the process operating conditions remain fixed, is demonstrated for an inhaler product utilizing multivariate projection space in Figure 1. The raw material is characterized by several physical and chemical properties. Raw material is produced at three supplier locations and depending on its origin, the data are coloured red, green and blue. The raw material properties are within univariate specifications, at all locations. Projected on a multivariate space, however, they form three clusters, indicating that in a multivariate sense the material possesses slightly different characteristics depending on the location it was produced (covariance structure changes with location). The material properties after micronization are projected on principal components and it can be observed that the material with red coloured origin projects on a different location than the green and blue. The filling performance of the material originating from the red location is different than the rest of the material. A note here that although the control ellipses shown are set by default in the vendor software, they are not interpretable when there is clustering; the assumptions for the calculation of these ellipses are for process monitoring and not for process exploration where there is intentional variation such that introduced by design of experiments.

4. DESIGN SPACE MODELLING

The effect of raw material on the quality as it propagates through different unit operations is shown for a tableting process in Figure 2. When the raw material properties have certain characteristics (marked black) the material projects on a different area. The properties of granules produced from raw material with such characteristics (black) are different from the rest, and the final quality also shows differences.

The difference in the quality can be theoretically explained based on the physical phenomena that govern the whole process. The idea of the design space is to express these phenomena by a model.

The design space can be established as a model that relates input material and process parameters to quality. The model may be theoretical (based on first principles) or empirical, derived from design of experiments or, a hybrid. Together with the model one has to specify the range of parameters for

which the model has been verified. The model may cover one unit operation or a series of unit operations.

The design space for the entire tableting process can be derived by relating quality to the raw material properties as well as to the process parameters of the unit operations (Figure 3). One row in the database depicted in Figure 3 would include the process conditions and quality experienced by the material as it is processed through the units. Multivariate projection methods can be used for the empirical modelling.

It should be emphasised here that the Design Space is a collection of models that relate 1) the final quality to all previous units, raw material and intermediate quality 2) intermediate quality to previous unit operations and raw material.

The empirical models derived are causal and based on carefully designed experiments (DOE). Some DOE's will also be necessary to estimate parameters even if mechanistic models are used.

Batch processes are very common in the pharmaceutical industry. Empirical methods for modelling and monitoring batch processes are discussed in Nomikos and MacGregor (1994) and in Kourti (2003).

Foundations for multiblock analysis necessary for multi unit operation systems can be found in Westerhuis et al (1998).

The level of detail in the models varies depending on the depth of process understanding one wishes to achieve. For example the variable trajectories of a granulation may be described by summary data (min, max, slopes, etc) or by the full variable trajectories aligned against time or another indicator variable (Kourti, 2003).

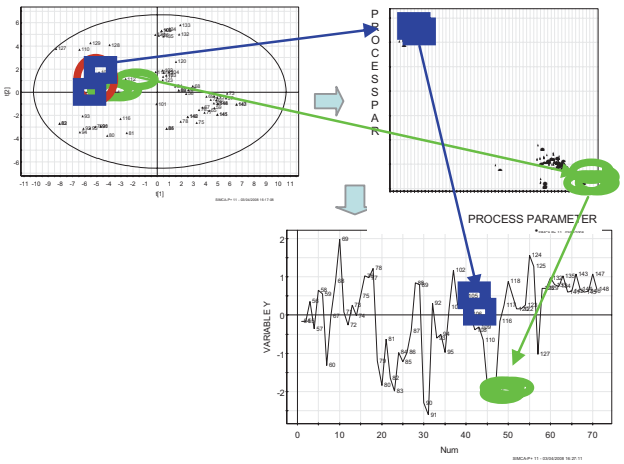


Fig. 4 Control Strategy using Projection Space.

5. CONTROL STRATEGY

Based on the process understanding derived from the design space, control strategy can be derived to assure final quality.

An example in Figure 4 is used to illustrate the new concepts. Control Strategy is devised once the Design Space is established. The example here illustrates a feed forward control scheme for Unit N+1 based on input information on the "state-of-the-intermediate product" from unit N. The settings are calculated and adjusted such that the target value for Quality Y is met.

A multivariate model was built (from batch data) to relate product quality to the process parameters of unit N+1 and the "state-of-the-intermediate product" from Unit N. From this model, a quantitative understanding was developed showing how process parameters in N+1 and the state-of-the-intermediate product from N interact to affect Quality.

Control of batch processes on multivariate space is discussed by Flores-Cerrillo and MacGregor (2004), while product transfer is discussed by García-Muñoz et al (2005).

6. PROCESS SIGNATURE AND MSPC.

It is known from other industries, that sometimes it is not sufficient to characterize a product with "end point quality measurements". The reason is that for some products we do not measure all the possible quality properties (example, downstream processability). The same "measured" quality properties may sometimes be achieved by taking different process paths. In these situations, these different paths may affect the properties that are not measured (i.e. processability). To achieve consistency in all the product properties (measured quality and ability to process down the stream) the process conditions (path to end point) must also be kept in statistical control. When this is not the case, although the measured product properties are on target, the properties that determine other characteristics (i.e., the processability of the product) may not be within acceptable limits. Therefore the "process path to the end point" must also be examined. This "process path to the end point" is also discussed in the European Regulatory Perspective (Graffner, 2005) where it is reported that "during discussions within the industry, the term process signature has been mentioned regularly". To get a common understanding of this, the EU PAT Team had invited public comments on the following definition: "A collection of batch specific information that shows that a batch has been produced within a design space of the product." The EU PAT team mentions as examples of process signatures the amount of water added in relation to time (wet massing), air flow rate, and bed temperature during fall rate drying (fluidized bed drying). They concluded that their understanding is that there is no unique process signature, but instead a family of process signatures with common characteristics (salient features).

It should be pointed out here that the process signature in the multivariate statistical process control context is nothing else

but the two multivariate indices Hotelling's T^2 and SPE. As a matter of fact, these indices take it to account not one feature (e.g., water addition rate or, drying rate) but the combination of all the variables affecting the process and product and their correlations both at each time interval but also their time correlations for the duration of the process (auto and cross correlations for the entire batch). They are therefore a more powerful tool to describe the "overall process signature".

Furthermore, these indices can be directly related to the concept of the design space, as outlined here. The design space model relates raw material characteristics, process conditions and quality. Given the characteristics of the raw material and the desired quality, the design space model can be solved to determine appropriate operating conditions. Maintaining T^2 and SPE within their good operation limits for these appropriate process conditions is nothing more than ensuring that the operation is within the design space.

7. PAT and SOFT SENSORS

Accurate on-line measurements of quality variables are essential for the successful monitoring and process control. However, due to measurement difficulties, sometimes process variables may be used to "infer" product quality in real time and therefore replace an analyzer. This is the idea of soft sensors. In many monitoring and control situations we are often lacking real time sensors capable of measuring many of the responses of interest, because the measurement equipment for such quality variables may be very expensive, or difficult to put on-line, or costly to maintain. As a result we often try to develop soft sensors or inferential models which use other readily available on-line measurements such as temperatures, and can be used to infer the properties of interest in a real time manner. In a recent paper it was demonstrated through application to a benchmark simulation of a fed-batch fermentation process that mutli-way PLS can provide accurate inference of quality variables, such as biomass concentration, that are often difficult to measure using on-line sensors. It was also demonstrated that the same PLS model can be used to provide early detection and isolation of fault conditions within a fermenter (Zhang and Lennox, 2004).

The soft sensors can either replace the hardware sensor (analyzer) or be used in parallel with it to provide redundancy and verify whether the hardware sensor is drifting or has failed; when used in parallel the soft sensor will either estimate the property and compare its value with that of the analyser, or it will keep track of the correlation between the analyser reading and the process measurements. An example where a soft sensor is used to assess the reliability of an analyser was presented in Kourti (2005). Latent variable modeling was used for this purpose.

This idea of using process measurements as a safety net to verify analyser reliability but also to monitor an index of wellness for the process, to check for unforeseen disturbances, is a crucial and important issue for real time release (Kourti, 2006a).

8. INTEGRATION OF CLINICAL TRIALS

As more complex structures of data are being generated, the multivariate analysis offers great opportunities for information integration and analysis.

Manufacturing Data as well as patient histories can be integrated and then incorporate into design space the clinical trial responses. (Kourti 2006b).

Figure 5 shows an example of the possibilities that can be explored. Quality in product Y can be related to past information of raw materials, preprocessing and holding times, the type of the vessel used, the operator that run the process, and other recipe information as well as process measurement trajectories and analyzer information.

The quality Y (and details of manufacturing) as well as the patient medical histories and clinical responses can be used to establish a better understanding of the design space.

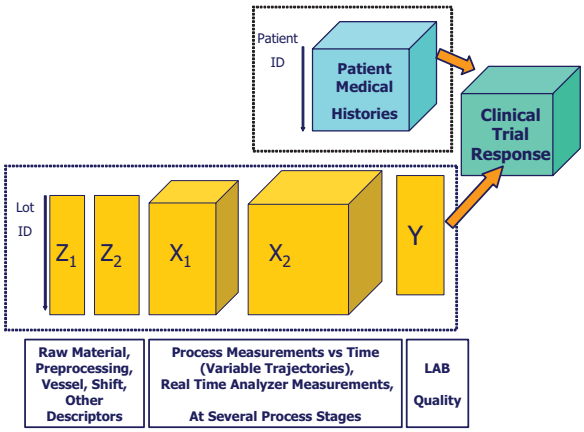


Figure 5 : Examples of complex data structures emerging in industry, that can be mined for a wealth of information. (Kourti, 2006b).

9. QUALITY BY DESIGN IN ANALYTICAL METHODS

The methodology described for design space can be applied in analytical methods. Chromatography, is a laboratory method but also a Unit operation in Bio – Pharmaceuticals.

Process Transfer ideas can be also applied in method transfer ideas, that is method transfer – and site transfer could be treated with similar principals (García-Muñoz et al., 2003)

ACKNOWLEDGEMENTS

I would like to thank John F. MacGregor, McMaster University and Gordon T. Muirhead, GlaxoSmithKline, for their support and dedication to innovation.

REFERENCES

Abboud, L., and Hensley. S (2003) Factory Shift: New Prescription For Drug Makers: Update the Plants; After Years of Neglect, Industry Focuses on Manufacturing; FDA Acts as a Catalyst; The Three-Story Blender. Wall Street Journal. Eastern edition). New York, N.Y.: Sep 3, 2003. pg. A.1, The Wall Street Journal

Flores-Cerrillo J., and MacGregor J. F. (2004) Control of batch product quality by trajectory manipulation using latent variable models, Journal of Process Control 14, 539 -553.

García-Muñoz, S, MacGregor, J.F., Kourti, T. (2005) Product Transfer between Sites using Joint Y_PLS. Chemometrics and Intelligent Laboratory Systems, 79, 101-114.

Graffner, C (2005). PAT- European Regulatory Perspective, The Journal of Process Analytical Technology, 2 (6) 8-11.

ICH (2008a). International Conference on Harmonization, ICH Draft Step 4, Q8 (R1) Pharmaceutical development revision 1, 11-November-2008.

(ICH, 2008b) International Conference on harmonization of technical requirements for registration of pharmaceuticals for human use, ICH harmonized tripartite guideline, pharmaceutical quality system Q10 Step 4, 9-June-2008.

Kourti, T. (2002) Process analysis and abnormal situation detection: from theory to practice. IEEE control systems, 22 issue 5 : 10-25

Kourti, T. (2003) Multivariate dynamic data modelling for analysis and statistical process control of batch processes, start–ups and grade transitions. J. Chemometrics, 17, 93-109

Kourti, T. (2005) Application of Latent Variable Methods to Process Control and Multivariate Statistical Process Control in Industry. International Journal of Adaptive Control and Signal Processing; 19, 213-246.

Kourti, T., (2006a). Process Analytical Technology and Multivariate Statistical Process Control. Index of wellness of product and process. Part 3. Process Analytical technology, volume 3 (3), 18-24.

Kourti, T. (2006b) Process Analytical Technology Beyond Real-Time Analyzers: The Role of Multivariate Analysis Critical Reviews in Analytical Chemistry, Critical Reviews in Analytical Chemistry, 36:257–278.

T. Kourti (2009) Multivariate Statistical Process Control and Process Control, Using Latent Variables. In: Brown S, Tauler R, Walczak R (eds.) Comprehensive Chemometrics, volume 4, pp. 21-54 Oxford: Elsevier

Miletic, I., S. Quinn, M. Dudzic, V. Vaculik, M. Champagne. (2004) Journal Process Control, 14, 821-836.

Miletic, I.; Boudreau, F.; Dudzic, M.; Kotuza, G.; Ronholm, L.; Vaculik, V.; Zhang, Y. (2008) Experiences in Applying Data-Driven Modelling Technology to Steelmaking Processes. Can. J. Chem. Eng. 86, 937–946.

Nomikos, P. and J.F. MacGregor (1994)"Monitoring of batch processes using multi-way principal component analysis," AIChE Journal, 40 , no. 8: 1361-1375.

Westerhuis, J.A., Kourti, T., and MacGregor, J.F. (1998) Analysis of multiblock and hierarchical PCA and PLS models. Journal of Chemometrics, 12, 301-321

Zhang, H., Lennox, B. (2004) Integrated condition monitoring and control of fed-batch fermentation processes, Journal of Process Control 14, 41–50.

Plantwide Optimizing Control for the Bio-ethanol Process

Silvia Ochoa*, Jens-Uwe Repke*, Günter Wozny*

**Chair of Process Dynamics and Operation, Berlin Institute of Technology, Sekr. KWT9, Strasse 17. Juni. 135, Berlin 10623, Germany*

Abstract: In this work, the Plantwide Control (PWC) problem of a continuous bio-ethanol process is investigated from a Plantwide Optimizing Control (PWOC) perspective. PWOC addresses the plantwide control problem integrating real-time optimization and control for optimal operation. Two different PWOC approaches have been considered: A Single-Layer Direct Optimizing Control approach (one-layer) and a Multi-Layer without Coordination approach (two-layer). The performance of these two PWOC approaches is compared with more traditional Decentralized architectures, demonstrating the benefits of using Plantwide Optimization-based Control strategies in bioprocesses.

Keywords: Plantwide Control, Optimizing Control, Dynamic Real Time Optimization, Ethanol.

1. INTRODUCTION

Nowadays, bioprocess industry is an important part of the worldwide economy. Specifically, the bio-ethanol industry has experienced a significant growth in the last years since ethanol, as an environmentally friendly fuel, is considered an attractive alternative energy source. Ethanol production has been continuously improved in very different ways in order to assure the economical and environmental feasibility of the process. Examples of these improvements include purification technologies for reducing energy consumption during the separation of the ethanol-water mixture (Arifeen et al., 2007), and genetic modifications of the microbial strains for building more ethanol-tolerant yeast and strains capable of carrying out simultaneously saccharification and fermentation tasks (Olofsson et al., 2008). From a process control point of view, different works have been done regarding the modelling, estimation, and control problem for the fermentation stage. However, only relatively few works (e.g. Meleiro et al., 2008, Costa et al., 2001) have addressed the control problem from the Process Systems Engineering point of view, considering the process as an integrated dynamic production system taking into account more than one single process unit (i.e. accounting for interactions between fermentation, cells recycle and flash units). In this work, the Plantwide Control (PWC) problem for the ethanol process is addressed as a large-scale real-time dynamic optimization problem due to the following facts: the nature of the process is highly nonlinear and dynamic; the process is characterized by the coupling of slow and fast dynamics; interactions between different operating units can not be neglected; and finally, economical feasibility of the process can be effectively assured if this is the main control objective of the plantwide strategy. PWC has attracted the attention of the process control community for more than 40 years, since the pioneer work by Buckley (1964). Through these years, different architectures have been used for tackling the problem of controlling a complete process. The intention of this section is to present a brief review of the several options reported for addressing PWC. A proposal of classification for different PWC architectures is shown in Fig 1, which agrees in some points with that presented by Scatollini (2009).

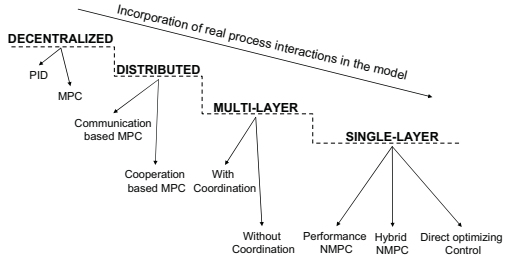


Fig 1. Plantwide Control Architectures

In the *Decentralized* scheme, many different individual regulators are used in the process without sharing any kind of information between them (i.e. each controller operates independently of the others), even though the selection of the manipulated and controlled variables might take into account the interactions in the process. The decentralized approach consists usually of SISO PID loops, although MPC controllers can also be used. As noted by Stephanopoulous and Ng (2000), most of the research activities in the topic of PWC up to year 2000, addressed the PWC problem as the selection of the best input-output pairing for the formation of SISO PID loops. Of course, as in any rule there are exceptions, and one of the most relevant examples in this case is the work by Garcia and Morari (1984), in which a multivariable control scheme based on a *Multi-layer* PWC architecture was proposed for controlling a benzene plant. Some of the many works that have addressed the PWC in a decentralized manner are Araujo et al. (2007), Larsson et al. (2003), Robinson et al. (2001), Zhen et al. (1999), Lausch et al. (1998), Luyben et al. (1997) and McAvoy and Ye (1993). Most of the works in the remaining three architectures shown in Fig 1 make use of a multivariable controller. Two main reasons motivated to move the PWC problem from the

paradigm of decentralized PID towards different alternatives: the performance limitations of the decentralized architecture, and the broad industrial impact of the Model Predictive Control (MPC) framework (Venkat et al. 2007). In the *Distributed* architecture some information is exchanged between the multiple MPC controllers. Two Distributed-MPC approaches worthy of mention are the communication- and cooperation-based, which mainly differ that in the first, each controller has a local objective function, whereas in the latter the objective function in each controller is a copy of the total objective function for the complete plant (Rawlings and Stewart, 2008). Representative works addressing the PWC from the Distributed perspective are those by Sun and El-Farra (2008), Venkat et al. (2007), Mercangöz and Doyle (2007) and Venkat (2006). *Multi-layer* architecture is a hierarchical structure that follows the guidelines given by Findeisen et al. (1980), which classified the hierarchical control into multilayer and multilevel. According to Findeisen's work, in the multilayer case, the control of a system is split into algorithms (layers), whereas in the multilevel case control is divided into local goals and the action of each local control unit is coordinated by an additional supremal unit. In Fig 1 it is proposed to sub-divide the Multi-layer (or hierarchical) architecture into: *With Coordination* (denoted as Multilevel approach by Findeisen) and *Without Coordination*. Multilayer architectures should be composed by at least two different layers, in which the task of finding the control actions that should be applied to the process is split usually into: a *Real Time Optimization* (RTO) layer that computes optimal set point values for the controlled variables, and a *Control* layer which is in charge of tracking the optimal set point values (Kadam et al., 2002). In the control layer, a PID or MPC controller can be used (Kadam and Marquardt, 2004). It is important to notice that as mentioned by Biegler and Zavala (2009), the "connection" between RTO and MPC layers may suffer inconsistencies due to model mismatch (non-linear steady state vs. linear dynamic) and conflicting objectives. Therefore, in the last years a proposal for replacing the steady state RTO by a Dynamic Real Time Optimization (D-RTO) layer has emerged (Kadam et al. 2003; Kadam and Marquardt, 2004). On the other hand, regarding the *Multi-layer with coordination* architecture, the reader is referred to the work by Tosukhowong et al. (2004) in which a coordination collar is used to find for each MPC a locally feasible set point close to the global solution found by the RTO layer; and to the work by Cheng et al. (2007), in which a price-driven method is used for coordination between the RTO and the MPC layers. Additionally to the references already mentioned, the following works include examples of PWC using *Multi-layer* architecture: Ochoa et al. (2009), Kadam and Marquardt (2007), Lu (2003), Duvall and Riggs (2000) and Ying and Joseph (1999). A final mention should be done regarding the difference between the Multilayer with coordination and the Distributed architectures. As both schemes include coordination, in the Distributed case the coordination consists on exchanging some information between the local MPCs, whereas in the Multilayer with coordination, the local MPCs are not communicated between them but communicated to the RTO layer. The last PWC architecture in the classification shown in Fig 1 is the *Single-layer* scheme. Despite the very common belief that a *Single-layer* or centralized structure will be intractable for PWC (Venkat et al., 2007), in the last years some publications from both the industrial and the academic side have shown that such monolithic approach it is not only possible to implement but also gives very good results from an economic point of view (Bartusiak, 2007; Zavala et al., 2007; Franke and Doppelhamer, 2007). Works using this architecture solve online a moving horizon optimization problem, but differ in the type of objective function optimized. A first group of works denoted as *Performance NMPC* uses a performance-type objective function (in which mainly the tracking of a reference value is penalized). The second scheme includes besides the performance term, an economic penalization term in the formulation of the objective function and therefore it is denoted here as *Hybrid NMPC* (Economic+Performance). The final scheme denoted in the literature as *Direct Optimizing Control* (Engell, 2007) uses a pure economic objective function in which the usual control specifications enter as constraints and not as set points, and therefore no tracking term is penalized. References showing examples of the application of the Single-layer architecture are: Biegler and Zavala (2009), Roman et al (2009), Ochoa et al. (2009), Engell (2007), Franke and Doppelhamer (2007), Zavala et al. (2007), Bartusiak (2007), Manenti and Rovaglio (2007), Franke and Vogelbacher (2006), Toumi and Engell (2004) and Jockenhövel et al (2003). The main purpose of this paper is to present a novel approach for the PWC of the bio-ethanol process, in which the main control objective is to maximize the profitability of the whole process. The paper is organized as follows: Section 2 gives a description of the ethanol continuous process from starch, including a brief description of the relevant works that have addressed the control of the process considering it as composed of more than one process unit. Section 3 presents the Plantwide Optimizing Control (PWOC) concept proposed in this work and describes the main steps of this approach. A new method for shrinking the search region during the optimization problem that arises when applying PWOC is proposed in Section 4. The Multi-layer without coordination and the Single-layer direct optimizing architectures are used for addressing the PWC problem in the continuous bio-ethanol process. These approaches are compared in Section 5 to conventional decentralized architectures.

2. BIO-ETHANOL PRODUCTION PROCESS

The case study addressed is based on the extractive alcoholic fermentation process shown in Fig 2. A detailed description of this process is found elsewhere (Meleiro et al., 2008). The process includes saccharification, fermentation, cells recycle, flash separation, distillation and rectification. The end product considered is the ethanol obtained at the top of the rectification column, which in a further step must be sent to a dehydration unit (e.g. molecular sieves). A nonlinear dynamic model of the process has been simulated using Simulink®. The model consists of a nonlinear DAE system comprising 69 differential states and 173 algebraic equations. pH, temperature and liquid levels are regulated as usually done in industry by means of local SISO loops, which in the

following will be denoted as *basic control*. After closing these basic loops, 13 input variables are left, 3 of which are identified as disturbances: starch (S_0), enzymes (Enz_1) and fresh yeast concentration (X_3) fed into the process. The remaining 10 inputs are available for improving the control strategy in the process. The process with its basic level control loops is shown in Fig. 2 (for simplicity, the pH and temperature loops are not shown). In addition to the basic loops, an internal biomass control strategy (Ochoa et al., 2009) is also shown. The combination of the traditional *basic control loops* with this biomass internal strategy is denoted in the following as *Local Control Strategy*. Two main reasons motivated implementing the biomass control. First, an optimal biomass concentration in the fermentor should be always guaranteed in order to avoid a misuse of the substrate if a higher concentration than the optimal is available. Additionally, if biomass concentration is below the optimum, a slower metabolite production rate will occur, affecting the productivity of the process. Second, yeast is only involved in a closed mass loop comprising fermentation, filter and cells recycle; i.e. no biomass is found on the streams up the fermentor nor downstream the filter. As already mentioned, the process has 10 manipulated variables available for improving the control strategy; however, 3 of them (F_3, F_7, F_{10}) are used as manipulated variables in the biomass strategy. The remaining 7 manipulated variables (F_0, F_1, F_{13}, VB_1, R_1, VB_2, R_2, which are the starch input flow, enzymes input flow, recycle flow from the flash to the fermentor and vapour and reflux rates for each column) are potential manipulated variables denoted as "Plantwide variables".

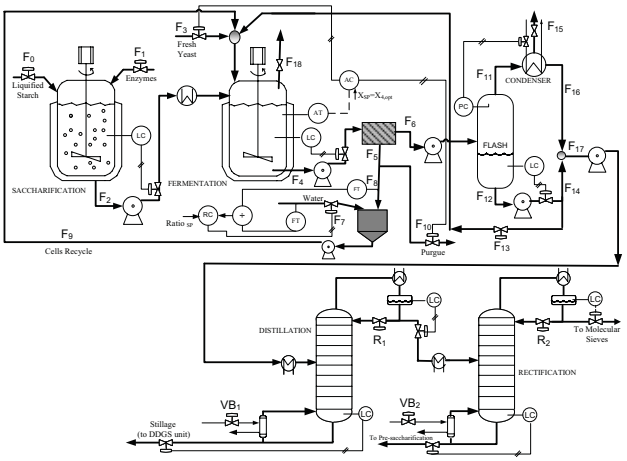


Fig 2. Bio-ethanol Process from Starch: Local Control Strategy (Basic loops + Internal Biomass Strategy).

Additionally, it should be noticed that despite the rapid increase of the bio-ethanol industry in the last 30 years and the high economic risk that this industry faces, no much effort has been done in order to improve the efficiency of the process from the optimization and control points of view. Several works have been published regarding mainly the control of the fermentation unit in the process, but to the author's knowledge, only few works have addressed the control of the process considering more than the fermentation stage. Costa et al. (2001) used Dynamic Matrix Control (DMC) for controlling the substrate or the product concentrations in the fermentor manipulating the substrate input flow or the cells recycle rate. A second contribution by Costa et al. (2002), proposes a SISO NMPC for controlling the substrate concentration in the fermentor, manipulating the substrate input flow. Meleiro et al. (2008) presented a multivariate NMPC to control simultaneously the ethanol, substrate and biomass concentrations in the fermentor. Although the process modelled in these works considers interactions fermentor-cells recycle-flash, the control task is still focused on tracking or regulating the main state variables in the fermentor without considering the optimal economic operation of the whole process. Finally, Bartee et al.(2008), propose using MPC for controlling the process including milling, cooking, distillation etc.; however, no details regarding algorithms and implementation are given.

3. PLANTWIDE OPTIMIZING CONTROL

Online optimizing control optimizes an economic objective over a finite moving horizon during plant operation based upon a rigorous nonlinear dynamic model (Küpper and Engell, 2008). Plant limitations and product specifications are included in the optimization as constraints. This definition is used in this section as key concept for developing the basic steps of a Plantwide Optimizing Control (PWOC) approach. PWOC addresses PWC as a nonlinear dynamic online problem, in which the available manipulated variables in the process are used for achieving maximum profitability in the plant in spite of disturbances. In this way, PWOC calculates optimal values for the set of selected manipulated variables, in order to maximize a Plantwide Profitability Objective function Φ, instead of maintaining a set of controlled outputs at predefined set points. A key feature of PWOC is that input-output pairing is avoided because the output actually controlled in the process is the Plantwide Profitability and the available manipulated variables are simultaneously used for satisfying that purpose. Online optimizing control has been gaining increasing attention in the last years in different chemical process applications (Engell, 2007). However, not much work has been reported in the open literature on the on-line optimizing control of bioprocesses. In this work PWC of a bioprocess is addressed from an optimizing control perspective, considering a large-scale nonlinear Dynamic Real-Time Optimization (D-RTO) problem. The proposed PWOC approach comprises six main stages, as shown in Fig 3. In the following, a description of each stage is presented.

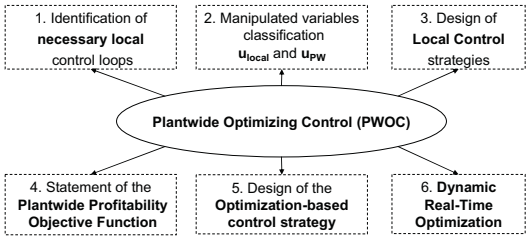


Fig 3. Plantwide Optimizing Control Stages

Stage 1: Identification of necessary control loops

Even though the goal of any chemical or biochemical process is to return a maximum profit, there are additional control

objectives that should be taken into account before establishing a PWOC structure for satisfying this economic goal. These objectives are mainly related to safe operation, equipment and environmental protection and should be achieved independently of the economical performance of the plant, i.e. by using local control loops.

Stage 2: Classification of the Manipulated Variables

Manipulated variables in the process can be used in the local control loops or for the PWOC of the process. Those manipulated variables used for satisfying the local control set points are denoted as *Local manipulated* (u_{Loc}), whereas the *Plantwide manipulated* variables (u_{Pw}) are those that remain available after selecting the u_{Loc}, and that are used for maximizing the plantwide profitability objective function.

Stage 3: Design of Local Control Strategies

After identifying the necessary local control loops in the process and the local manipulated variables required for satisfying the control objectives at the local control loops, it is then necessary to address the design of those local loops (i.e. pairing manipulated-controlled variables, selection of controller type, controller tuning, etc.), as traditionally done.

Stage 4: Statement of Plantwide Profitability Function (Φ)

The next step is to establish a plantwide profitability function Φ and its constraints, in order to formulate a D-RTO problem. Statement of the objective function Φ will depend upon the specific process addressed. However, it may contain terms related to productivity of the process, raw materials and energy consumption, economic losses, etc. Constraints in the optimization problem are determined by plant and product specifications, and by limitations in the state and input variables. Since PWOC addresses the optimizing control problem for a complete plant over a finite moving horizon during plant operation, it is also important to select the prediction horizon Δt_{opt} over which the objective function and constraints will be evaluated. Δt_{opt} should not be shorter than the characteristic response time of the slowest relevant dynamic in the process (to avoid unexpected long-term performance deterioration), while at the same time it should be as short as possible to minimize computational load.

Stage 5: Design of the Optimization-Based Control Strategy

PWOC is addressed here using two different architectures: *Single-Layer Direct Optimizing Control* and *Multi-Layer without Coordination*. These frameworks will be referred in the following as the one-layer and the two-layer approaches, respectively. The structures for both approaches are shown in Fig. 4. A detailed description of the building blocks for each framework can be found elsewhere (Ochoa et al., 2009). Comparing the schemes for the two frameworks (Fig. 4), it is possible to see that both approaches have very much in common. For example, both approaches are driven by a D-RTO layer, in which the objective function to be maximized is the plantwide profitability Φ. The main difference between the two frameworks is that in the one-layer approach, the input variables applied to the real plant are given by the optimization layer ($u_{Pw}=u_{opt}$), whereas for the two-layer, the inputs applied to the real plant are calculated by a control layer ($u_{Pw}=u_{mpc}$) that uses as set points, the optimal values of the states given by the optimization layer (x_{opt}). In both cases, the decision variables of the optimization problem are the plantwide manipulated variables u_{Pw}. In the two-layer case however, a second layer (NMPC controller) is used, in which an optimization problem is also solved for minimizing a performance-type objective function Γ, which can be composed of three terms: a penalization of the deviation of the main state variables from their set points (x_{opt}), a term that prevents large changes in the manipulated variables from one sample time to the next, and a term that constraints the manipulated variables to a small envelope around the reference trajectories u_{opt}, given by the optimization layer.

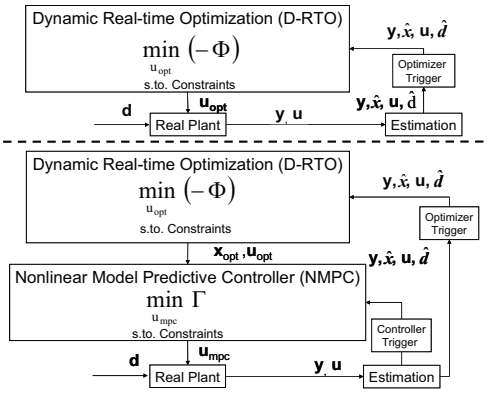


Fig. 4. Optimization-Based Control Strategies: One-layer (top) and Two-layer (bottom)

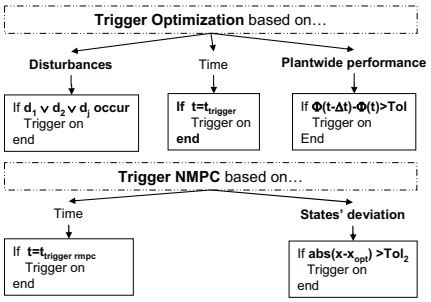


Fig. 5. Trigger for re-calling the D-RTO (top) and Controller layers (bottom)

Finally, trigger blocks in Fig. 4 deserve special mention due to their importance in the implementation of optimization-based control strategies. These trigger blocks act like switches for re-calling the optimization and control layers. An optimization-trigger for recalling the D-RTO layer can work based on a time criterion (e.g. the optimization is called periodically at a predetermined frequency), based on the disturbances dynamics (occurrence of a disturbance) or based on the performance of the plantwide profitability objective function (when Φ decreases below a certain tolerance). On the other hand, the controller-trigger can be based on a time criterion or on the state variables deviations from their optimal set points. Fig. 5 shows schematically the different criteria for activating the optimization and controller triggers.

Stage 6: Dynamic Real Time Optimization (D-RTO)

Because a nonlinear dynamic large-scale optimization problem arises in the last stage of the PWOC, an efficient feasible optimization method should be used in order to solve the problem in real time. For this purpose, different types of optimization algorithms can be used. However, the use of stochastic or evolutionary algorithms is considered here because of their reduced computational load (they do not need information about derivatives as required by gradient-based methods) and their relatively simple implementation. In this work, a stochastic method (i.e. localized random search) is used for solving the optimization problem in the PWOC. Independently of the optimization algorithm used, the method will search for the optimal solution in the space of the decision variables, which is a region bounded by the lower and upper limits of each manipulated variable (which are the decision variables of the optimization problem). This search region may be too large, resulting in long calculation times for finding an optimal solution, making difficult the solution of the PWOC problem in real time. In order to improve the efficiency of the optimization method for solving the large scale D-RTO problem, in the following section, a new stochastic-based approach for shrinking the search region of the optimization problem is introduced.

4. STOCHASTIC-BASED SHRINKING OF THE SEARCH REGION OF THE D-RTO PROBLEM

The main idea of the stochastic shrinking approach, is that for a sample time Δt (during which a disturbance took place in the process or the profitability function decreased), the changes on each plantwide manipulated variable (Δu_{PWi}) required for rejecting a disturbance, should be calculated as a function of the changes in the disturbances (Δd_j) and in the profitability objective function ($\Delta \Phi$). Mathematically, this can be written as shown in (1).

$$\Delta u_{PWi} = u_{PWi,t+\Delta t} - u_{PWi,t} = f_i(\Delta d_1, \Delta d_2, ...\Delta d_j, \Delta\Phi) \quad (1)$$

Where i is the number of plantwide manipulated variables and j is the number of disturbances that can be present in the process. f_i is a function that represents how much the manipulated variable i should change to reject disturbances. Specifically in this work, the use of a Gaussian distribution for describing function f_i is proposed. In this way, the changes on the manipulated variables are given in (2).

$$\Delta u_{PWi} = \xi_i\left(0\ ,\ \sigma_{u_i}\right) \quad (2)$$

Where $\xi_i(0,\sigma_{ui})$ represents a random number obtained from a Gaussian distribution with zero mean and standard deviation σ_{ui}. This standard deviation can be calculated as the maximum between different contribution terms, which represent the capability of the manipulated variable i for rejecting the different known disturbances of the process at time t, and for rejecting a decrease in Φ (that can be caused by both known and unknown disturbances), as shown in (3),

$$\sigma_{u_i} = \max(w_{i1}\Delta d_1, w_{i2}\Delta d_2, ...w_{ij}\Delta d_j, w_{i\Phi} z_\Phi \Delta\Phi) \quad (3)$$

where w_{ij} are gain factors that express how much a change in the manipulated variable u_{Pwi} can reject (or counteract) the occurrence of disturbance d_j, $w_{i\Phi}$ is the gain factor for the manipulated variable i rejecting the decrease in the profitability objective function Φ, and z_Φ is a dummy variable that is only activated when the objective function Φ decreases below a given tolerance *Tol*, that is:

$$z_\Phi = \begin{cases} 0, & \Phi(t-\Delta t) - \Phi(t) \le \text{Tol} \\ 1, & \Phi(t-\Delta t) - \Phi(t) > \text{Tol} \end{cases} \quad (4)$$

A final mention should be done, regarding the calculation of the gain factors used for obtaining the standard deviation σ_{ui}. It would be desirable to calculate these gains mathematically, from the nonlinear model of the process as expressed in (5)

$$w_{ij} = \sum_{k=1}^{n} \frac{\partial u_i}{\partial x_k} \times \frac{\partial x_k}{\partial d_j} \quad (5)$$

where $\partial u_i/\partial x_k$ represents the inverse of the open loop gain between state variable x_k and input u_i; and $\partial x_k/\partial d_i$ represents the open loop disturbance gain between x_k and disturbance d_j. As the complexity of the process model increases, the complexity for calculating the w_{ij} factors analytically also increases. For this reason, these gain factors are proposed to be calculated by using Digraphs. Information regarding digraph models is found elsewhere (Maurya et al., 2003).

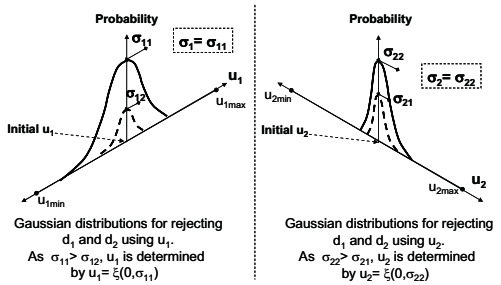


Fig. 6. Shrinking approach: Gaussian distributions for a system with two manipulated variables and two disturbances.

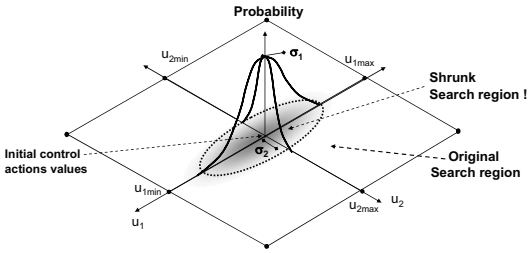


Fig. 7. Shrinking approach: Final Shrunk Search Region vs. Original Search Region.

For concluding this section, a graphical representation of the shrinking approach for a system with two manipulated variables and two disturbances that occur at the same time, is presented in Fig. 6 and 7. Fig. 6 shows the Gaussian distributions with standard deviation σ_{11} and σ_{22} for describing the manipulated variables u_1 and u_2 respectively, when disturbances occur in the process. Fig. 7 shows the Shrunk Search Region for the optimization problem, formed by the Gaussian distributions for u_1 and u_2. It is important to notice that despite the maximum standard deviation has been

selected for each case, a reduction of the search space for the optimization algorithm is achieved because the original search region of the optimization problem was only bounded by the upper and lower bounds of u_1 and u_2 (see Fig. 7). The stochastic-based shrinking approach is used in section 5 for reducing the search region of the optimization problem that arises when the PWOC concept is applied to the ethanol case study. As it will be shown through this example, the PWOC problem has been solved more efficiently by applying the shrinking approach than without shrinking.

5. PWOC FOR THE ETHANOL PROCESS: RESULTS AND COMPARISON

The main purpose of this section is to show the application of PWOC to the bio-ethanol process described in Section 2 and to compare the obtained results to a typical decentralized SISO loops scheme. The decentralized architecture implemented for comparison uses seven PID control loops, in addition to the Local Control strategy introduced in Section 2. The paired PID loops (controlled-manipulated variable) are the following: E_4-G_2-F_0, G_4-F_{13}, x_{DE1}-R_1, x_{BE1}-VB_1, x_{DE2}-R_2, x_{BE2}-VB_2, where E_4, G_2, G_4, are the ethanol concentration in the fermentor and the glucose concentration in the saccharificator and in the fermentor, respectively. x_{DE1}, x_{BE1}, x_{DE2}, x_{BE2}, corresponds to the mol fractions of ethanol in the top and bottoms of the distillation and rectification columns, respectively. A special mention should be done regarding the control loop E_4-G_2-F_0, which is a cascade proposed due to the fact that the ethanol to be produced depends strongly on the glucose concentration (G_2) that comes from the saccharificator. Finally, it should be noticed that following recommendations given by Araujo (2007), and in order to do a fair comparison to the PWOC results, the controlled variables for the distillation and rectification columns in the decentralized loops are concentrations and not temperatures (or temperature differences), which are usually the real controlled variables at an industry level. On the other hand, the main objective of PWOC is to control the profitability at its maximum value, and therefore the pairing controlled-manipulated variable is avoided. In the following, the PWOC stages are applied in detail to the bio-ethanol process.

Stages 1-3: Identification and design of necessary control loops.

For the bio-ethanol process, the following control loops has been identified as *necessary local loops*: level control in all tanks, pH and temperature control in saccharificator and fermentor, and pressure control in flash, distillation and rectification. Additionally to these loops, as explained in section 2, a biomass control strategy is used. With exception of the loops involved in the biomass strategy, all local loops are SISO (e.g. PI or PID). After implementing the local loops, the process still has 7 available manipulated variables that are used as plantwide manipulated for maximizing the profitability of the process. These plantwide manipulated variables are: F_0, F_1, F_{13},VB_1, R_1,VB_2, R_2, corresponding to starch and enzymes input flow, recycle flow from the flash to the fermentor and vapour and reflux rates for each column.

Stage 4: Statement of Plantwide Profitability Function (Φ).

The following profitability objective function is proposed to be maximized for the ethanol process addressed in this work:

$$\begin{aligned}\Phi = {} & w_1 \int_{t_0}^{t_0+\Delta t_{opt}} x_{ED_2} D_2 \, dt - w_2 \int_{t_0}^{t_0+\Delta t_{opt}} F_0 S_0 dt + w_3 \int_{t_0}^{t_0+\Delta t_{opt}} x_{ED_2} dt \\ & - w_4 \int_{t_0}^{t_0+\Delta t_{opt}} F_4 dt - w_5 \int_{t_0}^{t_0+\Delta t_{opt}} VB_1 dt - w_6 \int_{t_0}^{t_0+\Delta t_{opt}} VB_2 dt \\ & - w_7 \int_{t_0}^{t_0+\Delta t_{opt}} x_{WD_2} D_2 dt - w_8 \int_{t_0}^{t_0+\Delta t_{opt}} x_{EB_1} B_1 dt - w_9 \int_{t_0}^{t_0+\Delta t_{opt}} x_{EB_2} B_2 dt\end{aligned} \quad (6)$$

where w_i are weighting factors. The first term in (6) is related to the productivity of the process (expressed as the product between the ethanol concentration and the distillate flow rate in the top of the rectification column); the second term penalizes raw material consumption; the third term is a quality soft constraint; the following three terms in the second line of the equation (accompanied by w_4, w_5 and w_6) are used for penalizing the energy consumption in the process (pumping power and steam consumption). Last part of the equation contains a term that penalizes the presence of water at the top of the rectification column (related to post-processing costs in the dehydration unit) and two terms associated to economic losses due to the presence of ethanol in the bottom of the columns. t_0 is the initial time for the optimization routine and Δt_{opt} is the prediction horizon over which the objective function and constraints are evaluated. Δt_{opt}=15 hours has been selected taking into account the slow dynamic response of the process to changes in its inputs.

Stage 5: Design of the Optimization-Based Control Strategy

In order to compare the one- and two-layer approaches, PWOC for the ethanol case study is addressed using these two approaches shown in Fig. 8. It must be noticed that the biomass control is run in cascade with the D-RTO layer (in both frameworks), from which it receives the optimal set point value that should be locally tracked. In both cases, the objective function to be maximized in the D-RTO layer is given by (6). The complete formulation of the optimization problem addressed in the D-RTO layers is given in (7). As can be seen, the decision variables of the optimization problem are the values for the u_{Pw}. The last inequality constraint is used inside the optimization loop for assuring that the solution of the optimization problem will guarantee a long-term ethanol concentration at the top of the rectification column (x_{ED2}) equal or higher than the concentration obtained if the plantwide manipulated variables were kept constant at u_{Pw}^* (values of the manipulated variables at the time t_0). The performance-type objective function Γ in the NMPC layer of the two-layer approach (bottom of Fig. 8) penalizes deviations of the ethanol concentration in the fermentor (E_4) and in the top of the rectification column (x_{ED2}), respectively, from their optimal set points values given by the D-RTO layer during a prediction horizon Δt_{mpc}=2 hours, as stated in (8). The terms Q and R are weighting matrices, which can be seen as tuning parameters for the NMPC. Schwartz et al. (2006) present a method for determining MPC tuning parameters that lead to optimal results from either an

operational or financial standpoint. Finally, the trigger conditions used in the simulation study for addressing the PWOC problem of the ethanol process are shown in Fig. 9.

$$\min_{u_{PW}} (-\Phi(\dot{x}, u, t_0, \Delta t))$$
$$\text{s.t. } f(\dot{x}, x, u, d, t) = 0 \quad (7)$$
$$x_i(t_0) = x_{0i}\,,\; u_{min} \le u \le u_{max}$$
$$x_{ED_2}\left(t_0 + \Delta t_{opt}, u_{PW}\right) \ge x_{ED_2}\left(t_0 + \Delta t_{opt}, u_{PW}^{*}\right)$$

$$\Gamma = \int_{t_{0,mpc}}^{t_{0,mpc}+\Delta t_{mpc}} Q(x_{ED2} - x_{ED2,opt})^2 dt + \int_{t_{0,mpc}}^{t_{0,mpc}+\Delta t_{mpc}} R(E_4 - E_{4,opt})^2 dt \quad (8)$$

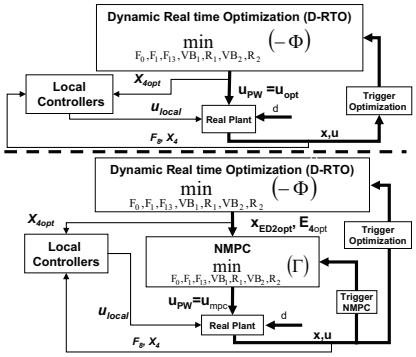


Fig. 8. Optimization-Based Control Strategies for the ethanol process: One-layer (top) and Two-layer (bottom)

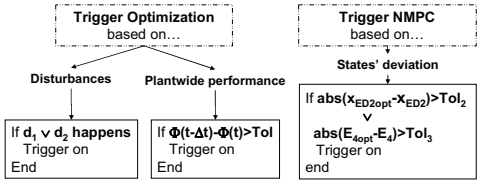


Fig 9. Trigger conditions for the ethanol PWOC: Optimization (left) and NMPC (right) layers.

The D-RTO problem was solved by the direct Sequential approach using a Monte Carlo localized random search optimization method, which is simple to implement, have broad applicability and do not require the computation of gradients (Spall, 2003). Basically, the algorithm consists of three main steps. First, an initial guess θ_0 of the optimal point is randomly picked and the number of iterations k, is set to zero. Second, an independent random vector d_k is generated, and added to the current optimal value θ_k. Third, it is checked if $-\Phi(\theta_k+d_k) < -\Phi(\theta_k)$; if this condition is satisfied, the new optimal value is set as $\theta_{k+1}=\theta_k+d_k$, otherwise, the second step is repeated (random generation of d_k). The algorithm stops when either, the maximum number of iterations has been reached or a convergence criterion has been fulfilled. For testing the PWOC approach, simulation studies were carried out using the nonlinear model of the process as the real plant. Results presented in this Section correspond to the simulation of the system starting at an optimal steady state. After 6 hours of operation at this steady state, a disturbance on the starch feed concentration enters the process (20% reduction of the starch concentration). At this moment, the optimization trigger is switched on and the D-RTO layer is called in order to calculate the new values for the plantwide manipulated variables that drive the process to optimal operation (maximal profitability). The localized random search method was used as previously explained for maximizing the profitability objective function, subject to the constraints given in (7). The optimization algorithm was selected to be run each time during 50 iterations after making a balance between performance and computational time for real-time implementation. The shrinking approach described in Section 4 was used for reducing the search space of the optimization problem. Specifically, the gain factors w_{ij} and $w_{i\Phi}$ in (3) were calculated using Digraphs. After calculating the gain factors, the standard deviation σ_{ui} of the Gaussian distribution that describes the probability of change of each manipulated variable for rejecting the disturbances was calculated as the maximum between different contribution terms. Then, precisely this Gaussian distribution for each manipulated variable was used for generating the vector d_k, in order to allow the optimization algorithm to make moves only in the region described by these distributions. Fig. 10 and 11 show the simulation results obtained of applying the PWOC to the ethanol case study, in presence of a disturbance on the feed concentration. PWOC was run using the two optimization-based control frameworks shown in Fig. 8. The first of these frameworks is the PWOC-one-layer (solid line) and the second is the PWOC-two-layer (dashed line). These two approaches are compared to the behaviour of the process when two different decentralized PID schemes are used, which in the following are denoted as: Decentralized 1 (described at the beginning of this section) and Decentralized 2. The only difference between both decentralized schemes is that in Decentralized 2, the E_4-G_2-F_0 loop is replaced by a D_2-F_0 loop, in order to keep constant the flow of product that goes to the dehydration unit. At this point it is important to remark that all the control approaches compared in this section use the Local Control Strategy mentioned in the Stages 1-3 of this Section, as part of the local control loops in the regulatory level. In the following Figures, the term *Two-layer SP* is used for denoting the set point values of the state variables E_4, x_{ED2} in the NMPC layer and for X_4 in the local control loop. These set point values are given by the D-RTO layer: $E_{4sp}=E_{4,opt}$, $x_{ED2,sp}= x_{ED2,opt}$ and $X_{4sp}=X_{4,opt}$. Also, the set point values for the state variables controlled in the decentralized schemes (including biomass concentration in the fermentor) correspond to the starting steady state values. Results shown in Fig. 10 are related to the fermentation section. It can be seen that using PWOC (both the one- and two-layer) results in a lower ethanol concentration in the fermentor than when using Decentralized schemes. Decentralized 1 achieves the highest ethanol concentration (E_4), which is not surprising because one of its control objectives is precisely to keep E_4 at its set point (original steady state value); and it is doing so by feeding a lower substrate flow rate (F_0), as shown in Fig 10-top-right. Fig 10-bottom-left shows the dynamic behaviour of the biomass concentration. It can be seen that both PWOC approaches keep a lower biomass concentration, due to the fact that they are actually tracking the optimal set point value, given by the D-RTO layer, and not just maintaining a fixed set point value (as done in the Decentralized schemes). At this point, analyzing the process as conventionally done as if it were

conformed just of a fermentation unit, wrong conclusions could arise, in which decentralized strategies will be claimed suitable enough or even, much better than the optimization-based approaches. However, it must be noticed that independently of the control scheme, the product of a bio-ethanol plant (after dehydration) is ethanol at purity higher or equal than 99.6% wt. Since the profitability of the plant is closely related to the net flow of this final product, and the latter is proportional to the net flow of ethanol in the fermentor output, which is given by $E_4 \times F_4$ (ethanol concentration × total output flow in the fermentor), then it can be concluded that the one- and two-layer approaches would lead the process to higher cumulative profitability values than the decentralized, because their total net flow of ethanol at the fermentor output is higher (See Fig. 10-bottom-right) in spite of their lower ethanol concentration. These results are confirmed in Fig 11, where the benefits of the PWOC approaches and the drawbacks of the decentralized schemes become evident when the profitability objective function (Fig 11-top) is compared for the different control approaches. Analyzing Fig 11, it is possible to conclude that PWOC results in a much more effective response to the disturbance than the Decentralized schemes, from an economical point of view. Specifically, the PWOC-one-layer reaches the highest cumulative profitability for the process, having at the same time the highest cumulative production rate and a higher ethanol concentration than the decentralized architectures. The PWOC-two-layer allows at the beginning a small decrease of its objective function (when compared to the initial value), despite the fact that this approach drives the process towards the highest product concentration (Fig 11-bottom-left). However, for the first 20 hours after the disturbance appearance, the two-layer approach kept the profitability on average at the same value than the starting steady state, and after this, it was able to improve its objective function value, reaching even at the end the same optimal value than the one-layer. Furthermore, it can be seen that the Decentralized schemes result not only in lower profitability, but also in lower product concentration and in the case of Decentralized 1, in the lowest cumulative flow of product. Of course, it can be argued that the product concentration resulting in the decentralized schemes is lower because precisely these controllers are doing their jobs regulating the controlled variables at their set point values. However, it must be noticed that regulation of the controlled variables at fixed set points might deteriorate the profitability of the process, because when a disturbance enters the process, the optimal operating point may also move. How much this point moves can not be generalized because it depends on the process and the nature of the disturbances, and in many cases disturbances can not be predicted. If well it is completely true that over the years process industry has been operating under fixed set point policies relying on PID SISO loops, without reporting enormous economical losses, it is also true as stated by Prett and García (1988), that the apparent savings in doing so (i.e. minimization of both design effort and maintenance) are in majority of cases nonexistent and in the long run result in more costs than the use of multivariate techniques. Following the analysis in Fig 11, comparing only the decentralized approaches, it is clear that Decentralized 2 is more convenient from an economic view, because at the end of the period, at least it reaches a profitability value close to the starting point, despite of resulting in less ethanol concentration in the fermentor (Fig 10-top-left).

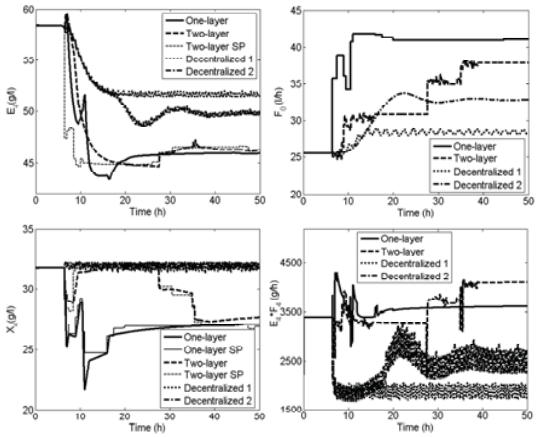


Fig. 10. PWOC results vs. decentralized control for the fermentation section: Ethanol Concentration (top-left), Starch Input Flow (top-right), Biomass Concentration (bottom-left) and Ethanol Net Flow (bottom-right).

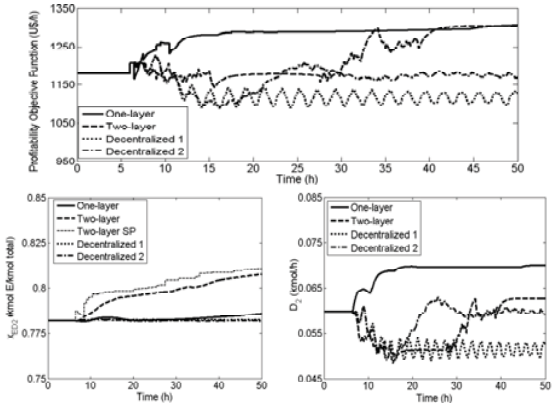


Fig. 11. PWOC results vs. Decentralized Control: Plantwide Profitability (top), Ethanol concentration in the distillate (left) and Distillate flow rate (right) at rectification section.

A final remark about the PWOC schemes should be done. If well both approaches reach the same profitability value at the end of the period (50 hours), and the two-layer has an optimal behaviour from a performance point of view (which is its "final" objective function), it might not be satisfactory at all from an economic view. Finally, it is important to highlight that independently of the optimization-based control framework selected (one or two-layer), the PWOC improves the profitability of the process when compared to the decentralized control strategies, and thus, it is a promising alternative for addressing the plantwide control problem of chemical or biochemical processes in which the profitability of the process is at risk when disturbances appear. On the other hand, in order to evaluate the performance of the shrinking approach, several simulation studies applying the PWOC with shrinking and without shrinking the search

region of the optimization problem were carried out. Fig. 12 shows the advantages of the shrinking approach. By running the optimization algorithm using the same number of iterations in the two cases, the shrinking approach has not only achieved a higher profitability, but also applied smoother control actions, which is an important fact for the stability of the process and is in general desirable.

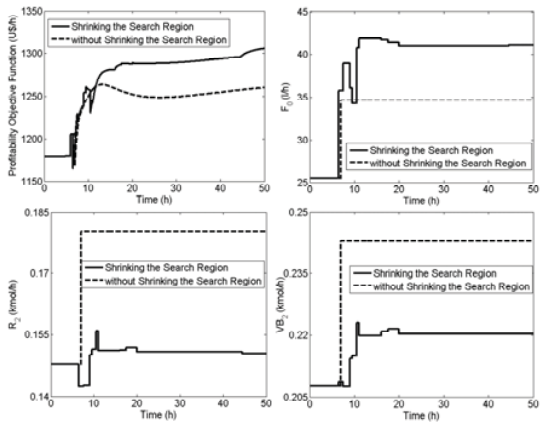


Fig. 12. One- layer PWOC: Shrinking vs. without Shrinking the Search Region. Profitability (top-left), Starch Input Flow (top-right), Reflux rate in the rectification (bottom-left) and Vapour flow rate to the rectification (bottom-right).

Analyzing the profiles for F_0, R_2 and VB_2 (Fig 12), it can be seen that when no shrinking is used (dashed lines) each manipulated variable change in a step-type policy with higher amplitude and longer period than when using shrinking (solid lines). Finally, it must be noticed that the main advantage of using the shrinking approach is that the probability of change for each manipulated variable is a function of the capability that each of them has for rejecting each disturbance (or the number of disturbances that occur at the same time, including those unknown), that means that the optimization algorithm does not waste time testing large changes in the manipulated variables that just reject in a weak way (or are not able to reject) the disturbances. When no shrinking is used, each manipulated variable is allowed to change from its lower to its upper bound, without any restriction, whereas the shrinking approach bounds the search region according to the standard deviation calculated for each manipulated variable, and it is precisely this standard deviation that contains the information about the cause-effect relationship between each manipulated variable and each disturbance.

6. CONCLUSIONS

A Plantwide Optimizing Control (PWOC) approach for bioprocesses has been presented based on the Optimizing Control concept. The main stages for PWOC and a stochastic-based shrinking approach for reducing the search space of the optimization problem have been introduced. PWOC has been applied to the bio-ethanol process, showing much better results from an economical point of view than when the process is only controlled by conventional control loops. It has been shown that PWOC is a very promising alternative for controlling chemical or biochemical processes in which the economical feasibility is at risk when disturbances appear. Finally, the shrinking approach was successfully tested resulting in an improvement of the optimization routine for real-time applications (i.e. higher productivities were obtained for the same number of iterations during optimization). Future work will be directed towards extending the shrinking approach for being applied with deterministic optimization methods.

ACKNOWLEDGMENTS

The authors acknowledge support from the Cluster of excellence “Unifying Concepts in Catalysis” coordinated by the Berlin Institute of Technology and funded by the German Research Foundation–Deutsche Forschungsgemeinschaft. Silvia Ochoa gratefully acknowledges financial support from DAAD and the Antioquia University of Colombia.

REFERENCES

Araujo, A.C.B., Hori, E.S. and Skogestad, S. (2007). Application of plantwide control to the HDA process. II Regulatory control. *Ind. Eng. Chem. Res.*,46, 5159-5174.

Araujo, A.C.B. (2007). *Studies on Plantwide Control*. Ph.D. Thesis. Norwegian University of Science & Technology.

Arifeen, N., Wang, R., Kookos, I. K, Webb, C., and Koutinas, A.A. (2007). Process Design and Optimization of Novel Wheat-Based Continuous Bioethanol Production System. *Biotechnol. Prog.* 23, 1394. 1403

Bartee, J.F., Macharia, M.A., Noll, P.D. and Tay, M.E. (2008). Integrated model predictive control of batch and continuous processes in a biofuel production process. *US Patent Application US2008/0109200 A1*.

Bartusiak, R.D. (2007). NLMPC: A platform for optimal control of feed- or product-flexible manufacturing. In: Findeisen, R. et al. (eds.). *Assessment and future directions of NMPC*. Springer-Verlag, Berlin, 367-381.

Biegler, L.T. and Zavala, V.M. (2009). Large-scale nonlinear programming using IPOPT. *Comp. Chem. Eng.*, 33, 575.

Buckley, P.S. (1964). *Techniques of Process Control*. John Wiley & Sons, Inc. New York.

Cheng, R., Forbes, J.F. and Yip, W.S. (2007). Price-driven coordination method for solving plant-wide MPC problems. *Journal of Process Control*, 17, 429-438.

Costa, A.C., Meleiro, L.A.C. and Filho, R.M. (2002). Non-linear predictive control of an extractive alcoholic fermentation process. *Process Biochemistry*, 38, 743-750

Costa, A.C., Atala, D.I.P., Filho, R.M. and Maugeri, F. (2001). Factorial design and simulation for optimization and determination of control structures for extractive alcoholic fermentation. *Proc. Biochem.,* 37, 125-137.

Duvall, P.M. and Riggs, J.B. (2000). On-line optimization of the Tennessee Eastman challenge problem. *Journal of Process Control*, 10, 19-33.

Engell, S. (2007). Feedback control for optimal process operation. *Journal of Process Control,* 17, 203-219.

Findeisen, W., Bailey, F.N., Brdys, M., Malinowski, K., Tatjewski, P. and Wozniak, A. (1980). *Control and coordination in hierarchical systems*. Wiley.

Franke, R. and Doppelhamer, J. (2007). Integration of advanced model based control with industrial IT. In

Findeisen, R. et al. (eds.). *Assessment and future directions of NMPC*. Springer-Verlag, Berlin, 399-406.

Franke, R. and Vogelbacher, L. (2006). Nonlinear model predictive control for cost optimal startup of steam power plants. *Automatisierungstechnik*, 54, 630-637.

Garcia, C.E. and Morari, M. (1984). Optimal operation of integrated processing systems. Part II: Closed-loop on-line optimizing control. *AIChE Journal*, 30, 226-234.

Jia, D. and Krogh, B.H. (2001). Distributed model predictive control. *IEEE Proc. American Control Conference*.

Jockenhövel, T., Biegler, L.T. and Wächter, A. (2003). Dynamic optimization of Tennessee Eastman process using OptControlCentre. *Comp. Chem. Eng.*, 27, 1513.

Kadam, J.V. and Marquardt, W. (2007). Integration of economical optimization and control for transient process operation. In: Findeisen, R. et al. *Assessment and future directions of NMPC*. Springer, Berlin, 419-434.

Kadam, J.V. and Marquardt, W. (2004). Sensitivity-based solution updates in closed-loop dynamic optimization. *DYCOPS 7 Proceedings*.

Kadam, J.V., Marquardt, W., Schelegel, M., Backx, T., Bosgra, O.H., Brouwer, P.J., Dünnebier, G., van Hessem, D., Tiagounov, A. and de Wolf, S. (2003). Towards integrated dynamic real-time optimization and control. *FOCAPO Conference Proceedings,* 593-596.

Kadam, J.V., Schlegel, M., Marquardt, W., Tousain, R.L., v. Hessem, D.H., Berg, J.v.d. and Bosgra, O.H. (2002). A two-level strategy of integrated dynamic optimization and control of industrial processes. *ESCAPE 12*, 511-516

Küpper, A. and Engell, S. (2008). Engineering of Online Optimizing Control - A Case Study: Reactive SMB Chromatography. *Proc. 17th IFAC World Congress.*

Larsson, T., Govatsmark, M.S., Skogestad, S. and Yu, C.C. (2003). Control structure selection for reactor, separator, and recycle processes. *Ind. Eng. Chem. Res.*, 42, 1225.

Lausch, H.R., Wozny, G., Wutkewicz, M. and Wendeler, H. (1998). Plant-wide control of an industrial process. *Trans. IChemE Part A*, 76, 185-192.

Lu, J.Z. (2003). Challenging control problems and emerging technologies in enterprise optimization. *Control Engineering Practice*, 11, 847-858.

Luyben, M.L., Tyreus, B.D. and Luyben, W.L. (1997). Plantwide control design procedure. *AIChE Journal*, 43, 3161-3173.

Manenti, F. and Rovaglio, M. (2007). Integrated multilevel optimization in large-scale poly(ethylene terephthalate) plants. *Ind. Eng. Chem. Res.*, 47, 92-104.

Maurya, M.R., Rengaswamy, R. and Venkatasubramanian, V. (2003). A systematic framework for the development and analysis of signed digraphs for chemical processes. *Ind. Eng. Chem. Res.,* 42, 4811-4827.

McAvoy, T.J. and Ye, N. (1993). Base control for the Tennessee Eastman process. *Comp. Chem. Eng.*, 18, 383.

Meleiro, L.A.C., Von Zuben, F.J. and Filho, R.M. (2008). Constructive learning neural network applied to identification and control of a fuel-ethanol fermentation process. *Eng. Applic. Artificial Intelligence,* 22, 201-215.

Mercangöz, M. and Doyle, F.J. (2007). Distributed model predictive control of an experimental four-tank system. *Journal of Process Control*, 17, 297-308.

Ochoa, S., Repke, J.U. and Wozny, G. (2009). Integrating Real-Time Optimization and Control for Optimal Operation: Application to the Bio-ethanol Process. *Biochem. Eng. J.* In press. doi:10.1016/j.bej.2009.01.005

Olofsson, K., Rudolf, A. and Liden, G. (2008). Designing SSF for improved xylose conversion by a recombinant strain of *S. cerevisiae*. *J. Biotechnol.* 134, 112–120

Prett, D.M. and Garcia, C.E. (1988). *Fundamental Process Control*. Butterworth. Stoneham.

Rawlings, J.R. and Stewart, B.T. (2008). Coordinating multiple optimization-based controllers: New opportunities and challenges. *J. Proc. Control*, 18, 839.

Robinson, D., Chen, R., McAvoy, T. and Schnelle, P.D. (2001). An optimal control based approach to designing plantwide control system architectures. *Journal of Process Control*, 11, 223-236.

Roman, R., Nagy, Z.K., Cristea, M.V. and Agachi, S.P. Dynamic modelling and nonlinear model predictive control of a fluid catalytic cracking unit. *Computers and Chemical Engineering*, 33, 605-617.

Scattolini, R. (2009). Architectures for distributed and hierarchical model predictive control – a review. *J. Proc. Control*, In press. doi:10.1016/j.jprocont.2009.02.003

Schwartz, J.D. Wang, W. and Rivera, D.E. (2006). Simulation-based optimization of process control policies for inventory management in supply chains. *Automatica*, 42, 1311-1320.

Spall, J.C. (2003). *Introduction to Stochastic Search and Optimization*. John Wiley & Sons, Inc., Hoboken, NJ.

Stephanopulous, G., and Ng, C. (2000). Perspectives on the synthesis of plant-wide control structures. *Journal of Process Control*, 10, 97-111.

Sun,Y. and El-Farra, N.H. (2008). Quasi-decentralized model-based networked control of process systems. *Computers and Chemical Engineering*, 32, 2016-2029.

Toumi, A. and Engell, S. (2004). Optimization-based control of a reactive simulated moving bed process for glucose isomerization. *Chem. Eng. Sci.*, 59, 3777-3792.

Tosukhowong, T., Lee, J.M., Lee, J.H. and Lu, J. (2004). An introduction to dynamic plant-wide optimization strategy for an integrated plant. *Comp. Chem. Eng.*, 29, 199-208.

Venkat, A.N., Rawlings, J.B. and Wright, S.J. (2007). Distributed model predictive control of large-scale systems. In: Findeisen, R. et al. (eds.). *Assessment and future directions of NMPC*. Springer, Berlin, 591-605.

Venkat, A.N. (2006). *Distributed model predictive control: Theory and applications*. Ph.D. Dissertation, University of Wisconsin-Madison.

Ying, C.M. and Joseph, B. (1999). Performance and stability analysis of LP-MPC and QP-MPC cascade control systems. *AIChE Journal*, 45, 1521-1534.

Zavala, V.M., Laird, C.D. and Biegler, L.T. (2007). Fast implementations and rigorous models: Can both be accommodated in NMPC?. *International Journal of Robust and Nonlinear Control*, 18, 800-815.

Zhen, A. Mahajanam, R.V. and Douglas, J.M. (1999). Hierarchical procedure for plantwide control systems synthesis. *AIChE Journal*, 45, 1255-1265.

A new approach for the modelling of crystallization processes in impure media using Population Balance Equations (PBE)

François Févotte[*] and Gilles Févotte[**]

** 99, Avenue de Verdun. 92130 Issy les Moulineaux (France)*
*** Ecole des Mines de Saint Etienne, Centre SPIN. 158, Cours Fauriel. 42000 Saint Etienne & Université Lyon 1, 43, Avenue A. Einstein, 69622 Villeurbanne, Cedex (France). fevotte@emse.fr*

Abstract: For obvious industrial and theoretical reasons the problem of accounting for the effect of impurities in the population balance modelling of solution crystallization processes is a very important issue, and yet it has never been reported until today. Meanwhile, several kinetic models are proposed in the literature that relate the effect of impurities on the crystal growth and could be used for PBE modelling. The goal of the present paper is to address this issue and to present a new method, based on characteristics, which is shown to efficiently solve the difficulties raised by the specificity of the mathematical formulation of the Population Balance Equation (PBE) in the presence of impurities. Indeed, as far as hindering effects of the impurities on the crystal growth are concerned, it turns out that the "age" of the particles (i.e. the time they spent in the presence of impurities) might plays a key-role in the overall dynamic crystallization process. Accounting for such a new internal variable required a specific PBE resolution algorithm to be developed and evaluated.

Keywords: Chemical industry, Crystallization, Characteristic curves, Nucleation, Modeling, Population balance equations, Numerical simulation, Batch processes.

1. INTRODUCTION

1.1 Population Balance Equations (PBE) and crystallization

The formalism of Population Balances Equations (PBEs) is a widely used modelling tool in engineering, with applications including crystallization, powder technologies, polymerization processes, biotechnologies, etc (Ramkrhisna and Mahoney, 2002). PBEs allow describing the time variations of properties of a large number of separate entities, such as particles, bubbles or droplets, interacting with each other and/or with their environment which usually consists of a continuous phase. The dynamics of complex distributed particulate systems is related through the evolution of appropriate distribution functions evolving in a p-dimensional space where p represents the number of internal coordinates required to characterize the particles. Internal coordinates refer to continuous properties of the individual dispersed entities such as size, composition, cristallinity, etc, or to discrete features such as the number of primary crystals in agglomerates or the number of free radicals in a polymerizing particle during emulsion polymerization reactions. In addition to internal coordinates, external coordinates are necessary to describe the physical location of the distributed entities.

As far as dispersed phases are concerned (i.e. separate entities in a continuous fluid phase), the governing equations involve the number density of particles, which is defined as follows:

$$E[n(\boldsymbol{x},\boldsymbol{r},t)] \equiv \psi(\boldsymbol{x},\boldsymbol{r},t), \quad \boldsymbol{x} \in \Omega_x \ , \qquad \boldsymbol{r} \in \Omega_x \tag{1}$$

As already mentioned, external and internal coordinates (i.e. $\boldsymbol{r}$ and $\boldsymbol{x}$, respectively) are necessary to characterize the "location" and the properties of the particles. Equation (1) actually means that the average number of particles in the particle state subspace $dV_x dV_r$ with coordinates $(\boldsymbol{x},\boldsymbol{r})$ is given by $\psi(\boldsymbol{x},\boldsymbol{r},t)\ dV_x dV_r$.

For the sake of simplicity $(\boldsymbol{x},\boldsymbol{r})$ is usually referred to as the particle state and, as outlined by Ramkrishna (2000), it is worth noting that the further definition of PBEs requires the average number density function $\psi(\boldsymbol{x},\boldsymbol{r},t)$to be sufficiently smooth for allowing differentiation with respect to the coordinates and time.

According to the previous definitions, the number $\mathcal{N}$ of particles belonging to a given subset $A_x \subset \Omega_x$ is given by:

$$\mathcal{N}(\boldsymbol{r},t) = \int_{A_x} \psi(\boldsymbol{x},\boldsymbol{r},t) dV_x \tag{2}$$

The previous mathematical formalism will now be applied to the time variations of crystals (i.e. solid particles generated during a crystallization process) characterized by some internal coordinates x. The rate of variation of $\boldsymbol{x}$ is referred to as $\dot{\boldsymbol{X}}$ in the following:

$$\frac{dx}{dt} = \dot{\boldsymbol{X}}(\boldsymbol{x},\boldsymbol{r},\boldsymbol{y},t) \tag{3}$$

where y represents any scalar variables required to quantify the possible interactions (e.g. through heat or mass transfer) between the particles and the continuous phase.

Considering that the following operating conditions are verified:

1. The solution crystallization process takes place in a well mixed batch reactor,

2. new crystals are generated through nucleation phenomena only (i.e. according to the "classical" nucleation theory, the size of new particles appearing in the dispersed phase is the critical size L^* which can be assumed negligible. Agglomeration and breakage of the particles are both neglected),

the population balance equation relating the time variations of the particle state is:

$$\frac{\partial}{\partial t}\psi(\boldsymbol{x},t) + \nabla_x \dot{\mathbf{X}}(\boldsymbol{x},y,t)\,\psi(\boldsymbol{x},t) = h(\boldsymbol{x},y,t) = 0 \qquad (4)$$

with the following boundary conditions:

$$\begin{aligned}\dot{\mathbf{X}}(0,\boldsymbol{x}',y,t)\,\psi(0,\boldsymbol{x}',t) &\cong \dot{\boldsymbol{X}}(L^*,\boldsymbol{x}',y,t)\,\psi(L^*,\boldsymbol{x}',t)\\ &= R_N(y,t)\end{aligned} \qquad (5)$$

Where the vector of internal coordinates is decomposed as:

$$\boldsymbol{x} = (L,\boldsymbol{x}')$$

The first assumption above implies that the number density function does not depend on space coordinates while assumption 2 means that, in order to express the source of new particles in the system, one has only to define boundary conditions to account for the expression of the rate(s) of nucleation of crystals (i.e. $h(\boldsymbol{x},y,t)$=0 in (4)).

In the mono-dimensional case where one characteristic size of the particle only is considered (e.g. the diameter L of a fictitious spherical particle exhibiting the same projected area than the crystal under consideration), (4) reduces to the well-known following partial differential equation allowing to compute the time variations of the Crystal Size Distribution (CSD). In the following, y is the supersaturation of the continuous liquid phase σ, defined by (9), which will now be omitted for the sake of simplicity:

$$\begin{cases} \dfrac{\partial\psi(L,t)}{\partial t} + G(t)\dfrac{\partial\psi(L,t)}{\partial L} = 0 & (6)\\ \psi(L,0) = 0 \quad or \quad \psi(L,0) = \psi_{seed} & (7)\\ \psi(0,t) \cong \psi(L^*,t) = \dfrac{R_N(t)}{G(t)} & (8)\end{cases}$$

The initial condition (7) accounts for the possibility of the crystallization to start through primary nucleation (i.e. no solid phase is initially present in the crystallizer) or through seeding. Seeding consists in the introduction of small amount of particles, usually sieved, in the supersaturated solution. The seed particles initiate the crystallization process and are characterized by their size distribution ψ_{seed}.
R_N is the rate of nucleation in #.s^{-1}.m^{-3} and G is the crystal growth rate in m/s.

1.2 Growth rates and impurities

In most published PBE modelling studies —according to McCabe's hypothesis— the growth rate *G(t)* is assumed not to depend on the particle size but essentially on the driving force of crystallization, the following absolute definition of the supersaturation σ is now defined as:

$$\sigma(t) = C(t) - C^* \qquad (9)$$

where C^* is the equilibrium concentration (i.e. the temperature-dependant solubility of the crystallizing compound) and *C(t)* is the solute concentration.

Several theoretical or phenomenological expressions can be found in the literature to express the supersaturation-dependency of the growth rate which, more or less, turns out to obey the following kinetic law:

$$G(t) = \frac{dL}{dt} = k_g(C(t) - C^*)^i = k_g\sigma(t)^i \qquad (10)$$

exponent i was shown to depend on the involved growth mechanism(s) which, in particular, depend(s) on the level of supersaturation (Mersmann A, 2002; Mullin J.W., 1993; Chernov, 2004). In practice, consistently with usual theoretical models, most published values of i are given between 1 and 2.

Actually, modeling and control papers published in the field of crystallization engineering deal essentially with pure solute/solvent systems. As far as one considers the context of industrial processes, this is obviously an unrealistic assumption. Indeed, it is worth noting that industrial processes cannot avoid undesirable impurities to be generated during the many chemical reactions preceding the crystallization steps. It is well-known that even minute concentrations of impurities present in the initial solution can affect the crystallization processes (Chernov, 2004; Sangwal, 1996; Wood, 2001) and induce significant reductions of the growth rate (Keshra & Sangwal, 1996; Kubota et al., 2000; Kubota, 2001).

It is also known that impurities can lead to supersaturation thresholds below which the development of crystallization is completely inhibited (see e.g. Sangwal, 2002). To the best of our knowledge, such key-features of "real" industrial crystallization processes (i.e. processes performed in the unavoidable presence of impurities) were investigated through the observation of single crystals, and never described using PBEs. Therefore it remains important to evaluate quantitatively the distribution and the time-variations of the detrimental effects of impurities during crystallization processes.

Now, if one considers the variety of the techniques which were proposed to solve the PBEs in the case of crystallization processes, it appears that few of these methods are based on

the method of characteristics (MOC). It is however known that MOCs avoid numerical diffusion errors and oscillatory solutions caused by the discretization of the involved growth term, especially when steep or discontinuous particulate phenomena take place in suspension (Kumar and Ramkrishna, 1997; Briesen, 2006).

Quamar and Warnecke (2007) have proposed a numerical method for solving PBEs involving nucleation, growth and aggregation processes. The scheme combines a method of characteristics for computing the growth term, with a finite volume technique for calculating aggregation terms. The method is compared to a finite volume scheme through the modelling of "academic" situations for which analytical solutions are available (i.e. combination of crystal growth with aggregation or nucleation). The authors show that the numerical scheme based on MOC is more efficient than pure finite volume schemes, and that it better tracks steep variations of number density functions. This interesting feature of MOC is attributed to the disappearance of the advection term $\partial G\psi/\partial L$ from the main PBE.

Sotowa et al. (2000) compared the numerical resolution of a simple crystallization PBE using a finite difference method and the method of characteristics. The study aimed at evaluating the impact of numerical dispersion on the design of feedback controllers. It was finally concluded that, as far as the simulation of control systems is concerned, the method of characteristics is recommended as a numerical technique for simulating crystallization processes.

More recently, in order to simulate the growth of anisotropic particles, Briesen (2006) proposed a reduced two-dimensionnal PBE model. Here, the MOC approach is used to validate the calculations. The application deals with the crystallization of potassium dihydrogen phosphate which is assumed to exhibit the shape of parallelepipeds terminated by two tetragonal pyramids. However, the simulation assumes initial seeding of the crystallization process only, i.e. no primary or secondary nucleation is taken into account, which is a rather questionable assumption. It should also be noticed that no specific information is reported about the MOC used.

It is the goal of the present paper to address the problem of accounting for the "birthdate" of crystals in the governing crystallization PBE, and to propose a new numerical scheme, based on MOC, for the resolution of the latter PBE. In fact, it is clear that the approach proposed by Kumar and Ramkrishna (1996a,b, 1997) in their series of three papers is much more "advanced" than the approach presented here, in terms of the accuracy of the used size integration technique and with respect to the ability of the method to describe agglomeration and breakage phenomena. Nevertheless, the present algorithm offers another way of considering nucleation phenomena and, through its great simplicity, could be valuable for applications where fast computation is required (i.e. for in-line feedback control applications for example).

2. MODELING THE CRYSTAL GROWTH RATE IN THE PRESENCE OF IMPURITIES.

2.1 The pinning mechanism

With respect to the growth of crystals in pure solvent, the time-averaged advancement velocity of a step in impure media appears to be hindered by the adsorption of impurity species on the growing crystal surface. Indeed, as Fig. 1 schematically shows, during the step advancement, kink sites can be blocked by foreign species that cannot easily be incorporated in the crystal lattice. To allow further crystal growth, the growth-step has to circumvent the pinned impurity, which obviously reduces the overall growth rate. Several models describing such a pinning mechanism were early described in the literature (see e.g. Cabrera & Vermilyea, 1958). Moreover, it is worth noting that many convincing observations of the pinning mechanism were reported, using e.g. advanced imaging techniques such as AFM (Atomic Force Microscopy, see e.g. Land et al., 1999, Thomas et al., 2004).

Kubota-Mullin's model (1995) was proposed to describe the pinning mechanism through Γ, the ratio between the step velocities in pure (u_0) and impure (u) media. Γ is given by the following expression:

$$\Gamma = \frac{u}{u_0} = 1 - \left[\frac{\gamma a}{kT\sigma d}\right]\theta = 1 - \alpha\theta \qquad (11)$$

where γ is the edge free energy, a is the size of the growth unit, T is the absolute temperature, k is the Boltzmann constant and θ is the fraction of coverage of active growing crystal surface by adsorbed impurities, d is the average distance between actives growth sites.

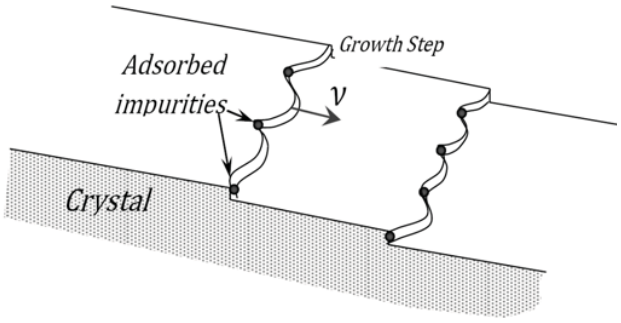


Fig. 1. Adorption of impurities at kink sites on the growing steps after Kubota (2001).

Parameter α is an effectiveness factor which quantifies the efficiency of the impurity specie in hindering the crystal growth. It is very important to notice that α does not only depend on properties of the involved solid, but also on supersaturation.

The coverage of the crystal surface by impurities is itself a stable dynamic process which therefore reaches a steady-state

θ^*. According to the hypotheses set to describe the adsorption process, various theoretical approaches can be used to compute the equilibrium coverage parameter θ^*.

In Kubota-Mullin's Model (Kubota & Mullin, 1995; Kubota et al., 1997), the equilibrium coverage of the growing surface is estimated thanks to Langmuir's adsorption theory:

$$\theta^* = KC_i/(1+KC_i) \tag{12}$$

where K is the Langmuir adsorption constant and C_i is the concentration of impurity.

Even though the adsorption process is often regarded as instantaneous (i.e., the steady-state coverage θ^* is reached instantaneously), it was shown that the dynamics of the adsorption of impurity species on the crystal surface cannot always be neglected. This is the reason why, as a first phenomenological approximation, the transient behavior of the coverage process was proposed by Kubota (2001) to obey a first-order dynamics:

$$\theta = \theta^* (1-\exp(-t/\tau)) \tag{13}$$

where τ is the time constant of the coverage dynamic process.

As the crystal growth rate is usually assumed to be proportional to the step velocity, it finally turns out that G depends on both time and supersaturation while in "usual" crystallization approaches dealing with pure media, G is assumed to depend only on σ(t). Combining equations (10) to (13) leads to the following expression where ν is the time at which the crystal surface is set in contact with impure liquid phase (i.e. the time of nucleation):

$$\begin{aligned} G(t) &= k_g \sigma(t)^i \left(1 - \alpha \frac{KC_i}{1+KC_i}\left[1 - \exp\left(-\frac{(t-\nu)}{\tau}\right)\right]\right) \\ &= G_0(t)\left(1 - \alpha \frac{KC_i}{1+KC_i}\left[1 - exp\left(-\frac{(t-\nu)}{\tau}\right)\right]\right) \end{aligned} \tag{14}$$

2.2 Expression of the PBEs accounting for Kubota-Mullin's model of impurities adsorption.

Applying the previous impurity adsorption model (14) is not straightforward as it increases the dimension of the problem: the time (t-ν) spent by the crystals in contact with impurities should now be accounted for. To this effect, we introduce a population density function ϕ depending on the "classical" variables, L and t, as well as ν:

$$\begin{cases} \dfrac{\partial \phi(L,t,\nu)}{\partial t} + G(t,\nu)\dfrac{\partial \phi(L,t,\nu)}{\partial L} = 0 & (15) \\ \phi(L,0,\nu) = 0 & (16) \\ \phi(0,t,\nu) = \dfrac{R_N(t)}{G(t,\nu)}\delta(t-\nu) & (17) \end{cases}$$

The standard definition of the crystal size distribution can still be retrieved as:

$$\psi(L,t) = \int_0^\infty \phi(L,t,\nu)d\nu = \int_0^t \phi(L,t,\nu)d\nu \tag{18}$$

3. A METHOD OF CHARACTERISTICS FOR SOLVING POPULATION BALANCE EQUATIONS ACCOUNTING FOR IMPURITY EFFECTS.

3.1 A method of characteristics for monodimensional PBEs without impurities.

Actually, the supersaturation σ(t) given by Eq. (9), is the driving force of the crystallization process. The decrease of the solute concentration $C(t)$ is caused by the generation of crystals: the molecules of solute initially present in the liquid phase are transferred through crystallization to the dispersed solid phase. The total amount of solid is therefore given by the total volume of solid after integrating the whole CSD:

$$\begin{aligned} C_s(t) &= \rho_s \varphi_p \int_{L^*}^\infty \psi(L,t) L^3 dL \\ &\cong \rho_s \varphi_p \int_0^\infty \psi(L,t) L^3 dL \end{aligned} \tag{19}$$

where ρ_s (kg/m^3) is the density of the solid compound, and φ_p is a volumetric particle shape factor (φ_p is equal to $\pi/6$ if one assumes ideally spherical particles.)

An elementary mass balance of the solute allows computing the evolutions of $C(t)$ and consequently yields σ(t) through (9), provided that the solubility curve is known.

Assuming first that the crystallization takes place in pure solvent, the PBE system (6-8) is expressed as follows where the growth rate $G(t)$ is a complex function of physical and kinetic variables depending on σ(t) and, through the indirect size-dependency of the solute concentration $C_s(t)$, on the overall current size distribution $\psi(L,t)$:

$$\begin{cases} \dfrac{\partial \psi(L,t)}{\partial t} + G(t)\dfrac{\partial \psi(L,t)}{\partial L} = 0 & (20) \\ \psi(L,0) = 0 & (21) \\ \psi(0,t) = \dfrac{R_N(t)}{G(t)} & (22) \end{cases}$$

In the sequel, it is clear that the process is operated in supersaturated conditions (i.e. $\sigma > 0$), the following condition is therefore always fulfilled:

$$\forall t \in \mathbb{R}^+, \forall \nu \in [0,t\,], \quad G(t) > 0 \tag{23}$$

Now, the following characteristic curves are considered:

$$\forall t \in \mathbb{R}^+, \quad \forall \nu \in [0,t\,] \quad , \quad \tilde{L}_\nu(t) = \int_\nu^t G(t')dt' \tag{24}$$

As represented in Figure 2, the CSD along a given characteristic curve is defined as follows:

$\tilde{\psi}_\nu(t) = \psi\left(\tilde{L}_\nu(t), t\right)$, so that one can write:

$$\frac{d\tilde{\psi}_\nu(t)}{dt} = \frac{\partial\psi\left(\tilde{L}_\nu(t),t\right)}{\partial t} + \frac{d\tilde{L}_\nu(t)}{dt}\ \frac{\partial\psi\left(\tilde{L}_\nu(t),t\right)}{\partial L}$$
$$= \frac{\partial\psi\left(\tilde{L}_\nu(t),t\right)}{\partial t} + G(t)\ \frac{\partial\psi\left(\tilde{L}_\nu(t),t\right)}{\partial L}$$
$$\frac{d\tilde{\psi}_\nu(t)}{dt} = 0 \qquad (25)$$

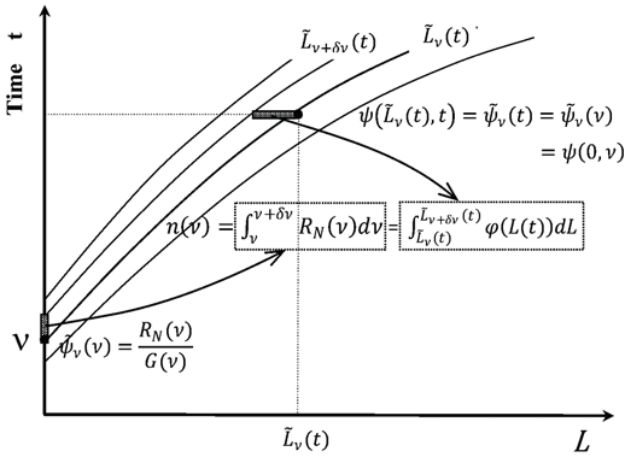


Fig. 2. Schematic representation of the relationship between the number of particles nucleated at time ν and the overall distribution at time t.

It therefore turns out that $\tilde{\psi}_\nu$ does not depend on t, which implies that the solution of (20) is fully determined by the boundary condition (22) and the resolution of (24) describing the time evolutions of the characteristic curves:

$$\forall t \in \mathbb{R}^+,\ \ \forall\, \nu \in [0,t\,]\ \ , \qquad (26)$$
$$\psi\left(\tilde{L}_\nu(t),t\right) = \tilde{\psi}_\nu(\nu) = \tilde{\psi}_\nu(t) = \psi(0,\nu)$$
$$\psi\left(\tilde{L}_\nu(t),t\right) = \frac{R_N(\nu)}{G(\nu)} \qquad (27)$$

Now, let us show that equation (27) allows determining the CSD for every size and time, *i.e.* every point (L,t) in the phase space can be represented as $\left(\tilde{L}_\nu(t),t\right)$. To this effect, the following application is considered:

$\lambda_t\colon [0,t\,] \to \mathbb{R}^+, \quad \nu \mapsto \tilde{L}_\nu(t)$

$\forall t \in \mathbb{R}^+$, λ_t is clearly continuous and

$\forall t \in \mathbb{R}^+,\ \forall\, \nu \in [0,t\,]\ ,\ \frac{d\lambda_t}{d\nu}(\nu) = -G(\nu)$

which, given (23), shows that λ_τ is strictly decreasing and therefore invertible from [0,*t*] to [0, $\tilde{L}_0(t)$]. It follows that the characteristic curves do not exhibit shock or rarefaction.

A means of computing the distribution density function is therefore given by:

$\forall t \in \mathbb{R}^+, \forall L \in [0,\tilde{L}_0(t)]$,

$$\psi(L,t) = \frac{R_N(\lambda_t^{-1}(L))}{G\left(\lambda_t^{-1}(L)\right)} \qquad (28)$$

3.2 Semi-discretization of the size population density function.

Considering successive sampling times, the time variable ν is discretized as follows:

$\forall t \in \mathbb{R}^+, \forall\, \nu \in [0,t\,],\ L_i(t) = \tilde{L}_{\nu_i}(t)$

$$\Psi_{\mathrm{i}}(\mathrm{t}) = \int_{\mathrm{L_i(t)}}^{\mathrm{L_{i-1}(t)}} \psi(L,t)dL \qquad (29)$$

From (28), and setting the following change of coordinates:

$$\begin{cases} \nu = \lambda_t^{-1}(L) \\ L = \tilde{L}_\nu(t), \qquad \dfrac{dL}{d\nu} = -G(\nu) \end{cases}$$

we get:

$$\Psi_{\mathrm{i}}(\mathrm{t}) = \int_{\mathrm{L_i(t)}}^{\mathrm{L_{i-1}(t)}} \psi(L,t)\,dL = \int_{\nu_{\mathrm{i-1}}}^{\nu_{\mathrm{i}}} \psi\left(\tilde{L}_\nu(t),t\right) G(\nu)\,d\nu = \int_{\nu_{\mathrm{i-1}}}^{\nu_{\mathrm{i}}} R_N(\nu)\,d\nu \qquad (30)$$

As illustrated by Fig.2, it finally turns out that integrating ψ in size between L_i (t) and $L_{i\text{-}1}$ (t), at a given time *t,* amounts to integrating ψ in ν between $\nu_{\text{i-1}}$ and ν_{I} , for a given size, and that the result of this integration does not depend on time t.

Consequently, one simply has now to solve the following two systems which are coupled by the growth rate *G*:

$$\begin{cases} \dfrac{dL_i}{dt}(t) = G(t) \\ L_i(\nu_i) = 0 \end{cases} ; \quad \begin{cases} \dfrac{d\Psi_i}{dt}(t) = 0 \\ \Psi_i(\nu_i) = \int_{\nu_{i-1}}^{\nu_i} R_N(\nu)d\nu \end{cases} \qquad (31)$$

3.3 Time-discretization.

Actually, any time-discretization algorithm can be used to solve jointly the two coupled systems defined by (31). As an example, *t* can be discretized in the same way as ν, which leads to the very simple numerical scheme displayed below.

Any numerical integration technique can also be used for the computation of the time variations of both *L* and ψ. The global accuracy of the final numerical solution will mostly be limited by the order of the applied integration scheme.

$k = 0$;
while $t_k < t_{end}$ do

  for $i = 1, k-1$ do
    $L_i(t_k) = L_i(t_{k-1}) + \int_{t_{k-1}}^{t_k} G(t)\,dt$;
  end

  $L_k(t_k) = 0$;
  $\Psi_k = \int_{t_{k-1}}^{t_k} R_N(\nu)\,d\nu$;

  $k = k + 1$;
end

3.4 Method of characteristics for monodimensional PBEs accounting for impurity effects.

Now, in order to account for the distribution of growth rates resulting from the adsorption of impurities, the general system (15-18) is considered. As already explained, the growth rate $G(t)$ is a complex function of physical and kinetic variables depending on time and, through the indirect size-dependency of the solute concentration, on the whole current size distribution:

$$G(t) = f\left(t, \nu, C^*, C_s \propto \int_0^\infty \phi(L', t, \nu) L'^3\,(t) dL'\right) \tag{32}$$

where C_S is the overall concentration of crystallized solid given by (19).

The nucleation time of every crystal is introduced in (32) because, as explained in Part 2, the growth rate G now depends on the time spent by the growing crystal surface in the presence of adsorbing impurities.

As in the "classical" case, it is obvious in the following that during the crystallization process the supersaturation remains positive:

$$\forall t \in \mathbb{R}^+, \forall\, \nu \in [0, t\,], \quad G(t, \nu) > 0 \tag{33}$$

Now, let us consider characteristic curves defined as follows:

$$\forall t \in \mathbb{R}^+, \ \ \forall\, (\nu, \mu) \in [0, t\,]^2,$$
$$\tilde{L}_\mu(t, \nu) = \int_\mu^t G(t', \nu) dt' \tag{34}$$

The distribution along a given characteristic curve is noted as follows:

$$\tilde{\phi}_\mu(t, \nu) = \phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big),$$

and one can write:

$$\frac{\partial \tilde{\phi}_\mu(t, \nu)}{\partial t} = \frac{\partial \phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big)}{\partial t} + \frac{\partial \tilde{L}_\mu(t, \nu)}{\partial t}\ \frac{\partial \phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big)}{\partial L}$$
$$= \frac{\partial \phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big)}{\partial t} + G(t, \nu)\ \frac{\partial \phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big)}{\partial L} \tag{35}$$

$$\frac{\partial \tilde{\phi}_\mu(t, \nu)}{\partial t} = 0 \tag{36}$$

As before, it therefore turns out that $\tilde{\phi}_\mu(t, \nu)$ does not depend on t. It can also be concluded that:

$$\forall t \in \mathbb{R}^+, \ \ \forall\, (\nu, \mu) \in [0, t\,]^2,$$

$$\phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big) = \tilde{\phi}_\mu(t, \nu) = \tilde{\phi}_\mu(\mu, \nu) = \phi(0, \mu, \nu) \tag{37}$$

$$\phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big) = \delta(\mu - \nu)\frac{R_N(\mu)}{G(\mu, \nu)} \tag{38}$$

$$\phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big) = \delta(\mu - \nu)\frac{R_N(\mu)}{G(\mu, \nu)} \tag{39}$$

3.5 Semi-discretization of the size population density function taking the nucleation time into account.

The time variable μ is discretized considering successive sampling times: $\mu \in \{\mu_i, i \in \mathbb{N}\}$, and one can define the following distribution function:

$$\forall \mathrm{t} \in \mathbb{R}^+, \forall\, \nu \in [0, \mathrm{t}\,]$$
$$L_i(t) = \tilde{L}_{\mu_i}(t, \nu) \tag{40}$$
$$\Phi_\mathrm{i}(\mathrm{t}, \nu) = \int_{\mathrm{L_i(t,\nu)}}^{\mathrm{L_{i-1}(t,\nu)}} \phi(\mathrm{L}, \mathrm{t}, \nu)\mathrm{dL}$$
$$\Psi_\mathrm{i}(\mathrm{t}) = \int_0^\mathrm{t} \Phi_\mathrm{i}(\mathrm{t}, \nu)\mathrm{d}\nu \tag{41}$$

Using (40) and setting:

$$\begin{cases} \mu = \lambda_{t,\nu}^{-1}(L) \\ L = \tilde{L}_\mu(t, \nu), \end{cases}$$

$$L = \tilde{L}_\mu(t, \nu), \ \frac{dL}{d\mu}(\mu) = -G(\mu, \nu) \tag{42}$$

yields:

$$\Phi_\mathrm{i}(\mathrm{t}, \nu) = \int_{\mathrm{L_i(t,\nu)}}^{\mathrm{L_{i-1}(t,\nu)}} \phi(L, t, \nu) dL \tag{43}$$
$$= \int_{\mu_{\mathrm{i-1}}}^{\mu_\mathrm{i}} \phi\big(\tilde{L}_\mu(t, \nu), t, \nu\big) G(\mu, \nu) d\mu \tag{44}$$

$$\Phi_i(t,\nu) = \int_{\mu_{i-1}}^{\mu_i} \delta(\mu-\nu)\, R_N(\mu)\, d\mu \quad (45)$$

$$= \begin{cases} R_N(\nu) & if\ \nu \in [\mu_{i-1};\mu_i] \quad (46) \\ 0 & otherwise \quad (47) \end{cases}$$

It follows that:

$$\Psi_i(t) = \int_0^t \Phi_i(t,\nu)d\nu = \int_{\mu_{i-1}}^{\mu_i} R_N(\nu)d\nu \quad (48)$$

The following two systems are thus obtained which are coupled by the growth rate G:

$$G(t) = f\left(t, \mu_i, C^*, C_s \propto \sum_i \Psi_i(\mu_i)\, L_i^3(t)\right) \quad (49)$$

$$\begin{cases} \frac{dL_i}{dt}(t) = G(t) \\ L_i(\mu_i) = 0 \end{cases} \quad \begin{cases} \frac{d\Psi_i}{dt}(t) = 0 \quad (50) \\ \Psi_i(\mu_i) = \int_{\mu_{i-1}}^{\mu_i} R_N(\mu)\, d\mu \quad (51) \end{cases}$$

The principle of the resolution numerical method is the same as previously (see Fig. 2 and Part 3.3).

4. APPLICATION: SIMULATION OF THE CRYSTALLIZATION OF CITRIC ACID MONOHYDRATE IN THE PRESENCE OF IMPURITIES.

In order to illustrate the resolution method, the crystallization of citric acid monohydrate is simulated using kinetic data previously published by Févotte *et al*. (2007). In the absence of reported experimental results in impure media, the parameters of the Kubota-Mullin model were set arbitrarily in order to compare the features of crystallization operations performed with and without impurities. The corresponding parameters are summarized in Table (1).

In the following, no effect of the impurities on the nucleation kinetics is simulated, which is probably a very rough assumption. Actually published data about nucleation in the presence of impurities are really lacking and the goal here is rather to show the usefulness of the resolution method than to investigate real solute/solvent/impurity systems.

Isothermal desupersaturation crystallization operations were simulated at 15°C. In order to initiate the crystallization in the supersaturated zone (i.e. $C_{init.}$ >C* at 15°C)), a seed mass of 10 kg (2% of the expected final mass of solid) is supposed to be introduced in a 1 m^3 pilot-scale well-mixed crystallizer initially feed with a supersaturated citric acid solution. The intial solute concentration is: C_{init} = 1.825 kg/kg water. After seeding, the number of particles increases, due to secondary nucleation, and the initial supersaturation is consumed through the growth of crystals. Despite the adsorption of impurity species at the crystal surfaces, the overall concentration C_i is assumed to remain constant during the crystallization (i.e. the amount of adsorbed molecules is clearly negligible with respect to the dissolved impurities.

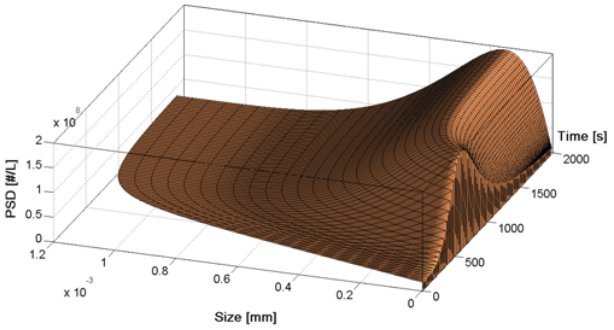


Fig. 3. Simulation of seeded isothermal crystallization of citric acid monohydrate in pure water at 15°C.

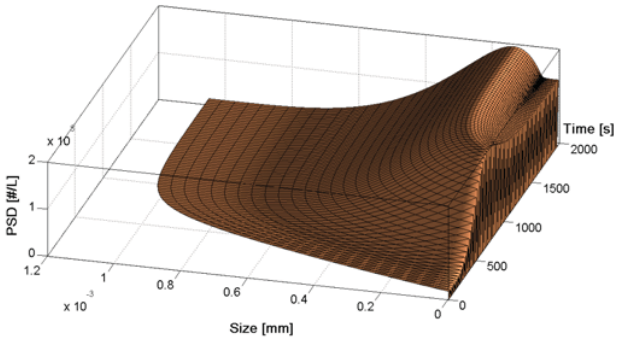


Fig. 4. Simulation of seeded isothermal crystallization of citric acid monohydrate in water and in the presence of impurity at 15°C. The parameters of Kubota-Mullin's are given in Table 1.

As one can see in Figs. 3 and 4, the computed CSD is smooth and does not exhibit oscillatory behaviour, even when coarse time intervals are used for the numerical simulation. As expected, the presence of impurities has a clear effect on the development of the CSD. Fig. 3 shows that the size of the biggest particles obtained in pure water is about 1.2 mm while it is only 1 mm in the presence of impurities (Fig.4). However, Fig.5c shows that the main difference between the two final CSDs can be observed in a rather significant increase in the number of fines which is expected to have a very detrimental effect on the the downstream operation such as filtration.

It turns out that the most significant effects of the impurities on the development of the batch process arise from the reduction of the supersaturation decrease, as displayed in Fig.5a.

Indeed, as outlined in Part 1.2, due to the pinning mechanism, the level of supersaturation remains higher when impurities are present in the crystallizing solution. As during the present simulation no impurity effect is assumed to affect the secondary nucleation of new citric acid particles, higher levels of supersaturation lead to a much higher overall

number of particles (Fig.5d) while, due to growth rate reductions and to the final supersaturation threshold outlined previously, the overall production of solid is clearly reduced. Figure 5b shows that only 70% of the expected solid is obtained at the end of the batch process performed in the presence of impurities (0.5 kg/L of crystals was expected from the selected values of C_{init} and C*).

Table 1. Kinetic equations and parameters used for the simulation of the crystallization of Citric Acid monohydrate.

$$\frac{dL}{dt} = G(t) = k_g(C(t) - C_m^*)^i$$
$$= 7.18\ 10^{-6}(C(t) - C_m^*)^{1.58} \quad (52)$$
$$R_N(t) = K_2\ C_S(t)^{i_m}(C(t) - C_m^*)^{j_m}$$
$$= 1.72\ 10^8\ C_S(t)^{0.47}(C(t) - C_m^*)^{1.14} \quad (53)$$

where $R_N(t)$ is the rate of secondary nucleation of monohydrate citric acid,

K_2 is a "lumped" kinetic constant for secondary nucleation,

C_m^* is the solubility of monohydrate citric acid at 15°C (1.35 kg/kg of water),

i & j_m are exponents expressing the supersaturation dependency of the growth rate and nucleation rate,

i_m is the exponent accounting for the impact of solid already present in suspension on the secondary nucleation rate.

$$C_s(t) = \rho_s \varphi_p \int_{L^*}^{\infty} \psi(L,t) L^3 dL$$

where ρ_s= 1545 kg/m^3

φ_p= π/6 (spherical particles)

Parameters of Kubota-Mullin's Model:

K = 1 m^3/kg
C_i = 0.01 kg/m^3
τ = 500 s
α = 10/σ

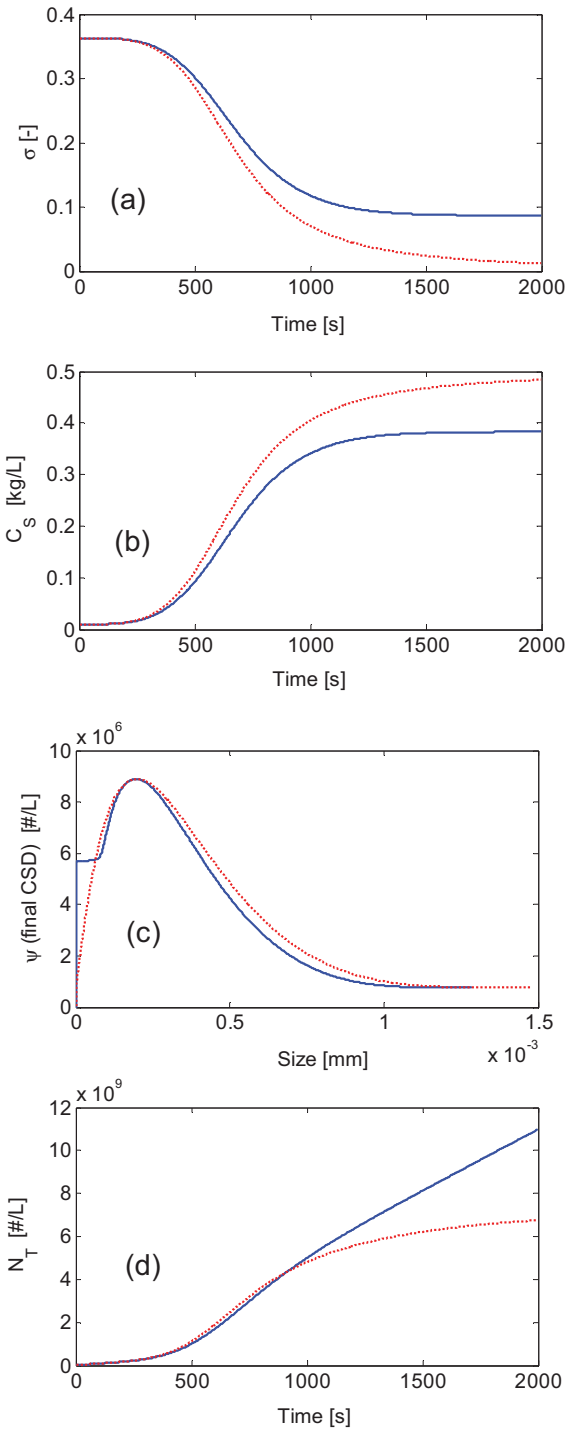


Fig. 5 . Simulation of seeded isothermal crystallization of citric acid monohydrate in water with (dotted line) and without (continuous line) the presence of impurities, at 15°C.

(a) Desupersaturation profile. (b) Generation of total crystallized solid during time. (c) Nucleation rate assuming negligible effect of the impurity species on the generation of particles. (d) Overall number of crystals.

5. CONCLUSIONS

A method of characteristics for the resolution of population balance equations was developed and evaluated using published kinetic data on the crystallization of citric acid in water. The method can be applied to crystallization processes without agglomeration and breakage, and it is intended to allow the simulation of growth rate reductions observed during solution crystallizations performed in the presence of industrial impurities.

Indeed, the effect of impurities was shown by Kubota and Mullin (1995) to depend on the time spent by every crystal in the impure liquid medium and, consequently, to depend on the "age" of the crystals. Such a particular problem requires accounting for an additional time variable in the expression of the PBEs and finding a way of solving the resulting PBE system.

From a physical viewpoint, the simulation results are shown to be consistent and demonstrate the ability of the model to simulate the development industrial crystallization processes in the presence of impurities. Such simulation could be applied, for example, to the design of optimal temperature trajectories aimed at minimizing the detrimental effect of the concentration of impurities on the yield of industrial crystallization operations.

As outlined by several authors (Kumar and Ramkrishna, 1997; Briesen, 2006), despite the apparent simplicity of these two processes, the discretization of crystal nucleation and growth raises numerical diffusion and stability issues which arise from the hyperbolic features of the governing equations (6) to (8). From this latter viewpoint it is clear that the proposed resolution method allows one to account for nucleation and growth rates in a very straightforward way.

NOMENCLATURE

C	Solute concentration	kg solute/kg solvent
C^*	Solubility concentration	kg solute/kg solvent
C_i	Impurity concentration	$kg.m^{-3}$
C_S	Solid concentration	$kg.m^{-3}$
G	Growth rate	$m.\ s^{-1}$
i	Exponent of the supersaturation dependency of the crystal growth rate	[-]
i_m	Exponent of the dependency of the nucleation rate on the concentration of solid in suspension	[-]
j_m	Exponent of the supersaturation dependency of the nucleation rate	[-]
K	Langmuir's constant	$m^3.\ kg^{-1}$
k_g	Growth rate constant	[-]
K_2	Kinetic nucleation parameter	
L	Particle size	m
$\mathcal{N}(\boldsymbol{r},t)$	Number of particles at time t in a given subset	$\#.m^{-3}$
R_N	Nucleation rate	$\#.s^{-1}.m^{-3}$
t	Time	s
u	Step velocity	$m.\ s^{-1}$
	Greek letters	
α	Impurity effectiveness factor	[-]
θ	Fraction of coverage of growing crystal surface by adsorbed impurity	[-]
θ^*	Fraction of coverage of growing crystal surface by adsorbed impurity at the equilibrium	[-]
ν	Nucleation time	s
σ	Supersaturation	[-]
τ	Adsorption time constant	s
$\psi(L,t)$	Population density function	$\#.m^{-1}.m^{-3}$

REFERENCES

Briesen, H. (2006) Simulation of crystal size and shape by means of a reduced two-dimensional population balance model. *Chem. Engng. Sci.,*61, 104-112.

Cabrera, N. and Vermilyea, D.A. (1958). in *Growth and Perfection of Crystals*, New York, 393 p.

Chernov, A.A. (2004). Notes on interface growth kinetics 50 years after Burton, Cabrera and Frank. *J. of Crystal Growth,* 264(4), 499-518.

Févotte, G., Caillet, A. and Sheibat-Othman, N. (2007). A population balance model of the solution-mediated phase transition of citric acid. *AIChE Journal*, 53(10), 2578-2589.

Kubota, N. (2001). Effect of Impurities on the Growth Kinetics of Crystals. *Cryst. Res. Technol*., 36(8-10), 749–769.

Kubota, N. (2001). Effect of Impurities on the Growth Kinetics of Crystals. *Cryst. Res. Technol*., 36, 749–769.

Kubota, N., Yokota, M. and Mullin, J.W. (2000). The combined influence of supersaturation and impurity concentration on crystal growth. *J. Crystal Growth*, 212(3-4), 480-488.

Kubota, N. and Mullin, J.W. (1995). A kinetic model for crystal growth from aqueous solution in the presence of impurity. *J. Crystal Growth,* 152(3), 203-208.

Kubota, N., Yokota, M. and Mullin, J.W. (1997). Supersaturation dependence of crystal growth in solutions in the presence of impurity. *J. Crystal Growth*, 182(1-2), 86-94.

Kumar, S. and Ramkrishna, D. (1996). On the solution of population balance equations by discretization--I. A fixed pivot technique, *Chem. Engng. Sci.,* 51, 8, 1311-1332.

Kumar,S., and Ramkrishna, D. (1996). On the solution of population balance equations by discretization--II. A

moving pivot technique, *Chem. Engng. Sci.,* 51, 8, 1333-1342.

Kumar, S. and Ramkrishna, D. (1997). On the solution of population balance equations by discretization- III. Nucleation, growth and aggregation of particles, *Chem. Engng. Sci*. 52, 24, 4659-4679.

Land, T.A., Martin, T.L., Potapenko, S., Palmore, G.T., De Yoreo, J.J. (1999). Recovery of surfaces from impurity poisoning during crystal growth. *Nature*, 399, 442-445.

Mersmann, A. (2002). Fundamentals of crystallization. *Crystallization Technology Handbook,* 2nd Ed., New York: Marcel Dekker.

Qamar,S.; Warnecke,G. (2007). Numerical solution of population balance equations for nucleation, growth and aggregation processes, *Comp. & Chem. Engng*. 31, 1576-1589.

Mullin, J.W. (1993). *Crystallization,* 3rd Ed., London (UK): Butterworth-Heinemann,

Ramkrishna, D. (2000). Population Balances: *Theory and Applications to Particulate Systems Engineering,* Academic Press.

Ramkrishna, D.; Mahoney, A.W. (2002). Population balance modelling. Promise for the future. *Chem. Engng. Sci*., 57, 595-606.

Sangwal, K. (1996). Effects of impurities on crystal growth processes. *Progress in Crystal Growth and Characterization of Materials*, 32(1-3), 3-43.

Sotowa, K.I., Naito, K., Kano, M., Hasebe, S., Hashimoto, I. (2000). Application of the method of characteristics to crystallizer simulation and comparison with finite difference for controller performance evaluation. *J. of Process Control*, 10, 203-208.

Thomas, T.N., Land, T.A., Martin, T., Casey, W.H., DeYoreo, J.J. (2004). AFM investigation of step kinetics and hillock morphology of the {1 0 0} face of KDP. *J. Crystal Growth,* 260(3-4), 566-579.

Wood, W.M.L. (2001). A bad (crystal) habit--and how it was overcome. *Powder Technology*, 121(1), 53-59.

ACKNOWLEDGEMENTS

We greatly acknowledge the French research agency ANR for the support granted to the project “IPAPI” (Improving the Properties of Active Pharmaceutical Ingredients), ref. 07-BLAN-0183

Micro Process Engineering for Fine Chemistry and Fuel Processing - from Lab to Pilot/Production and First Issues on Dynamic Operation

Volker Hessel[1,2]

[1] *Institut für Mikrotechnik Mainz GmbH (IMM), Mainz, Germany*
[2] *Eindhoven University of Technology (TU/e), Eindhoven, the Netherlands; v.hessel@tue.nl*

This presentation tries to cover the following two aspects.

- To show up how milli and micro process technologies contribute to green processing and process intensification in chemical industry. Relevant trends in the IMM developments (with some supplements of TU/e research) will be given at the focal points catalyst / fabrication / reactor / plant / processes.
- As outlook and much more briefly, first thoughts on process control for micro processing will be given, showing some issues and examples for process analytics in fine chemistry (taken from Paul Watts, University of Hull) and some IMM results about dynamic operation in fuel processing.

Fine Chemical Applications – Process Intensification by Novel Process Windows

Draw on sustainability for chemical production processes demands the integration of sustainability aspects already during process development, whereas further environmental impacts and production costs become predefined. Micro and milli process technologies [1,2] can provide novel ways for process intensification combined with ecological [3] and economic [4] advantages and first assessments were made here [5], mainly by industry. Microstructured reactors have entered the field of fine chemistry with first pilot and production plants; some examples being reported [6]. To bring these innovative apparatus to their operational limit and thus to process with maximal cost competitiveness and environmental sustainability, the idea of "Novel Process Windows" [7] is discussed referring examples of actual research. A recent case study disclosed the key drivers for ecological and economic optimisation [5] when intensifying the aqueous Kolbe-Schmitt synthesis in a minicapillary reactor [8] by using microwaves as alternative energy source and ionic liquids as alternative solvents.

[1] Hessel, Hardt, Löwe / Hessel, Löwe, Müller, Kolb, Chemical Micro Process Engineering (2 Volumes), Wiley-VCH, Weinheim, 2004/2005. *[2]* Jähnisch, V. Hessel, Löwe, Baerns, Angew. Chem. Int. Ed. 2004, 43 (4), 406. *[3]* Kralisch, Kreisel, Chem. Eng. Sci., 2007, 62 (4), 1094. *[4]* Krtschil, Hessel, Kralisch, Kreisel, Küpper, Schenk, Chimia, 2006, 60 (9), 611. *[5]* Hessel, Kralisch, Krtschil, Energy Environ. Sci. 2008, 1 (4), 467. *[6]* Hessel, Löb, Löwe, Industrial Microreactor Process Development up to Production, in: Microreactors in Organic Chemistry and Catalysis (ed. T. Wirth), Wiley-VCH, Weinheim, 2008, pp. 211-275. *[7]* Hessel, Löb, Löwe, Curr. Org. Chem., 2005, 9 (8), 765. *[8]* Hessel, Löb, Löwe et al., Org. Proc. Res. Dev., 2005, 9 (4), 479.

Hydrogen for Fuel Cells by Fuel Processing – Process System Engineering

For catalysis using microreactors [1], it is essential to have total reaction control at all length scales, ranging from mm to nm, which includes smart engineering of reactor plates and microchannels as well as proper setting of catalyst coatings and metal clusters in mesopores. Most essential here is the reliable finding of activated catalysts and proper introduction of these into the microchannels [2-7]. In continuation, catalytic pilot and production microstructured reactors [8,9] demand for solutions on scale-out (system assembly) [8,9], control over flow distribution and heat management at multi-plate architecture [10,11], integration of reaction and heat exchange / separation, process flow with many coupled reactions and operations (sometimes involving recirculation) [8,9], reduction in expenditure for energy and apparatus required for work-up [12], and increase in the service life of the catalyst including concepts for catalyst change, e.g. by replacement of entire modules. This is accompanied by the development of new microfabrication and joining techniques applicable to large, meter-sized format such as rolling / embossing / etching and brazing / diffusion bonding [13]. System and process design is widely practised at IMM for applications in fuel processing for hydrogen production to feed fuel cells [14]. This has led to industrial implementation already, e.g. the 250 W LPG fuel processor-fuel cell *VEGA* for leisure vehicles and boats (Truma Gerätetechnik) and a 2 kW diesel fuel processor-fuel cell prototype as auxiliary power unit for trucks (Volvo, DAF; assembled by Tenneco) [14]. Applications using fossil fuels being reality, the next step is to explore chances of the technology for biofuels with its more complex processing schemes, different logistics and cost structures.

[1] Kolb, Hessel, Chem. Eng. J., 2004, 98 (1-2), 1. *[2]* Muraza, Rebrov, de Croon, Schouten et al., Chem. Eng. J., 2008, 135S, S99. *[3]* Mies, Rebrov, Jansen, de Croon, Schouten, J. Catal., 2007, 247 (2), 328. *[4]* Rebrov, Kuznetsov, de Croon, Schouten, Catal. Today, 2007, 125, 88. *[5]* Pennemann, Hessel, Kolb, Löwe, Zapf, Chem. Eng. J., 2008, 135 (1), S66. *[6]* Men, Kolb, Zapf, Hessel, Löwe, Catal. Today 2007, 125 (1-2), 81. *[7]* Men, Kolb, Zapf, Hessel, Löwe, Trans IChemE, 2007, 85 (B5), 1. *[8]* Men, Kolb, Zapf, Tiemann, Wichert, Hessel, Löwe, Int. J. Hydrogen Energy, 2008, 33 (4), 1374. *[9]* Kolb, Schürer, Tiemann, Wichert, Zapf, Hessel, Löwe, J. Power Sources, 2007, 171 (1), 198. *[10]* Rebrov, Ismagilov, Ekatpure, de Croon, Schouten AIChE J., 2007, 53 (1), 28. *[11]* Mies, Rebrov, Deutz, Kleijn, de Croon, Schouten, Ind. Eng. Chem. Res., 2007, 46 (12), 3922. *[12]* Delsman, Uju, de Croon, Schouten, Ptasinski, Energy Int. J., 2006, 31 (15), 3300. *[13]* Hessel, Kolb, Brandner, Microfabrication for Energy Generating Devices and Fuel Processors in: Microfabricated Power Generation Devices (eds.: Mitsos, Barton,), in print (2008). *[14]* G. Kolb, Fuel Processing, Wiley-VCH, Weinheim, 2008.

Integration of Real-time Optimization & and Model Predictive Control ⋆

V. Adetola * **M. Guay** *

* *Department of Chemical Engineering, Queen's University, Kingston, Ontario, Canada (e-mail: martin.guay@chee.queensu.ca)*

Abstract: This paper proposes a controller design approach that integrates RTO and MPC for the control of constrained uncertain nonlinear systems. Assuming that the economic function is a known function of constrained system's states, parameterized by unknown parameters and time-varying, the controller design objective is to simultaneously identify and regulate the system to the optimal operating point. The approach relies on a novel set-based parameter estimation routine and a robust model predictive controller that takes into the effect of parameter estimation errors. A simulation example is used to demonstrate the effectiveness of the design technique.

Keywords: Adaptive control, Real-time optimization, Model predictive control

1. INTRODUCTION

In this paper, we provide a formal design technique that integrates RTO and MPC for constrained uncertain nonlinear systems. The framework considered assumes the economic function is a known function of constrained system's states, parameterized by unknown parameters. The objective and constraint functions may explicitly depend on time, which means that our proposed method is applicable to both dynamic and steady state economic optimization. The control objective is to simultaneously identify and regulate the system to the operating point that optimizes the economic function. The control input may also be required to satisfy some constraints.

The method proposed solves the control and optimization problem at the same frequency. This eliminates the ensuing interval of "no-feedback" that occurs between economic optimization and thereby improving disturbance attenuation. The RTO layer is tackled via a computational efficient approach. The constrained economic optimization problem is converted to an unconstrained problem and Newton based optimization method is used to develop an update law for the optimum value. The integrated design distinguishes between the extremum seeking and the adaptive tracking of the reference trajectory.

While many advances have been made in nonlinear systems for the stabilization of one fixed operating point, few attempts have been made to address the stabilization problem for time-varying or non-fixed setpoints. In Magni (2002), a stabilizing nonlinear MPC algorithm was developed for asymptotically constant reference signals. By selecting a prediction horizon that is longer than the time the reference setpoint is assumed to have converged, the constant pre-programmed value is used to design the stabilizing controller parameters, *i.e*, the terminal stability constraint $\mathbb{X}_f$ and terminal penalty W. The result is limited to reference signals that converge to *a-priori* known constant setpoint. The method proposed in Findeisen et al. (2000), combines a pseudo-linearization technique with a nonlinear MPC strategy to stabilize a family of (known and constant) setpoints. While the method provides a possible solution for tracking changing setpoints, such pseudo-linearization transformation and feedback is in general difficult to obtain and involve cumbersome computation.

⋆ The authors would like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada.

2. PROBLEM DESCRIPTION

Consider a constrained optimization problem of the form

$$\min_{x\in\mathbb{R}^{n_x}} \; p(t,x,\theta) \tag{1a}$$

$$s.t. \qquad c_j(x) \le 0 \qquad j = 1\dots m_c \tag{1b}$$

with θ representing unknown parameters, assumed to be uniquely identifiable and lie within an initially known convex set $\Theta^0 \triangleq B(\theta^0, z_\theta^0)$. The functions p and c_j are assumed to be C^2 in all of their arguments (with locally Lipschitz second derivatives), uniformly for $t \in [0,\infty)$. The constraint $c_j \le 0$ must be satisfied along the system's state trajectory $x(t)$.

Assumption 1. The following assumptions are made about (1).

(1) There exists $\varepsilon_0 > 0$ such that $\frac{\partial^2 p}{\partial x^2} \ge \varepsilon_0 I$ and $\frac{\partial^2 c}{\partial x^2} \ge 0$ for all $(t,x,\theta) \in (\mathbb{R}^+ \times \mathbb{R}^{n_x} \times \Theta^\epsilon)$, where Θ^ϵ is an ϵ neighborhood of Θ.
(2) The feasible set

$$\mathbb{X} = \big\{x \in \mathbb{R}^{n_x} \mid \max_j \; c_j(x) \le 0\big\},$$

has a nonempty interior.

Assumption 1 states that the cost surface is strictly convex in x and $\mathbb{X}$ is a non-empty convex set. Standard nonlinear optimization results guarantee the existence of a unique minimizer $x^*(t,x,\theta) \in \mathbb{X}$ to problem 1. In the case of non-convex cost surface, only local attraction to an extremum could be guaranteed. The control objective is to stabilize the nonlinear system

$$\dot{x} = f(x,\xi,u) + g(x,\xi,u)\theta \triangleq \mathcal{F}(x,\xi,u,\theta) \quad (2a)$$
$$\dot{\xi} = f_\xi(x,\xi) \quad (2b)$$

to the optimum operating point or trajectory given by the solution of (1) while obeying the input constraint $u \in \mathbb{U} \in \mathbb{R}^{n_u}$ in addition to the state constraint $x \in \mathbb{X} \in \mathbb{R}^{n_x}$. The dynamics of the state ξ is assumed to satisfy the following input to state stability condition with respect to x.

Assumption 2. If x is bounded by a compact set $B_x \subseteq \mathbb{X}$, then there exists a compact set $B_\xi \subseteq \mathbb{R}^{n_\xi}$ such that $\xi \in B_\xi$ is positively invariant under 2.

3. EXTREMUM SEEKING SETPOINT DESIGN

3.1 Finite-time Parameter Identification

Let $\hat{x}$ denote the state predictor for (2), the dynamics of the state predictor is designed as

$$\dot{\hat{x}} = f(x,\xi,u) + g(x,\xi,u)\hat{\theta} + k_w(t)e + w\dot{\hat{\theta}}, \quad (3)$$

where $\hat{\theta}$ is a parameter estimate generated via any update law $\dot{\hat{\theta}}$, $k_w > 0$ is a design matrix, $e = x - \hat{x}$ is the prediction error and w is the output of the filter

$$\dot{w} = g(x,\xi,u) - k_w w, \qquad w(t_0) = 0. \quad (4)$$

Denoting the parameter estimation error as $\tilde{\theta} = \theta - \hat{\theta}$, it follows from (2) and (3) that

$$\dot{e} = g(x,\xi,u)\tilde{\theta} - k_w\, e - w\dot{\tilde{\theta}}. \quad (5)$$

The use of the filter matrix w in the above development provides direct information about parameter estimation error $\tilde{\theta}$ without requiring a knowledge of the velocity vector $\dot{x}$. This is achieved by defining the auxiliary variable

$$\eta = e - w\tilde{\theta} \quad (6)$$

with η, in view of (4, 5), generated from

$$\dot{\eta} = -k_w\,\eta, \qquad \eta(t_0) = e(t_0). \quad (7)$$

Based on the dynamics (3), (4) and (7), the main result is given by the following theorem.

Theorem 3. Let $Q \in \mathbb{R}^{n_\theta \times n_\theta}$ and $C \in \mathbb{R}^{n_\theta}$ be generated from the following dynamics:

$$\dot{Q} = w^T w, \qquad Q(t_0) = 0 \quad (8a)$$
$$\dot{C} = w^T(w\hat{\theta} + e - \eta), \qquad C(t_0) = 0 \quad (8b)$$

Suppose there exists a time t_c and a constant $c_1 > 0$ such that $Q(t_c)$ is invertible *i.e.*

$$Q(t_c) = \int_{t_0}^{t_c} w^T(\tau)w(\tau)\,d\tau \succ c_1 I, \quad (9)$$

then

$$\theta = Q(t)^{-1}C(t) \qquad \text{for all } \; t \geq t_c. \quad (10)$$

Proof: The result can be easily shown by noting that

$$Q(t)\,\theta = \int_{t_0}^{t} w^T(\tau)w(\tau)\Big[\hat{\theta}(\tau) + \tilde{\theta}(\tau)\Big]\,d\tau. \quad (11)$$

Using the fact that $w\tilde{\theta} = e - \eta$, it follows from (11) that

$$\theta = Q(t)^{-1}\int_{t_0}^{t} \dot{C}(\tau)\,d\tau = Q(t)^{-1}C(t) \quad (12)$$

and (12) holds for all $t \geq t_c$ since $Q(t) \succeq Q(t_c)$. ■

The result in theorem 3 is independent of the control u and parameter identifier $\dot{\hat{\theta}}$ structure used for the state prediction (eqn 3). Moreover, the result holds if a nominal estimate θ^0 of the unknown parameter (no parameter adaptation) is employed in the estimation routine. In this case, $\hat{\theta}$ is replaced with θ^0 and the last part of the state predictor (3) is dropped ($\dot{\hat{\theta}} = 0$).

Let

$$\theta^c \triangleq Q(t_c)^{-1}\,C(t_c) \quad (13)$$

The finite-time (FT) identifier is given by

$$\hat{\theta}^c(t) = \begin{cases} \hat{\theta}(t), & \text{if } \; t < t_c \\ \theta^c, & \text{if } \; t \geq t_c. \end{cases} \quad (14)$$

3.2 Constraint Removal

An interior point barrier function method is used to enforce the inequality constraint. The state constraint is incorporated by augmenting the cost function p as follows:

$$p_a(t,x,\theta) \triangleq p(t,x,\theta) - \frac{1}{\eta_c}\sum_{j=1}^{m_c}\ln(-c_j(x)) \quad (15)$$

with $\eta_c > 0$, a fixed constant. The augmented cost function (15) is strictly convex in x and the unconstrained minimization of p_a therefore has a unique minimizer in $\text{int}\{\mathbb{X}\}$ which converges to that of (1) in the limit as $\eta_c \to \infty$ Bertsekas (1995).

3.3 Setpoint Update Law

Let $x_r \in \mathbb{R}^{n_x}$ denote a reference setpoint to be tracked by x and $\hat{\theta}$ denote an estimate of the unknown parameter θ. A setpoint update law $\dot{x}_r$ can be designed based on newton's method, such that $x_r(t)$ converges exponentially to the (unknown) $\hat{\theta}$ dependent optimum value of (15). To this end, consider an optimization Lyapunov function candidate

$$V_r = \tfrac{1}{2}\|\frac{\partial p_a}{\partial x}(t,x_r,\hat{\theta})\|^2 \triangleq \tfrac{1}{2}\|z_r\|^2 \quad (16)$$

For the remainder of this section, omitted arguments of p_a and its derivatives are evaluated at $(t,x_r,\hat{\theta})$. Differentiating (16) yields

$$\dot{V}_r = \frac{\partial p_a}{\partial x}\left(\frac{\partial^2 p_a}{\partial x \partial t} + \frac{\partial^2 p_a}{\partial x^2}\dot{x}_r + \frac{\partial^2 p_a}{\partial x \partial \theta}\dot{\hat{\theta}}\right). \quad (17)$$

Using the update law

$$\dot{x}_r = -\left(\frac{\partial^2 p_a}{\partial x^2}\right)^{-1}\left[\frac{\partial^2 p_a}{\partial x \partial t} + \frac{\partial^2 p_a}{\partial x \partial \theta}\dot{\hat{\theta}} + k_r\frac{\partial p_a^T}{\partial x}\right] \triangleq f_r(t,x_r,\hat{\theta}) \quad (18)$$

with $k_r > 0$ and $r(0) = r_0 \in \text{int}\,\{\mathbb{X}\}$ results in

$$\dot{V}_r \leq -k_r\|z_r\|^2, \quad (19)$$

which implies that the gradient function z_r converges exponentially to the origin.

Lemma 4. Suppose $(\theta,\hat{\theta})$ is bounded, the optimal setpoint $x_r(t)$ generated by (18) is feasible and converges to $x^*_{pa}(\hat{\theta})$, the minimizer of (15) exponentially.

Proof: Feasibility follows from the boundedness of $(\theta, \hat{\theta})$ and Assumption 1.1 while convergence follows from (19) and the fact that z_r is a diffeomorphism. ■

4. ONE-LAYER INTEGRATION APPROACH

Since the true optimal setpoint depends on θ, the actual desired trajectory $x_r^*(t, \theta)$ is not available in advance. However, $x_r(t, \hat{\theta})$ can be generated from the setpoint update law (18) and the corresponding reference input $u_r(x_r)$ can be computed on-line.

Assumption 5. $x_r(t, \hat{\theta})$ is such that there exists $u_r(x_r)$ satisfying

$$0 = f(x_r, u_r, \hat{\theta}) \tag{20}$$

The design objective is to design a model predictive control law such that the true plant state x tracks the reference trajectory $x_r(t, \hat{\theta})$. Given the desired time varying trajectory (x_r, u_r), an attractive approach is to transform the tracking problem for a time-invariant system into a regulation problem for an associated time varying control system in terms of the state error $x_e = x - x_r$ and stabilize the $x_e = 0$ state. The formulation requires the MPC controller to drive the tracking error x_e into the terminal set $\mathbb{X}_{e_f}(\tilde{\theta})$ at the end of the horizon. Since the system's dynamics is uncertain, we use the finite-time identifier (34) for online parameter adaptation and incorporate robust features in to the adaptive controller formulation to account for the impact of the parameter estimation error $\tilde{\theta}$ in the design.

4.1 Min-max Adaptive MPC

Feedback min-max robust MPC is employed to provide robustness for the MPC controller during the adaptation phase. The controller maximizes a cost function with respect to θ and minimizes it over feedback control policies κ.

The integrated controller is given as

$$u = \kappa_{mpc}(t, x_e, \hat{\theta}) \triangleq \kappa^*(0, x_e, \hat{\theta}) \tag{21a}$$

$$\kappa^* \triangleq \arg \min_{\kappa(\cdot,\cdot,\cdot,\cdot)} J(t, x_e, \hat{\theta}, \kappa) \tag{21b}$$

where $J(t, x_e, \hat{\theta}, \kappa)$ is the (worst-case) cost associated with the optimal control problem:

$$J(t, x_e, \hat{\theta}, \kappa) \triangleq \max_{\theta \in \Theta} \int_0^T L(\tau, x_e^p, u^p, u_r) d\tau \tag{22a}$$

$$+ W(\tau, x_e^p(T), \tilde{\theta}^p(T)) \tag{22b}$$

s.t. $\forall \tau \in [0, T]$

$$\dot{x}^p = f(x^p, \xi^p, u^p) + g(x^p, \xi^p, u^p)\,\theta, \;\; x^p(0) = x \tag{22c}$$

$$\dot{\xi}^p = f(x^p, \xi^p), \;\; \xi^p(0) = \xi \tag{22d}$$

$$\dot{x}_r^p = f_r(t, x_r, \theta), \;\; x_r^p(0) = x_r \tag{22e}$$

$$x_e^p = x^p - x_r^p \tag{22f}$$

$$\dot{w}^p = \beta(g^T(x^p, \xi^p, u^p) - k_w w^p), \;\; w^p(0) = w \tag{22g}$$

$$\dot{Q}^p = \beta({w^p}^T w^p), \;\; Q^p(0) = Q \tag{22h}$$

$$\dot{\hat{\theta}}^p = \Gamma\, Q^p\, \tilde{\theta}^p, \;\; \tilde{\theta}^p = \theta - \hat{\theta}^p, \;\; \hat{\theta}^p(0) = \hat{\theta} \tag{22i}$$

$$u^p(\tau) \triangleq \kappa(\tau, x_e^p(\tau), \hat{\theta}^p(\tau)) \in \mathbb{U} \tag{22j}$$

$$x_e^p(\tau) \in \mathbb{X}_e, \;\; x_e^p(T) \in \mathbb{X}_{e_f}(\tilde{\theta}^p(T)) \tag{22k}$$

where $\mathbb{X}_e = \{\, x_e^p : x^p \in \mathbb{X} \,\}$, $\mathbb{X}_{e_f}$ is the terminal constraint and $\beta \in \{0, 1\}$. The effect of the future parameter adaptation is incorporated in the controller design via (22a) and (22k), which results in less conservative worst-case predictions and terminal conditions.

4.2 Implementation Algorithm

Algorithm 1. The finite-time min-max MPC algorithm performs as follows: At sampling instant t_i

(1) **Measure** the current states of the plant $x = x(t_i)$, $\xi = \xi(t_i)$ and obtain the current value of the desired setpoint $x_r = x_r(t_i)$ via the update law (18)

(2) **Obtain** the current value of matrices w, Q and C from

$$\dot{w} = g(x, u) - k_w\, w, \;\; w(t_0) = 0, \tag{23}$$

and

$$\dot{Q} = w^T w, \;\; Q(t_0) = 0 \tag{24a}$$

$$\dot{C} = w^T(w\,\theta^0 + x - \hat{x} - \eta), \;\; C(t_0) = 0 \tag{24b}$$

respectively

(3) **If** $det(Q) = 0$ or $cond(Q)$ is not satisfactory update the parameter estimates $\hat{\theta}$ and the uncertainty set $\Theta(t) \triangleq B\left(\hat{\theta}(t), z_\theta(t)\right)$ according to Algorithm 3 in the Appendix.
Else if $det(Q) > 0$ and $cond(Q)$ is satisfactory, set $\beta = 0$ and update

$$\hat{\theta} = Q^{-1}(t_i) C(t_i), \;\; z_\theta = 0$$

End

(4) Solve the optimization problem (21,22) and apply the resulting feedback control law to the plant until the next sampling instant

(5) **Increment** $i = i + 1$. **If** $z_\theta > 0$, repeat the procedure from step 1 for the next sampling instant. **Otherwise**, repeat only steps 1 and 4 for the next sampling instant.

Since the algorithm is such that the uncertainty set Θ contracts over time, the conservatism introduced by the robustness feature in terms of constraint satisfaction and controller performance reduces over time and when Θ contracts upon θ, the min-max adaptive framework becomes that of a nominal MPC. The drawback of the finite-time identifier is attenuated in this application since the matrix invertibility condition is checked only at sampling instants. The benefit of the identifier, however, is that it allows an earlier and immediate elimination of the robustness feature.

4.3 Lipschitz-based Adaptive MPC

While the min-max approach provides the tightest uncertainty cone around the actual system's trajectory, its application is limited by the enormous computation required to obtain the solution of the min-max MPC algorithm. To address this concern, the robust tracking problem is reposed as the minimization of a nominal objective function subject to "robust constraints".

The model predictive feedback is defined as

$$u = \kappa_{mpc}(t, x_e, \hat{\theta}, z_\theta) = u^*(0) \tag{25a}$$

$$u^*(.) \triangleq \arg\min_{u^p_{[0,T]}} J(t, x_e, \hat{\theta}, z_\theta, u^p, u_r) \tag{25b}$$

where $J(t, x_e, \hat{\theta}, z_\theta, u^p, u_r)$ is given by the optimal control problem:

$$J(t, x_e, \hat{\theta}, z_\theta, u^p, u_r) = \int_0^T L(t, x_e^p, u^p, u_r) d\tau \tag{26a}$$

$$+ W(x_e^p(T), z_\theta^p(T)) \tag{26b}$$

$$s.t.\ \forall \tau \in [0,\ T]$$

$$\dot{x}^p = f(x^p, u^p) + g(x^p, u^p)\hat{\theta}, \quad x^p(0) = x \tag{26c}$$

$$\dot{\xi}^p = f(\xi^p, x^p), \quad \xi^p(0) = \xi \tag{26d}$$

$$\dot{x}_r^p = f_r(t, x_r, \hat{\theta}), \quad x_r^p(0) = x_r \tag{26e}$$

$$x_e^p = x^p - x_r \tag{26f}$$

$$\dot{z}_e^p = \beta(\mathcal{L}_f + \mathcal{L}_g \Pi) z_e^p + \|g(x^p, \xi^p, u^p)\| z_\theta, \tag{26g}$$

$$z_x^p(0) = 0 \tag{26h}$$

$$X_e^p(\tau) \triangleq B(x_e^p(\tau), z_e^p(\tau)) \subseteq \mathbb{X}_e, \quad u^p(\tau) \in \mathbb{U} \tag{26i}$$

$$X_e^p(T) \subseteq \mathbb{X}_{e_f}(z_\theta^p(T)) \tag{26j}$$

Since the Lipschitz-based robust controller is implemented in open-loop, there is no setpoint trajectory $x_r(\hat{\theta})$ feedback during the inter-sample implementation. Therefore, the worst-case deviation $z_e^p \geq \max_{\theta\in\Theta} \|x_e - x_e^p\| = \max_{\theta\in\Theta} \|x - x^p\|$. Hence z_e^p given in (26g) follows from

$$\dot{z}_x^p = (\mathcal{L}_f + \mathcal{L}_g \Pi) z_x^p + \|g(x^p, u)\| z_\theta, \quad z_x^p(t_0) = 0 \tag{27}$$

where $\Pi = z_\theta + \|\hat{\theta}\|$. We assume an appropriate knowledge of Lipschitz bounds as follows:

Assumption 6. A set of functions $\mathcal{L}_j : \mathbb{X} \times \mathbb{R}^{n_\xi} \times \mathbb{U} \to \mathbb{R}^+$, $j \in \{f, g\}$ are known which satisfy

$$\mathcal{L}_j(\mathbb{X}, \xi, u) \geq$$
$$\min\left\{\mathcal{L}_j \;\middle|\; \sup_{x_1, x_2 \in \mathbb{X}} \left(\|j(x_1, \xi, u) - j(x_2, \xi, u)\| - \mathcal{L}_j \|x_1 - x_2\|\right) \leq 0\right\},$$

4.4 Implementation Algorithm

Algorithm 2. The finite-time Lipschitz based MPC algorithm performs as follows: At sampling instant t_i

(1) **Measure** the current states of the plant $x = x(t_i)$, $\xi = \xi(t_i)$ and obtain the current value of the desired setpoint $x_r = x_r(t_i)$ via the update law (18)
(2) **Obtain** the current value of matrices w, Q and C from (23) and (24)
(3) **If** $det(Q) = 0$ or $cond(Q)$ is not satisfactory, set $\beta = 1$ and update the parameter estimates $\hat{\theta} = \hat{\theta}(t_i)$ and uncertainty bounds $z_\theta = z_\theta(t_i)$ and $z_\theta^p(T) = z_\theta^p(t_i + T)$ via equation (29)

$$\dot{\hat{\theta}} = \Gamma\,(C - Q\,\hat{\theta}), \quad \hat{\theta}(t_0) = \theta^0, \tag{29}$$

equation (A.1) and equation (30)

$$z_\theta^p(\tau) = \exp^{-\bar{\mathcal{E}}(\tau - t_i)}\, z_\theta(t_i) \quad \tau \in [t_i,\ t_i + T) \tag{30}$$

where

$$\bar{\mathcal{E}} \geq \mathcal{E}(t_i) = \lambda_{\min}\,(\Gamma Q(t_i))\,.$$

Else if $det(Q) > 0$ and $cond(Q)$ is satisfactory, set $\beta = 0$ and update

$$\hat{\theta} = Q^{-1}(t_i) C(t_i), \quad z_\theta = 0$$

End

(4) **Solve** the optimization problem (25,26) and apply the resulting feedback control law to the plant until the next sampling instant
(5) **Increment** $i = i + 1$. **If** $z_\theta > 0$, repeat the procedure from step 1 for the next sampling instant. **Otherwise**, repeat only steps 1 and 4 for the next sampling instant.

Implementing the adaptive MPC control law according to Algorithm 2 ensures that the uncertainty bound z_θ reduces over time and hence, the error margin z_x^p imposed on the predicted state also reduces over time and shrinks to zero when the actual parameter estimate is constructed in finite-time.

4.5 Robust Stability

Robust stability is guaranteed under the standard assumptions that $\mathbb{X}_{e_f} \subseteq \mathbb{X}_e$ is an invariant set, W is a local robust CLF for the resulting time varying system and the decay rate of W is greater than the stage cost L within the terminal set $\mathbb{X}_{e_f}$ in conjunction with the requirement for W to decrease and $\mathbb{X}_f$ to enlarge with decreased parametric uncertainty.

4.6 Enhancing Parameter Convergence

In min-max adaptive formulation, the terminal penalty is parameterized as a function of $\tilde{\theta}$. This ensures that the algorithm will seek to reduce the parameter error in the process of optimizing the cost function and will automatically inject some excitation in the closed-loop system, when necessary, to enhance parameter convergence. However, this is not the case in the Lipschitz-based approach since the control calculation only uses nominal model. To improve the quality of excitation in the closed-loop the proposed excitation cost is

$$J_{\mathcal{E}} = \frac{\beta}{1 + \mathcal{E}_\theta^p(T)} \tag{31}$$

where

$$\mathcal{E}_\theta^p(\tau) = \lambda_{\min}\{Q^p(\tau)\} \quad \text{or} \quad \mathcal{E}_\theta^p(\tau) = \nu^T\, Q^p(\tau)\, \nu \tag{32}$$

with $\nu \in \mathbb{R}^{n_\theta}$ a unit vector. Note that any reduction in the cost function due to $J_{\mathcal{E}}$ implies an improvement in the rank of Q^p. Though, the predicted regressor matrix Q^p differs from the actual matrix Q, a sufficient condition for $Q > 0$ is for $Q^p > z_Q \geq \|Q - Q^p\|$.

5. TWO-LAYER INTEGRATION METHOD

The integration task can also be posed as a two degree of freedom paradigm where the problem is divided into two phases. The first phase deals with generating a state trajectory that optimizes a given objective function while respecting the system's dynamics and constraints, and the second phase deals with the design of a controller that would regulate the system around the trajectory.

The MPC controller design follows that of (21) and (25). The only difference is that rather than solving the setpoint differential equation (18) inside the MPC loop, the measurement of x_r obtained at sampling instants

is used as the desired setpoint to be tracked, that is, equations (22e) and (26e) are replaced by

$$\dot{x}_r^p = 0, \;\; x_r^p(0) = x_r. \tag{33}$$

The adaptive controllers are implemented according to Algorithms 1 and 2.

6. MAIN RESULT

The integration result is provided in the following:

Theorem 7. Consider problem (1) subject to system dynamics (2), and satisfying Assumption 1. Let the controller be (21) or (25) with setpoint update law (18) and parameter identifier (34)

$$\hat{\theta}^c(t) = \begin{cases} \hat{\theta}(t), & \text{if } \; t < t_c \\ Q(t_c)^{-1}\, C(t_c), & \text{if } \; t \geq t_c. \end{cases} . \tag{34}$$

If the invertibility condition (equation 35)

$$Q(t_c) = \int_{t_0}^{t_c} w^T(\tau) w(\tau)\, d\tau \succ c_1 I, \tag{35}$$

is satisfied, then for any $\varrho > 0$, there exists constant η_c such that $\lim_{t\to\infty} \|x(t) - x^*(t,\theta)\| \leq \varrho$, with $x^*(t,\theta)$ the unique minimizer of (1). In addition $x \in \mathbb{X}$, $u \in \mathbb{U}$ for all $t \geq 0$.

Proof: We know from from triangle inequality that

$$\begin{aligned} \|x - x^*(\theta)\| \leq \|x - x_r(\hat{\theta})\| + \|x_r(\hat{\theta}) - x_{pa}^*(\hat{\theta})\| \\ + \|x_{pa}^*(\hat{\theta}) - x^*(\hat{\theta})\| + \|x^*(\hat{\theta}) - x^*(\theta)\| \end{aligned} \tag{36}$$

where $x_{pa}^*(\hat{\theta})$ denotes the unique minimizer of the unconstrained problem (15) for $\theta \equiv \hat{\theta}$. Since the MPC controllers guarantees asymptotic convergence of x_e to the origin, we have $\lim_{t\to\infty} \|x - x_r(\hat{\theta})\| = 0$. Also, it follows from Lemma 4, that $\|x_r(\hat{\theta}) - x_{pa}^*(\hat{\theta})\|$ converges exponentially to the origin. Moreover, it is well established that $x_{pa}^*(\hat{\theta})$ converges continuously to $x^*(\hat{\theta})$ as $\eta_c \to \infty$ (Bertsekas, 1995, Proposition 4.1.1). Therefore there exists a class $\mathcal{K}$ function [1] $\alpha_c(\cdot)$ such that

$$\lim_{t\to\infty} \|x_{pa}^*(\hat{\theta}) - x^*(\hat{\theta})\| \leq \alpha_c\left(\frac{1}{\eta_c}\right). \tag{37}$$

The finite-time identification procedure employed ensures that $\hat{\theta} = \theta$ for all $t \geq t_c$, with $t_c < \infty$ and thus $\lim_{t\to\infty} \|x^*(\hat{\theta}) - x^*(\theta)\| = 0$.

Finally, we have

$$\lim_{t\to\infty} \|x(t) - x^*(t,\theta)\| \leq \alpha_c\left(\frac{1}{\eta_c}\right) \tag{38}$$

and the result follows for sufficiently large η_c. The constraint satisfaction claim follows from the feasibility of the adaptive model predictive controllers. ■

7. SIMULATION EXAMPLE

Consider the parallel isothermal stirred-tank reactor in which reagent A forms product B and waste-product C DeHaan and Guay (2005). The reactors dynamics are given by

$$\begin{aligned} \frac{dA_i}{dt} &= A^{in}\frac{F_i^{in}}{V_i} - A_i\frac{F_i^{out}}{V_i} - k_{i1}A_i - 2k_{i2}A_i^2, \\ \frac{dB_i}{dt} &= -B_i\frac{F_i^{out}}{V_i} + k_{i1}A_i, \\ \frac{dC_i}{dt} &= -C_i\frac{F_i^{out}}{V_i} + k_{i2}A_i^2, \end{aligned}$$

where A_i, B_i, C_i denote concentrations in reactor i, k_{ij} are the reaction kinetic constants, which are only nominally known. The inlet flows F_i^{in} are the control inputs, while the outlet flows F_i^{out} are governed by PI controllers which regulate reactor volume to V_i^0.

The economic cost function is the net expense of operating the process at steady state.

$$\begin{aligned} p(A_i,\, s,\, \theta) = \sum_{i=1}^{2} [(p_{i1}s_i + P_A - P_B)k_{i1}A_iV_i^0 \\ + (p_{i2}s_i + 2P_A)k_{i2}A_i^2V_i^0] \end{aligned} \tag{39}$$

where P_A, P_B denote component prices, p_{ij} is the net operating cost of reaction j in reactor i. Disturbances s_1, s_2 reflect changes in the operating cost (utilities, etc) of each reactor. The control objective is to robustly regulate the process to the optimal operating point that optimizes the economic cost (39) while satisfying the following state constraints $0 \leq A_i \leq 3$, $c_v = A_1^2V_1^0 + A_2^2V_2^0 - 15 \leq 0$ and input constraint $0.01 \leq F_i^{in} \leq 0.2$. The reaction kinetics are assumed to satisfy $0.01 \leq k_i \leq 0.2$.

The two-layer approach was used for the simulation. The setpoint value available at sampling instant is passed down to the MPC controller for implementation. The robustness of the adaptive controller is guaranteed via the Lipschitz bound method. The stage cost is selected as a quadratic cost $L(x_e, u_e) = x_e^T\, Q_x\, x_e + u_e^T\, R_u\, u_e$, with $Q_x > 0$ and $R_u \geq 0$.

Terminal Penalty and Terminal Set Design Let $x = [A_1,\; A_2]^T$, $\theta = [k_{11},\, k_{12},\, k_{21},\, k_{22}]^T$ and $u = [F_1^{in},\; F_2^{in}]^T$, the dynamics of the system can be expressed in the form:

$$\begin{aligned} \dot{x} = \underbrace{-\begin{bmatrix} \dfrac{x_1k_{V1}(\xi_1 - V_1^0 + \xi_3)}{\xi_1} \\ \dfrac{x_2k_{V2}(\xi_2 - V_2^0 + \xi_4)}{\xi_2} \end{bmatrix}}_{f_{p1}} \\ + \underbrace{\begin{bmatrix} \dfrac{A_{in}}{\xi_1} & 0 \\ 0 & \dfrac{A_{in}}{\xi_2} \end{bmatrix}}_{f_{p2}} u - \underbrace{\begin{bmatrix} x_1 & 2x_1^2 & 0 & 0 \\ 0 & 0 & x_2 & 2x_2^2 \end{bmatrix}}_{g} \theta, \end{aligned}$$

where ξ_1, ξ_2 are the two tank volumes and ξ_3, ξ_4 are the PI integrators. The system parameters are $V_1^0 = 0.9$, $V_2^0 = 1.5$, $k_{v_1} = k_{v_2} = 1$, $P_A = 5$, $P_B = 26$, $p_{11} = p_{21} = 3$ and $p_{12} = p_{22} = 1$.

A Lyapunov function for the terminal penalty is defined as the input to state stabilizing control Lyapunov function (iss-clf):

$$W(x_e) = \frac{1}{2}\, x_e^T\, x_e \tag{40}$$

[1] A continuous function $\mu : \mathbb{R}^+ \to \mathbb{R}^+$ is of class $\mathcal{K}$ if it is strictly increasing and $\mu(0) = 0$.

Choosing a terminal controller

$$u = k_f(x_e) = -f_{p_2}^{-1}\Big(- f_{p_1} + k_1 x_e + k_2\, g\, g^T\, x_e\Big), \quad (41)$$

with design constants $k_1,\ k_2 > 0$, the time derivative of (40) becomes

$$\dot{W}(x_e) = -k_1\, x_e^T\, x_e - x_e^T\, g\, \theta - k_2 x_e^T\, g\, g^T\, x_e \quad (42)$$

$$\leq -k_1 \|x_e\|^2 + \frac{1}{4k_2}\|\theta\|^2 \quad (43)$$

Since the stability condition requires $\dot{W}(x_e(T)) + L(T) \leq 0$. We choose the weighting matrices of L as $Q = 0.5I$ and $R = 0$. The terminal state region is selected as

$$\mathbb{X}_{e_f} = \{x_e\ :\ W(x_e) \leq \alpha_e\} \quad (44)$$

such that

$$k_f(x_e) \in \mathbb{U}, \quad \dot{W}(T) + L(T)\ \leq 0, \quad \forall(\theta,\, x_e) \in\ (\Theta,\, \mathbb{X}_{e_f}) \quad (45)$$

Since the given constraints requires the reaction kinetic θ and concentration x to be positive, it follows that

$$\dot{W} + L = -(k_1 - 0.5)\|x_e\|^2 - x_e^T\, g\, \theta - k_2 x_e^T\, g\, g^T\, x_e \leq 0 \quad (46)$$

for all $k_1 > 0.5$ and $x_e > 0$. Moreover, for $x_e < 0$, the constants k_1 and k_2 can always be selected such that (46) is satisfied $\forall\ \theta \in \Theta$. The task of computing the terminal set is then reduced to finding the largest possible α_e such that for $k_f(.) \in \mathbb{U}$ for all $x \in \mathbb{X}_{e_f}$.

The simulation results are presented in Figures 1 to 3. The phase trajectories displayed in Figure 1 shows that the reactor states obeys the imposed constraints while Figure 2 shows that the actual, unknown setpoint cost $p(t, x_r, \theta)$ converges to the optimal, unknown $p^*(t, x^*, \theta)$. Figure 3 shows the convergence of the parameter estimates to the true values.

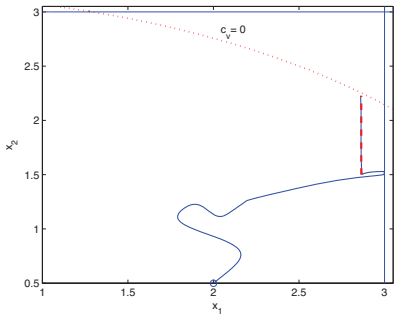


Fig. 1. Phase diagram and feasible state region

8. CONCLUSIONS

This paper provides a formal design technique for integrating RTO and MPC for constrained nonlinear uncertain systems. The solution is based upon the tools and strategies developed in the previous chapters. A single layer and two-layer approaches are presented.

REFERENCES

Dimitri P. Bertsekas. *Nonlinear Programming.* Athena Scientific, Belmont, MA, 1995.

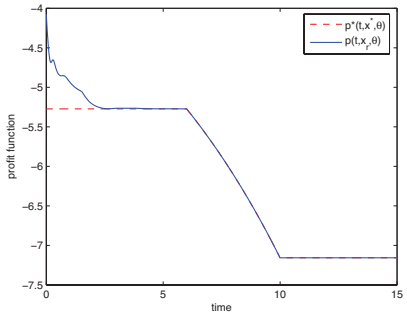


Fig. 2. Optimal and actual profit functions

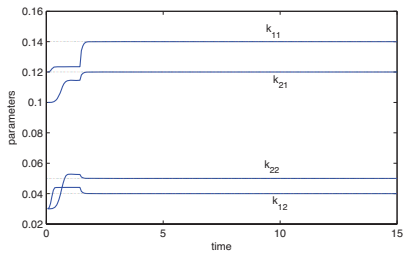


Fig. 3. Unknown parameters and estimates

D. DeHaan and M. Guay. Extremum seeking control of state constrained nonlinear systems. *Automatica*, 41(9): 1567–1574, 2005.

R. Findeisen, H. Chen, and F. Allgöwer. Nonlinear predictive control for setpoint families. In *American Control Conference*, pages 260–264, 2000.

L. Magni. On robust tracking with nonlinear model predictive control. *Int. J. Control*, 75(6):399–407, 2002.

Appendix A. ALGORITHMS

Algorithm 3. Let $\mathcal{E}(\sigma) = \lambda_{\min}(\Gamma Q(\sigma))$, beginning from time $t_{i-1} = t_0$, the parameter and set adaptation is implemented iteratively as follows:

(1) **Initialize** $z_\theta(t_0) = z_\theta^0$, $\Theta(t_0) = B(\hat{\theta}(t_0), z_\theta(t_0))$, $\bar{\mathcal{E}} = \mathcal{E}(t_0) = 0$

(2) **Implement** the following adaptation law over the interval $\tau \in [t_{i-1},\ t_i)$

$$\dot{z}_\theta(\tau) = -\bar{\mathcal{E}} z_\theta(\tau) \quad (A.1)$$

(3) At time t_i, **perform** the updates

$$\bar{\mathcal{E}} = \begin{cases} \mathcal{E}(t_i), & \text{if } \mathcal{E}(t_i) \geq \mathcal{E}(t_{i-1}) \\ \mathcal{E}(t_{i-1}), & \text{otherwise} \end{cases} \quad (A.2)$$

$$\Big(\hat{\theta},\ \Theta\Big) = \begin{cases} \Big(\hat{\theta}(t_i),\ \Theta(t_i)\Big), & \text{if } z_\theta(t_i) - z_\theta(t_{i-1}) \\ & \leq -\|\hat{\theta}(t_i) - \hat{\theta}(t_{i-1})\| \\ \Big(\hat{\theta}(t_{i-1}),\ \Theta(t_{i-1})\Big), & \text{otherwise} \end{cases} \quad (A.3)$$

(4) **Iterate** back to step 2, **incrementing** $i = i + 1$.

Dantzig-Wolfe decomposition for real-time optimization - applied to the Troll west oil rim

Vidar Gunnerud * **Bjarne Foss** * **Bjørn Nygreen** **
Randi Vestbø ** **Nina C. Walberg** **

* *Department of Engineering Cybernetics, NTNU, Trondheim Norway*
** *Department of Industrial Economics and Technology Management, NTNU, Trondheim Norway*

Abstract: This paper studies different decomposition approaches for real-time optimization of process systems with a decentralized structure where the idea is to improve computational efficiency and transparency of a solution. The contribution lies in the application and assessment of the Dantzig-Wolfe method which allows us to efficiently decompose a real-time optimization problem into parts. Furthermore, the nonlinear system is modeled by piecewise linear models with the added benefit that error bounds on the solution can be computed.
The merits of the method are studied by applying it to a semi-realistic model of the Troll west oil rim, a petroleum asset with severe production optimization challenges due to rate dependent gas-coning wells. This study indicates that the Dantzig-Wolfe approach offers an interesting and robust option for complex production systems. Moreover, the method compares favourable with earlier results using Lagrangian relaxation which again was favourable compared to a global approach.

Keywords: Optimization, Dantzig-Wolfe Decomposition, Petroleum production, Plantwide control.

1. INTRODUCTION

Development of a petroleum field asset requires planning on multiple horizons. On a life-cycle horizon, strategic decisions are made on field development such as the choice of technology and export options, and investment and recovery strategies. For offshore assets the choice of technology may include subsea solutions, and the issue of processing the reservoir fluid offshore or onshore (Nygreen et al., 1998). On a medium time horizon, typically three months to two years, production targets are decided. Depending on the life cycle of an asset, decisions may also involve a drilling program. During for instance the green field stage, it is important to plan, drill and commision new wells to reach some pre-defined plateau rate as soon as possible. A reservoir simulator, containing anything between 100.000 and $1,000,000$ states, is usually an important planning tool on the medium time horizon. A reservoir simulator will be quite complex if the geology is complex, due to heteregeneities like faults and shale layers, to represent flow patterns accurately.

On a shorter time horizon, typically days to weeks, production optimization where both the sub-surface part, like the reservoir and wells, *and* the surface part like the manifolds, pipelines and downstream production equipment, is taken into account is important. This is commonly called the real-time production optimization (RTPO) problem. Production may be constrained by reservoir conditions such as coning effects and/or the production equipment like pipeline capacity or downstream water handling capacity, and constraints may move from one part of the system to another part over time. Water production may e.g. be low early on and increase dramatically during the decline phase of a reservoir thereby making water handling capacity an issue. Decision variables in RTPO include production and possibly injection rates, and routing of well streams. A typical production system structure is shown in Fig.1. It has two separate non-connected reservoirs from which 11 wells feed into three manifolds and pipelines, and finally into the downstream facilities section. Manifold 1 and 2 belong to one cluster, while manifold 3 belongs to the other cluster.

RTPO is in use in the upstream industry today. Wang (2003), Saputelli et al. (2003) and Bieker et al. (2006) provide readable overviews. It might be noted that these references focus on the value chain from the reservoir to, and not including, the downstream processing equipment. The downstream boundary is typically a constant pressure on the inlet separator. A few publications on RTPO for the production chain from the reservoir to export are available; Foss and Halvorsen (2009), Selot et al. (2007). Commercial products for RTPO are available, but not widely used. Two of them are GAP and MaxPro. They model the well and near well region, and the pipeline system, and solve the optimization problem using a nonlinear programming (NP) algorithm like sequential quadratic programming (SQP).

There are several factors which complicate the RTPO problem.

- RTPO may give rise to optimization problems with both continuous and discrete decision variables. Discrete decision variables are found in routing when there is a choice to route the fluid from a manifold to one of several flow lines. The presence of discrete decision variables complicates the optimization problem by transforming a linear program (LP) or a nonlinear program (NP) to a mixed integer linear program (MILP) or a mixed integer nonlinear program (MINLP), respectively. Güyaguler and Byer (2007) discusses RTPO in the context of MILP while Kosmidis et al. (2005) use a MINLP formulation.
- The models in the optimization problem are often nonlinear, some of which may be highly nonlinear, as will be discussed later. This includes well models as well as pressure drop models for the pipelines which support multiphase fluid transport.
- The optimization problems are usually quite large and may include several hundred decision variables. An example is the rate allocation problem at Troll which in total includes more than one hundred wells (Hauge and Horn, 2005).

This paper focusses on the RTPO problem for systems with a decentralized structure meaning that common constraints are quite few. Such strucures are quite typical in the upstream petroleum industries as visualized by Fig.1. The contribution lies in the application and assessment of the Dantzig-Wolfe method which allows us to decompose a RTPO problem into parts meaning that we apply a divide-and-conquer strategy which is a sound engineering design principle. This principle has survived ever since complex systems came into making. The Dantzig and Wolfe principle dates back to 1960 (Dantzig and Wolfe, 1960).

A few recent publications apply Dantzig-Wolfe decomposition (DWD) to process systems. Alabi and Castro (2009) apply DWD to a refinery planning problem, formulated as a large LP problem, by decomposing it along the value chain. They show substantial savings in computation time. They also point to the inclusion of binary decision variables, which will be addressed in this paper, as a future task. Cheng et al. (2008) propose DWD as a means for designing a decentralized MPC for plant-wide MPC coordination. Again substantial computational savings are reported for the LP formulation chosen. This paper also gives a very readable introduction to DWD. Both papers state that DWD is particularly well suited for large problems with well-structured subproblems and a small number of linking constraints.

This study will show that the Dantzig-Wolfe approach offers an interesting and robust option for complex production systems wiht certain structural properties. Moreover, the method compares favourable with earlier results using Lagrange relaxation (LR) (Foss et al., 2009) on a realistic field case. The nonlinear system will be modeled by piecewise linear models with the added benefit that error bounds on the solution of the production optimization problem can be computed.

The remainder of this paper is organized as follows. First, the RTPO problem is presented in a mathematical context before the decomposition approach in general and the Dantzig-Wolfe method in particular are presented.

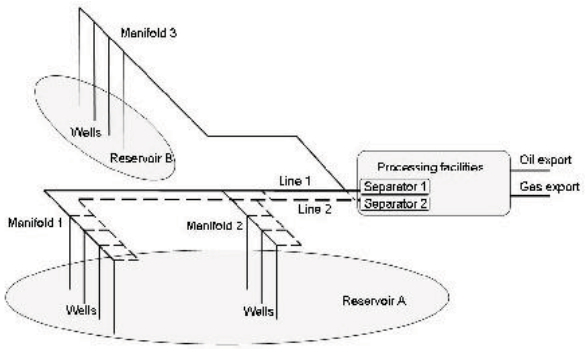


Fig. 1. A petroleum production system with two separate reservoirs from which 11 wells feed into 3 manifold and 3 pipelines. Manifold 1 and 2 belong to one cluster, while manifold 3 belongs to the other cluster. The pipeline flows provide input to the processing facilities where fluids are conditioned for export.

Subsequently, the Troll west oil rim case is presented and results are shown. Finally, results are discussed and some conclusions end the paper.

2. FORMALIZING THE RTPO PROBLEM

The RTPO problem will in most cases mean maximizing oil production while honouring system constraints like capacities in pipelines and wells, safety regulations and preventing damage on long-term effects, in particular recovery of available hydrocarbon resources. The latter point is important. An example of the interplay between short term production and long-term recovery was shown in Naus et al. (2006) in the sense that accelerated short-term production reduced long-term recovery.

The optimization problem is usually treated in a quasi-dynamic way by re-optimizing a stationary optimization problem, typically once a day. The solution of the mathematical RTPO will serve as a recommendation to the operating engineers who may or may not follow the advise. One reason for neglecting a recommendation may hinge on the fact that the transition cost of changing from one routing configuration to another is not included in the optimization problem. Therefore such a change will only be implemented if there is a substantial gain by doing this.

Referring again to Fig.1 to explain upstream systems closer, there are four wells connected to manifold 1 and 3, respectively, and three wells connected to manifold 2. Well streams from each well are connected to one pipeline. Hence, each of the wells in manifold 1 and 2 can be connected to either of the two pipelines transporting the reservoir streams to downstream processing. There is only one pipeline from manifold 3, and therefore no routing decision is necessary in this part of the system. The decision variables on each well are usually one production choke valve to adjust production and on-off valves linking a well to one of the pipelines. Further, different well completions may give rise to additional decision variables like the injected gas-rate for a gas-lift well, a commonly used technology to increase well lifetime as reservoir pressure decreases. Altogether this means that there will be both continuous as well as discrete decision variables in a typical

RTPO problem. The well and pipeline system is divided into clusters. There are two cluster in Fig.1, one covering manifold 1 and 2, and the second includes manifold 3. Hence, a cluster may include one or more manifolds.

We focus on production systems with a decentralized structure where common constraints may include downstream processing capacity limitations and common pipelines. In the following we present a system model which encompasses a large class of upstream production systems. Some simplifications are made to ease the explanation. For instance we assume only one manifold for each cluster in this section. An extension to several manifolds per cluster, as is the case for manifold 1 and 2 in Fig.1, is however straightforward. In fact in the Troll case treated later the clusters have two manifolds each.

We first present a system model for a subsystem, denoted cluster i, before the integrated optimization problem is described.

Indexes, constants and decision variables are explained in Table 1 and 2.

2.1 Modelling a subsystem

In the following we present and comment the model of *one* subsystem, cluster i.

- Mass balance is preserved for each phase, i.e. gas, oil and water, at each node. This means that no phase transition takes place at the surface of a cluster.
$$\sum_{j=1}^{J(i)} q_{ij}^p = q_i^p, \quad p \in \{g, o, w\} \tag{1}$$
- The routing problem is parameterized through binary variables for each well, one for each line; y_{ij}^l. If $y_{ij}^l = 1$ the well is connected to line l, if not it is zero. Each well cannot be connected to more than one line, hence
$$\sum_{l=1}^{L(i)} y_{ij}^l \leq 1, \quad y_{ij}^l \in \{0,1\}, \quad j \in \{1, \ldots, J(i)\} \tag{2}$$
This implies that the flow q_i^p from one cluster is divided onto $L(i)$ pipelines.
- The well model, or performance curve for gas, oil and water, are given by the following nonlinear structure
$$q_{ij}^p = d_{ij}^p(p_{ij}^{res}, p_{ij}^{wh}), \tag{3}$$
$$p \in \{g, o, w\}, \quad j \in \{1, \ldots, J(i)\}$$
where p_{ij}^{res} and p_{ij}^{wh} denotes the reservoir pressure locally at the well and the pressure at the wellhead, respectively. Depending on the reservoir conditions near a well the complexity of these well models vary a lot. The simplest version will be a linear model. In systems with rate-dependent gas coning however, as in the Troll oil case (Hauge and Horn, 2005), nonlinearities can be severe.
- The pressure drop across the production choke is given by
$$y_{ij}^l p_i^l \leq p_{ij}^{wh}, \tag{4}$$
$$j \in \{1, \ldots, J(i)\}, \quad l \in \{1, \ldots, L(i)\}$$
This constraint may only be binding if $y_{ij}^l = 1$ since it is always satisfied for $y_{ij}^l = 0$.
- Flow into the pipelines from cluster i to the platform is given by
$$q_i^{pl} = \sum_{j=1}^{J(i)} y_{ij}^l q_{ij}^p \tag{5}$$
$$p \in \{g, o, w\}, \quad l \in \{1, \ldots, L(i)\}$$
- The pressure drop in a pipeline segment from cluster i to the inlet separator depends nonlinearily on the flow of gas, oil and water in the pipe segment. The nonlinearities are particularly severe during the transition from one multiphase flow regime to another, and when a pipeline exhibits slugging. More on multiphase flow may e.g. be found in Brenne (2005)
$$p^{sep} - p_i^l = d_i^l(q_i^{gl}, q_i^{ol}, q_i^{wl}), \tag{6}$$
$$l \in \{1, \ldots, L(i)\}$$
- There are non-negativity conditions on all flow and pressure variables, i.e. backflow is not modeled.

It should be added that the downstream boundary condition is given by a fixed inlet separator pressure p^{sep}, and we assume that it is equal for all $L(i)$ pipelines. Further, it is straightforward to include additional local constraints like for instance the flowrate from a well due some external reason. This could be well-related problems like sand production, or reservoir based constraints as discussed earlier. Such contraints will typically induce relations like $q_{ij}^o + q_{ij}^w \leq q_{ij}^{\max}$

Table 1. The indexes used.

i	-	cluster i
I	-	no. of clusters
$p \in \{g, o, w\}$	-	phase index - gas, oil or water
ij	-	well j in cluster i
$J(i)$	-	no. of wells linked to cluster i
$l \in \{1, .., L(i)\}$	-	line index for cluster i
$L(i)$	-	no. of lines linked to cluster i

Table 2. The variables and data used to define the sub-problem.

q_i^p	-	total mass flowrate of phase p from cluster i
q_{ij}^p	-	mass flowrate of phase p from well j in cluster i
q_i^{pl}	-	mass flowrate of phase p through line l in cluster i
y_{ij}^l	-	binary variable equal to 1 if well ij is routed to line l
p_{ij}^{res}	-	reservoir pressure at well ij
p_{ij}^{wh}	-	wellhead pressure
d_{ij}^p	-	well performance model
d_i^l	-	pipeline pressure drop model
p_i^l	-	pressure in line l subsea in cluster i
p^{sep}	-	separator pressure

2.2 The integrated problem

The RTPO problem is specified below. The objective function is defined by the total oil production, and the global constraints are given by gas and water handling capacities in the downstream part of the value chain. Hence, the objective function and common constraints are given by

$$max \sum_{i=1}^{I} q_i^o \tag{7}$$

$$\sum_{i=1}^{I} q_i^g \leq \overline{q}^g \tag{8}$$

$$\sum_{i=1}^{I} q_i^w \leq \overline{q}^w \tag{9}$$

The objective function and common constraints are linear and additive and each term $q_i^o = \sum_{j=1}^{J(i)} q_{ij}^o$ is a function of only local variables.

The complete RTPO problem consists of I clusters, each modelled by (1)-(6), or an extension of these equations due to several manifolds in one cluster, and the integration through (7)-(9).

The actual decision variables are production choke openings and on-off valves linking a well to a pipeline. The production choke openings are not directly a part of the optimization problem. They are calculated using the pressure drop across the production choke (4) and the flowrate through the production choke $(q_{ij}^g, q_{ij}^o, q_{ij}^w)$ in an appropriate valve model.

2.3 Piecewise linearization and SOS2 sets

The optimization problem contains both continuous and discrete variables. Furthermore, nonlinear well and pressure drop models are present. Hence, this is basically a MINLP problem. We transform this into a MILP problem by replacing the nonlinear constraints by linear constraints and constraints on some auxiliary integer variables. The procedure is as follows: The nonlinear constraints, (3) and (6), are replaced by piecewise linear approximations. These piecewise linear approximations are modelled by linear constraints and discrete variables, i.e. integer constraints, using Special Ordered Sets of type 2 (SOS2). The discrete variables are necessary to assure interpolation between neighbouring points only, Williams (2005), as in any piecewise linear approximation of a nonlinear function. The number of linear constraints and integer constraints necessary to replace one nonlinear constraint depends on the nonlinearities and approximation accuracy. Higher accuracy means more interpolation points and hence more linear and integer constraints.

3. DECOMPOSITION

3.1 Principle

When a problem becomes too large or complicated to handle, a decomposition approach can be applied if the problem structure is suitable. The basic mechanism in all decomposition principles is to decompose the original problem into smaller sub-problems which are coordinated by a "master" problem. There exists multiple decomposition techniques to solve large problems. Two common methods are Lagrange relaxation and Dantzig-Wolfe decomposition.

Both LR and DWD are suited for problems with a block angular constraint structure which is the case for the RTPO problem described above. The structure is exploited when the original problem is split into sub-problems, while the common constraints remain in the master problem.

In LR (Beasley, 1993) the basic idea is to attach Lagrange multipliers to the common constraints in the model and relax these in the objective function, while DWD handles the common constraints in a master problem. The resulting integrated optimization problem will hence fall apart into I local optimization problems, one for each cluster i (Fisher, 1985). For (convex) LP problem the solution of I such local optimization problems provides the same solution as (1)-(9) provided that the Lagrange multipliers for the common constraints λ^g, λ^w are known. Hence, the Lagrange multipliers put a common cost to the use of a scarce resource by each local problem.

3.2 Dantzig-Wolfe decomposition

When applying DWD to the RTPO problem the sub-problems will be identical to LR. However, while the Lagrange multipliers are updated by a simple heuristic in the LR case, the update is now done by solving an LP-problem.

We start by assuming linear constraints and continous variables, i.e. an LP-problem instead of a MILP problem. The master problem is a reformulation of the integrated problem. By taking advantage of the fact that a convex combination of basic feasible points, which are corner points of the feasible set defined by the linear constraints of the integrated problem, also is a feasible solution, an alternative formulation can be achieved. Each basic feasible point in the integrated problem is then represented as a variable in the master problem. The number of basic feasible points for any practical problem can clearly be prohibitively high, and in reality only a small number of these basic feasible points will ever enter the basis in the master problem. The idea is then to restrict the master problem by reducing the number of basic feasible points. This is called a Restricted Master Problem (RMP).

Hence, we start with a few basic feasible points and check if the solution of the integrated problem is within a convex combination of these points. If this is not the case new basic feasible points are included in a structured way until the optimal solution has been found (Williams, 2005). This is usually called column generation and several procedures are proposed in the literature; either adding one or several columns, i.e. new basic feasible points, at each iteration (Dantzig and Thapa, 2002). Some details of the algorithm are given below with some related comments specific to the RTPO problem.

Algorithm structure

1. Choose two initial basic feasible points for each local optimization problem.

2. Specify the RMP as a LP for the given set of basic feasible points. Then solve it and compute values for the Lagrange multipliers for the global constraints, i.e. λ^g, λ^w. The RMP is specified in a separate section below.

3. Solve I local optimization problems by using the Lagrange multipliers computed in 2.

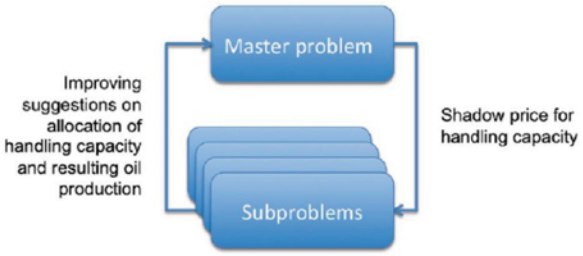


Fig. 2. Iteration structure for Dantzig-Wolfe Decomposition (DWD) and Lagrangian Relaxation (LR)

4a. For $i \in \{1, \ldots, I\}$: If the solution of a local optimization problem i extends the convex set defined by the basic feasible points used in 2, then add these basic feasible points to the RMP, and go to 2. (This implies that the feasible region of this new RMP is extended).

4b. If the solutions of all the local optimization problems are unchanged, the optimal solution has been found; and the algorithm terminates.

The main iteration loop is shown in Fig.2. This figure is also applicable for LR if the master problem box is understood as the updating algorithm for the the Lagrange multipliers for the global constraints.

In view of our RTPO problem item 1 above implies that two feasible solutions for each cluster must be determined to start the algorithm.

3.3 Restricted master- and sub-problem

The procedure is to update the Lagrange multiplier in a way that the consumption of the relaxed common constraints converge to their optimal values. Each subproblem is defined by (again only including two common constraints)

$$\begin{aligned} max \quad & q_i^o - \lambda^g q_i^g - \lambda^w q_i^w - \lambda_i^{CONVEX} \\ & \lambda \geq 0, \quad i \in \{1, \ldots, I\} \end{aligned} \tag{10}$$

and the local constraints (1)-(6). λ_i^{CONVEX} is the Lagrange multiplier for the convexity constraint in the RMP defined below. Since no sub-problem variables are associated with it, it will only act as a constant in the subproblem.

The RMP can now be formulated. z_{is}^p represents one basic feasible point s from sub-problem i. z_{is}^p could in principle include the optimal value of all decision variables for subproblem i after solving it given λ^g and λ^w. However, only the variables also present in the objective function and the common constraints will be relevant for the RMP. Hence, z_{is}^p will for this RTPO problem contain some flow variables (q_i^o, q_i^g, q_i^w), but no pressure variables. μ_{is} is the corresponding weight the master problem will give this basic feasible point. The objective function of the master problem is given in (11). Further, (12) and (13) represents the constrained common resources, while (14) is the convexity constraint.

$$max \sum_i \sum_{s \in S_i} z_{is}^o \mu_{is} \tag{11}$$

$$\sum_i \sum_{s \in S_i} z_{isp}^g \mu_{is} \leq \overline{q}^g \tag{12}$$

$$\sum_i \sum_{s \in S_i} z_{isp}^w \mu_{is} \leq \overline{q}^w \tag{13}$$

$$\sum_{s \in S_i} \mu_{is} = 1 \quad i \in \{1, \ldots, I\} \tag{14}$$

$$\mu_{is} \geq 0 \tag{15}$$

3.4 Integer variables

DWD will find exact optimal solutions for feasible LP problems. If it is extended to a MILP problem, however, Branch & Price (Desrosiers and Lubbecke, 2006) or some heuristics have to be applied to handle the integer properties. When solving the master problem, we have not imposed integer restrictions on μ_{is}, i.e. the RMP is solved as an LP to achieve Lagrange multipliers for (12) - (14). The resulting solution may then be infeasible with respect to the original MILP problem, since a convex combination of two different basic feasible points is not necessarily feasible. As mentioned, this could be handled in several ways. However, if a satisfying number of basic feasible points are generated up front, a feasible solution could simply be found by demanding integer values for μ_{is} and solve the RMP as an MIP problem. Vanderbeck (2006) adresses the use of DWD on mixed integer problems.

3.5 Solution quality

For both LP and MILP problems, upper and lower bounds on the objective function can be computed. The LP solution of the RMP plus the sum of the objective values of the sub-problems will act as an upper bound (Karlof, 2006). In the LP case, the solution of the RMP alone will give a feasible lower bound, while for the MILP problems a heuristic has to be applied to create the feasible lower bound. By using these bounds actively during the optimization process, it is possible to terminate the optimization problem when an acceptable gap is achieved.

4. RESULTS

The Troll field is a huge oil and gas field on the continental shelf west of Norway. Production allocation is complex as described in (Hauge and Horn, 2005). We study the Troll C production system shown in Fig.3 where primarily oil is produced from an oil rim through more than 50 wells. Well models and pressure drop models for multiphase flow in pipelines are based on typical models as encountered in this application. Hence, the models should be understood as approximations of the actual well and pipeline models. Each nonlinear model is approximated by a piecewise linear model. A well model (3) is typically divided into somewhere between 10 and 100 linear segments with the wellhead pressure as its input. The pipeline models require more linear segments since they depend on three inputs, cf. (6). There are 8 clusters, and each cluster has a complex structure in the sense that they contain two manifolds. Each cluster has $6-8$ wells and the total number of wells is 64. For the moment only the gas handling capacity is a binding constraint. Water handling will become an issue

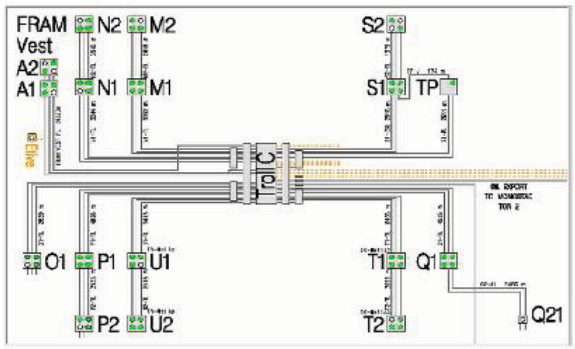


Fig. 3. Topology for the wells connected to the Troll C platform

in the near future as the reservoir drains and therefore produces more water.

The purpose of the numerical study is to investigate the DWD performance compared to a global strategy and the LR method proposed in Foss et al. (2009). Three different strategies were therefore defined:

(1) A global strategy where all clusters are solved in one large MILP problem.
(2) The LR method proposed in Foss et al. (2009).
(3) The DWD method proposed in this paper.

The computations were performed on an IBM Thinkpad T60P with a $2.33GHz$ processor and a tolerance bound of 0.5% for LR and DWD. The state-of-the-art XPress-MP software suite is used to solve the MILP problems. The main results are presented in Table 3. Results are presented column-wise for different system sizes starting with two cluster and ending with the full 64 well/8 cluster system. The gas capacity for each scenario increases with the number of clusters as shown in line 2. In the next three lines the number of variables and constraints are listed. Then follows the results in terms of computation time and oil production which is the ultimate goal, cf. (7).

Finally, it should be noted that the results in Table 3 represent typical values as observed after several test runs.

Table 3. Results from tests on the model of the Troll C production system.

No. of clusters	2	4	6	8
Gas cap.[Sm^3/day]	3000	12000	18000	24000
Continuous variables	13898	27805	41766	55688
Discrete variables	1029	2134	3725	4819
Constraints	491	981	1639	2185
Strategy 1 - Global				
Solution time [min]	0.26	7.38	237.0	720.0
Oil [Sm^3/day]	1777	6487	11641	14365
Strategy 2 - LR				
Solution time [min]	1.42	8.54	18.2	19.2
Oil [Sm^3/day]	1774	6467	11640	14440
Strategy 3 - DWD				
Solution time [min]	1.86	1.43	6.16	11.3
Oil [Sm^3/day]	1777	6458	11629	14473

5. DISCUSSION

The main observation to make is the fact that the decomposition stratgies, DWD and LR, outperform the global method for the combined rate allocation and routing problem for all but the smallest 2 cluster problem. The global method does not converge to its termination criteria after 12 hours for the eight cluster case and in general it has a hard time solving problems consisting of 6 clusters or more. Furthermore, DWD shows superior performance to LR.

By taking a closer look at the 2 cluster problem, the global method is actually fastest. This is not surprising as it is expected that the global method would be faster for small problems. For the medium size problem with 4 clusters, the global method is still working fine, and is actually faster than LR. DWD is in this case extremely quick, due to few main iterations. For the larger problems, we observe that the two decomposition methods are much faster than the global method, and that DWD is significantly faster than LR.

The reason why DWD is faster than LR is related to the updating of Lagrange multipliers. The DWD master problem finds good multipliers with fewer iterations than the LR master problem, and on average converges after fewer iterations. It should be mentioned that the computations involved in solving the DWD master problem, i.e. the LP-problem, is small compared to solving the local MILP-problems. Hence, this is no issue when comparing DWD and LD.

DWD is more stable with respect to solution time than LR. Furthermore, DWD has few tuning parameters, and works well for changing data sets. LR in contrast, is quite sensitive to perturbations of the data set. A minor change might result in a doubling of the solution time. However, extensive knowledge of the problem will give the operator a good feel for which parameter values result in fast convergence.

Focussing on solution quality, we observe that the global method finds the optimal solution for all except the full field problem with 8 clusters. In that case the method was stopped after 12 hours, with still a little more than 7.5% in duality gap. The decomposition methods terminate with less than 0.5% duality gap for all problems. The solution time does of course depend on the resolution of the piecewise linear models. A cruder approximation reduces run-time and vice versa.

DWD provides a framework for decomposing a problem and still keep track of the optimal solution for the integrated problem. This approach has potential advantages in terms of algorithmic efficiency as indicated by the test case in the previous section. The DWD algorithm, as well as LR, has some interesting properties. First, the sub-problems may be solved using different algorithms or even different software packages. This feature is interesting for integrated optimization applications which may encompass reservoir, wells, pipelines and processing facilities. It should be added that the duality gap can only be computed if upper and lower bounds on the solution can be found. This is in general not possible if the sub-problems are nonlinear programs as opposed to MILPs. A second useful property of the algorithmic structure is the potential for parallel computing since each sub-problem is self-contained and has no dependency on the other sub-problems. If the computational load between the sub-

problems is well-balanced a parallel implementation will be particularly efficient.

The optimization problem is usually treated in a quasi-dynamic way by re-optimizing the stationary optimization problem, typically once a day. More frequent disturbances may be handled by selecting a couple of wells for frequent production changes to compensate variations in for instance gas processing capacity. Well models are typically updated twice a year by running well tests to collect data to estimate well parameters. The use of dynamic models is an issue. Some applications may benefit from dynamic well models, in particular during start-up of wells. Start-up may occur quite often since many wells are shut-in from time to time due to maintenance or operational problems. Applications with long pipelines may also benefit from dynamic pipeline models provided the dynamics are important for optimal performance.

6. CONCLUSIONS

This paper argues that Dantzig-Wolfe Decomposition is well suited for the well allocation and routing problem in the upstream industries. There are several reasons for this. DWD clearly outperforms a global method. DWD has several similarities to LR. However, as the results show, DWD gives better performance than LR in all relevant cases tested herein. This is due to more efficient updating of the dual variables. Furthermore, an error bound on the solution of the production optimization problem can easily be computed. This is clearly information of interest to any user. Finally, the algorithm is efficient and can be parallelized for even higher efficiency.

Acknowledgments: We acknowledge the support of the Center for Integrated Operations at NTNU, Norway.

REFERENCES

Alabi, A. and Castro, J. (2009). Dantzig-wolfe and block coordinate-descent decomposition in large-scale integrated refinery planning. *Computers and Operations Research*, 36, 2472–2483.

Beasley, J. (1993). *Modern Heuristics Techniques for Combinatorial Problems*, 243–303. Halsted Press.

Bieker, H., Slupphaug, O., and Johansen, T. (2006). Real-time production optimization of offshore oil and gas production systems: A technology survey. *SPE 99446*.

Brenne, C.E. (2005). Fundamentals of multiphase flow. *Cambridge university press*.

Cheng, R., Forbes, J.F., and Yip, W. (2008). Dantzig-wolfe decomposition and plant-wide mpc coordination. *Computers and Chemical Engineering*.

Dantzig, G.B. and Wolfe, P. (1960). Decomposition pricniple for linear programs. *Operations Research*, 8, 101–111.

Dantzig, G. and Thapa, M. (2002). Linear programming 2: Theory and extensions. *Springer Verlag*.

Desrosiers, J. and Lubbecke, M.E. (2006). *Column Generation*, 1–32. Springer US.

Fisher, M.L. (1985). An application oriented guide to lagrangian relaxation. *Interfaces*, 15, 10–21.

Foss, B., Gunnerud, V., and Dueñas Díez, M. (2009). Lagrangian decomposition of oil production optimization - applied to the Troll west oil rim. *Accepted for SPE Journal*.

Foss, B. and Halvorsen, I.J. (2009). Dynamic optimization of the lng value chain. *Proceedings of the 1st Annual Gas Processing Symposium*.

Güyaguler, B. and Byer, T. (2007). A new production allocation optimization framework. *SPE 105200*.

Hauge, J. and Horn, T. (2005). The challenge of operating and maintaining 115 subsea wells on the troll field. *Offshore Technology Conference*.

Karlof, J.K. (2006). *Integer Programming: Theory and Practice*, chapter 4: Decomposition in Integer Linear Programming. CRC Press.

Kosmidis, V.D., Perkins, J.D., and Pistikopoulos, E.N. (2005). A mixed integer optimization for the well scheduling problem on petroleum fields. *Computers and Chemical Engineering*, 29, 1523–1541.

Naus, M., Dolle, N., and Jansen, J. (2006). Optimization of commingled production using infinitely variable inflow control valves. *SPE Production and Operations*, 21, 293–301.

Nygreen, B., Christiansen, M., Haugen, K., Bjorkvoll, T., and Kristiansend (1998). Modeling norwegian petroleum production and transportation. *Annals of Operations Research*, 82, 251 267.

Saputelli, L., Mochizuki, S., Hutchins, L., Cramer, R., Anderson, M., and Muller, J. (2003). Promoting real-time optimization of hydrocarbon production systems. *SPE 83978*.

Selot, A., Kuok, L.K., Robinson, M., Mason, T.L., and Barton, P.I. (2007). A model for short-term supply chain management of a lng production system. *IPTC, Dubai, U.A.E.*

Vanderbeck, F. (2006). Implementing mixed integer column generation. *Column Generation, chapter 12, Springer*, 331–358.

Wang, P. (2003). Development and applications of production optimization techniques for petroleum fields. *Ph. D. Dissertation, Standford University*.

Williams, H.P. (2005). *Model Building in Mathematical Programming*. Wiley.

Implementation and Validation of a Closed Loop Performance Monitoring System

Claudio Scali*(*), Marco Farnesi(**), Raffaello Loffredo(**), Damiano Bombardieri (***)

(*) Chemical Process Control Laboratory (CPLab), Department of Chemical Engineering (DICCISM), University of Pisa, Via Diotisalvi n.2, 56126, Pisa (Italy), e-mail: scali@ing.unipi.it
(**) ENI, Refining & Marketing, Raffineria di Livorno, Via Aurelia n.7, 57017, Livorno (Italy),
(***) Sysdat Informatica s.r.l.; Via Meucci 22 - Pisa (Italy)

Abstract: Main features of a performance monitoring systems operating on loops of refinery plants are illustrated together with examples of application and achieved results. The system analyses data recorded by the DCS during routine operation and originates verdicts about performance of control loops; also indications of causes of low performance and different strategies to adopt are given (retuning, valve maintenance, upstream actions). The system architecture is firstly illustrated, with characteristics of modules which accomplish different tasks of data acquisition and transfer, system configuration and priority assignment, performance analysis and verdicts emission, database query and operator support. Examples of field validation are then presented, with illustration of loop performance before and after actions suggested by the monitoring system. A synthesis of main techniques adopted in the system is finally presented.

Keywords: Closed Loop Performance Monitoring, Valve Diagnostic, Identification and Retuning.

1. INTRODUCTION

Closed Loop Performance Monitoring (CLPM) is widely recognized as a primary need in the process industry, as product quality, energy saving, waste minimization depend at a large extent on the efficiency of the control system.

The possibility of evaluating loop performance and being able to diagnose causes of deterioration brings to a direct improvement of plant performance both in the case of base control (SISO PID loops) and in the case of advanced control (MIMO, Optimization), which necessarily relies on good performance of low hierarchy control loops. Different causes of scarce performance, as incorrect tuning of controllers, anomalies and failures of sensors, presence of friction in actuators, external perturbations, should be detected and the right actions to perform should be suggested by the CLPM system.

This is a field where academic research and industrial applications should go ahead in tight contact in order to focus on real problems and to find solutions which are user friendly to be promptly accepted by plant operators. In large scale industrial processes, involving thousands of variables and hundreds of control loops, a monitoring system needs to operate automatically, leaving only key decisions to the operators. Also, results of performance analysis accomplished by the CLPM system should be presented in a very efficient manner to avoid to be seen as an additional work to be performed for plant supervision.

Anyway, even though a completely unattended system can be seen as the optimal solution, the right degree of interaction with plant personnel is of crucial important in the stage of parameter calibration and field validation. Also, the full potential of a CLPM system is fully exploited when the operator has at least a minimum knowledge of performance monitoring issues.

For these reasons, the final system architecture should take into account specific needs and requirements, as well as the skill of the user.

The evolution of academic research can be followed in the abundant literature produced in the last decade: two review papers (Qin, 1998; Thornhill and Horch, 2007) can give a flavour about. Certainly among more important open issues must be considered: the definition of significant performance indices, the development of simple and reliable techniques for automatic detection of causes, the diagnosis of root causes of perturbations in large scale plants. A very active research area is concentrated on methods for automatic diagnosis of valve stiction and many new techniques have been proposed in the last few years. A first comparison can be found in Horch (2006), while Jelali and Scali (2009) compare 11 recently proposed techniques on a benchmark of 93 industrial loops.

The paper illustrates architecture and application results of a Closed Loop Monitoring System recently developed and implemented in an Italian refinery and has the following structure: section 2 presents the overall system, with logic and interaction of different modules performing data acquisition, data transfer and verdicts archiving; section 3 illustrates main features of the system which accomplishes performance analysis and diagnostics; section 4 shows basic

features of the Data Base and the query system; section 5 presents some examples of field validation and gives examples of achieved improvements; finally conclusions and next steps are reported in section 6. An appendix will add some details about the techniques adopted for performance analysis.

2. THE SYSTEM ARCHITECTURE

A synthetic picture of the system architecture is depicted in Figure 1, where different modules, their interconnection and physical location are indicated.

The User Module (MU) starts the whole procedure by sending a message to the module of scheduling (MS) about the sequence of plants (and loops) to be analysed (the procedure is repeated periodically). In addition, it allows to see the state of advancement of operations and to send queries to the database (DB) for a synthesis of performance analysis. The user module also permits the configuration of the loops which is the very first step of the performance monitoring process. Loops configuration consists in the assignment of loop name, DCS address, loop info (for instance: single loop or cascade), priorities and constraints of the acquisition. More important loops can have higher frequency of acquisition, cascade loops are acquired simultaneously, loops of the same process unit are analysed in the same data acquisition run.

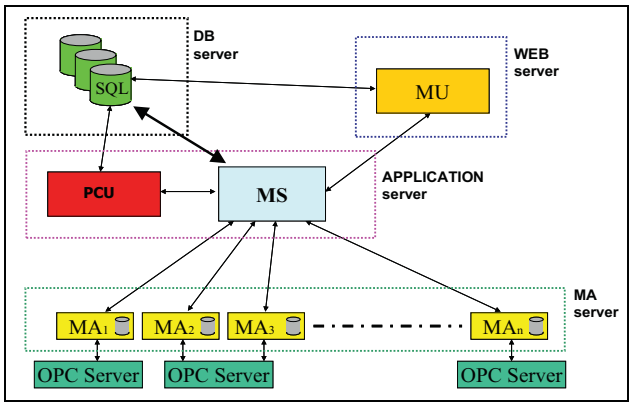


Fig.1: The system architecture

The Scheduling Module (MS), once activated by MU, sends a command to acquisition modules (MAi) which perform physically the acquisition of data from the DCS. For each loop, specific information are transferred to the Data Base (DB) trough MS, such as: loop tag name, controller settings, ranges of controlled variable (PV) and controller output (OP), saturation limits, loop hierarchy (e.g. master/slave of a cascade loop, loops under advanced control), Also information about default, minimum and maximum values for the duration of acquisition and sampling time (ts) are exchanged (default for ts= 10 seconds). Once acquisition is terminated, MS receives from MAi, data files which are sent to the DB input section. It activates the performance analysis accomplished sequentially by the PCU (Plant Check Up) module; finally, verdicts about loop status generated by PCU are transferred to the output section of the DB.

Acquisition Modules (MAi) interact with DCS, from which receive data and updated loop parameters at each sampling time; they act in parallel (up to a maximum number of 7 on a single server) and sequentially on scheduled loops, following priority and constraints indicated by MS. During the acquisition, the quality of each single datum and the change of status (man/auto, cascade open/closed) is checked and a flag is activated. In addition a first analysis is performed locally: mainly, the duration may be increased from the default (2 hours) to the maximum value (8 hours) in order to get a significant number of cycles in the case of very slow oscillating loops.

More about the PCU and DB modules are reported in next sections.

3. THE PCU MODULE

The PCU module contains the intelligence of the performance monitoring systems: it analyses each loop sequentially, interacting with the MS and with the DB from which receives raw data and to which send verdicts. A schematic representation is reported in Figure 2, where main steps and a simplified logical flow of data analysis modules are indicated. Further details about basic techniques adopted inside modules are given in the appendix.

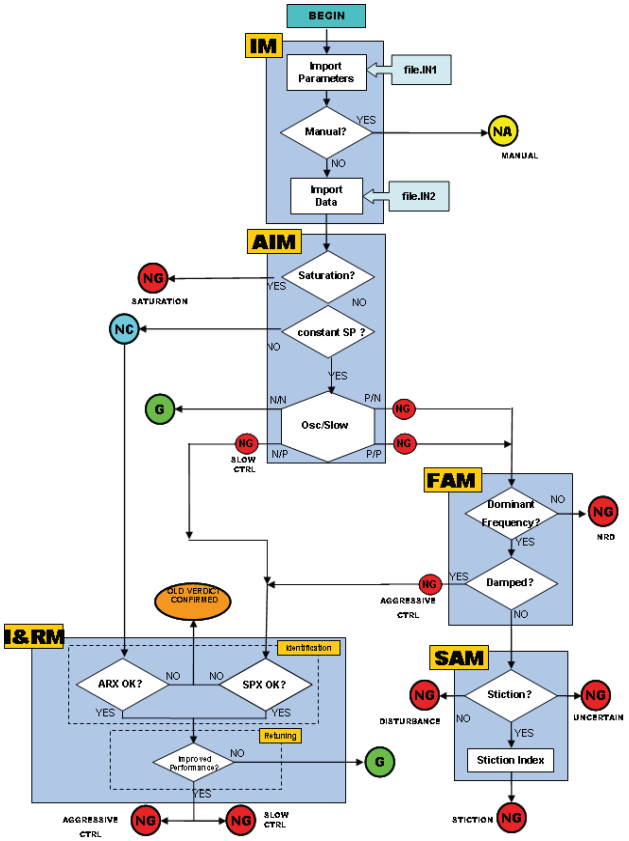


Fig.2: Schematic representation of the PCU module

IM: The Initialization Module imports parameters values from file IN1 and performs a first check about loop status; if the quality of data is not good, or a change of configuration is

detected, or the valve is operating manually (info contained in flags activated by MAi), the analysis stops. In this case, the loop receives a (definitive) label (NA: Not Analyzed) and the analysis is aborted. Otherwise, recorded data are imported from the IN2 file and the performance analysis begins.

AIM: The Anomaly Identification Module performs a first assignment of performance with verdicts: as G (*Good*), NG (*Not Good*). Loops subject to excessive set point changes (as amplitude or frequency) are temporary labelled as NC (*Not Classified*) and send to the identification module (I&RM). Valve saturation is checked first and, if detected, the label NG (and the cause) is definitive, without any further analysis (only duration is indicated). For loops not in saturation, after a data pre-treatment, tests to detect oscillating or sluggish loops are executed; these tests refer to the Hägglund approach (Hägglund, 1995, 1999), with suitable modifications of internal parameters, based on field calibration. In the case of both negative tests, the loop is classified as good performing and a definitive label G is assigned. Slow loops can only be caused by the controller: therefore they receive a NG label and are sent to the identification and Retuning Module (I&RM). Oscillating loops can be caused by aggressive tuning, external disturbance or valve stiction: for this reason, they are primarily sent to FAM, for a frequency analysis.

FAM: The Frequency Analysis Module has the scope of separating irregular oscillations from regular ones on the basis of a power spectrum which computes dominant frequencies; irregular loops are labelled NG, without any further enquiring about causes. Regular loops with decaying oscillations are sent to the I&R Module, otherwise (loops showing permanent oscillations) to the SAM for stiction/disturbance detection.

SAM: The Stiction Analysis Module analyzes data of NG oscillating loops and performs different tests to detect the presence of valve stiction. They mainly consist in the application of two techniques: the Relay based fitting of values of the controlled variable (PV) (*Rel*; Rossi and Scali, 2005) and the improved qualitative shape analysis (Ya^+; Scali and Ghelardoni, 2008). The two techniques are recalled in the appendix. Other techniques proposed for stiction diagnosis are also applied, when appropriate. Among them: the Cross-Correlation (*Cxy*; Horch, 1999), which is the simplest (and probably most widely used) test for a first discrimination between stiction and disturbance and the Bichoerence (*Bic*; Choudhury et al. 2005), which allows to put into evidence non linear characteristics of loop data. The appropriate application technique is automatically selected by the system, depending on type of loops, duration of acquisition, etc.. (for instance: Cxy is not used for Level Control, Ya^+ is reserved only to Flow Control loops). Final verdict takes into account indications coming from different techniques and from other auxiliary indices: to the exit loop, already tagged NG, the cause *Stiction* or *Disturbance* is assigned in the cases of strong evidence, otherwise the cause is *Uncertain*.

I&RM: The Identification & Retuning Module accomplishes process identification and, if successful, controller retuning and evaluation of performance improvements. It analyses loops tagged NG, owing to controller tuning (that is sluggish or too oscillating responses) and loops tagged NC. The two possibilities of constant and variable Set Point are treated differently, the second case being typical of secondary loops under cascade control. In the case of constant SP, recorded data represent a loop response under disturbance rejection: identification of process dynamics is carried out by means of a Simplex based search procedure (Rossi, 2007; Scali and Rossi, 2009). In the case of variable SP, recorded data represent a loop response under set point tracking: identification is performed by means of an ARX algorithm (Ljung, 1999). In both cases, if model identification is successful, new tuning parameters are calculated according to different techniques, the achievable performance improvement is evaluated by means of suitable upgrading indices and new controller settings are proposed. Otherwise, in the case of impossible identification, the previous assigned verdict is confirmed, without any additional suggestion.

To conclude this synthetic illustration, after the performance analysis by means of the PCU module, every loop is classified as:

- NA (Not Analysed): Manual valve, invalid data acquisition, change of loop configuration;

- NC: (Not Classified): impossible identification and no preliminary verdict;

- G (Good Performing);

- NG (Not Good performing): with an indication of cause (*saturation*, *sluggish, too oscillating, stiction, external disturbance*), or without indication for the cases of irregular disturbances or uncertainty between stiction and disturbance in the SAM.

As distinctive features of the PCU performance monitoring system, the following can be pointed out:

1) it is open to the adoption of new diagnosis techniques; once the algorithm has been built and checked in simulation, it is tested on archived and recent plant data: improvements of reliability of issued verdict lead to updating of algorithms and performance indicators;

2) it has been designed to operate completely unattended and for this reason a verdict is assigned and causes indicated only in the case of strong evidence; false indications are carefully avoided, at the expense of conservative (too cautious) indications: in case of uncertainty, the verdict is postponed to next data acquisitions;

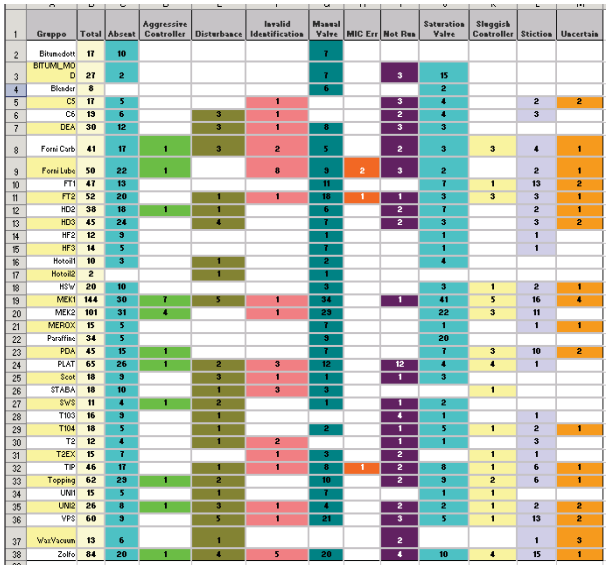
3) the calibration of values of key and auxiliary performance indicators is made on the basis of operator experience, in order to make verdicts as more homogeneous as possible with their practice.

4. THE DATA BASE ORGANISATION

The Data Base contains all information about each single loop: recorded data, controller parameters, loop configuration and diagnosis performed by PCU (verdicts). The possibility of a fast and efficient consultation from the operator is certainly one of main specifications to be achieved for the

success of the whole implementation. Therefore, operators suggestions and requirements are carefully taken into consideration in the design of the Data Base architecture. Some significant features incorporated in the DB management are illustrated in the sequel.

1) Analyzed loops and issued verdicts for a group of plants (or all of them) at a certain date, can be immediately summarized on the screen and this allows a first evaluation of loop status, that is the total number of good performing loops, causes of scarce performance, loops in manual, reasons for invalid acquisition etc.. (Figure 3).



	Gruppo	Total	Absent	Aggressive Controller	Disturbance	Invalid Identification	Manual Valve	MIC Err	Not Run	Saturation Valve	Sluggish Controller	Stiction	Uncertain
2	Bitumodott	17	10				7						
3	BITUMI_MOD	27	2				7		3	15			
4	Blender	8					6			2			
5	C5	17	5			1			3	4		2	2
6	C6	13	6		3	1			2	4		3	
7	DEA	30	12		3	1	8		3	3			
8	Forni Curb	41	17	1	3	2	5		2	3	3	4	1
9	Forni Lube	50	22	1		8	9	2	3	2		2	1
10	FT1	47	15				11			7	1	13	2
11	FT2	52	20		1	1	18	1	1	3	3	3	1
12	HD2	38	18	1	1		6		2	7		2	1
13	HD3	45	24		4		7		2	3		3	2
14	HF2	12	9				1			1		1	
15	HF3	14	5				7			1		1	
16	Hotoil1	10	3		1		2			4			
17	Hotoil2	2			1		1						
18	HSV	20	10				3			3	1	2	1
19	MEK1	144	30	7	5	1	34		1	41	5	16	4
20	MEK2	101	31	4		1	23			22	3	11	
21	MEROX	15	5				7			1		1	1
22	Paraffine	34	5				3			26			
23	PDA	45	15	1			7			7	3	10	2
24	PLAT	65	26	1	2	3	12		12	4	4	1	
25	Scot	18	3		3	1	1		1	3			
26	STABA	18	10		1	3	3				1		
27	SWS	11	4	1	2		1		1	2			
28	T103	16	3		1				4	1		1	
29	T104	18	5		1		2		1	5	1	2	1
30	T2	12	4		1	2			1	1		3	
31	T2EX	15	7			1	3		2		1	1	
32	TIP	46	17		1	1	8	1	2	8	1	6	1
33	Topping	62	29	1	2		10		2	9	2	6	1
34	UNI1	15	5		1		7			1	1		
35	UNI2	26	8	1	3	1	4		2	2	1	2	2
36	VPS	60	9		5	1	21		3	5	1	13	2
37	WaxVacuum	13	6		1				2			1	3
38	Zolfo	84	20	1	4	5	20		4	10	4	15	1

Fig.3: Global picture of all plants monitored at a certain date, with indication of loop status

2) The same type of visualization can be produced for a group of plants (or a single plant) for acquisitions repeated in a desired interval of time, thus allowing a first indication about the trend of loop performance (Figure 4).

Fig.4: Global performance of a single plant for repeated acquisition at different dates

3) All plant loops at a desired date can also be visualized with individual verdicts (Figure 5).

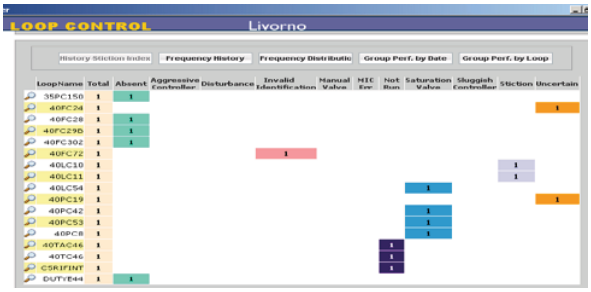


Fig.5: Single plant loop performance at a desired date

4) Single loop performance can be easily investigated, by means of plots and significant performance indices. In the case of successful identification of a loop with incorrect tuning, the trends of SP, PV, OP variables show possible improvements and required control effort with best tuned PI(D) controllers (Figure 6). On this basis (together with values of upgrading indices which allow a quantification of improvements (shown in a separate page), the operator can take a decision about the opportunity of a controller retuning.

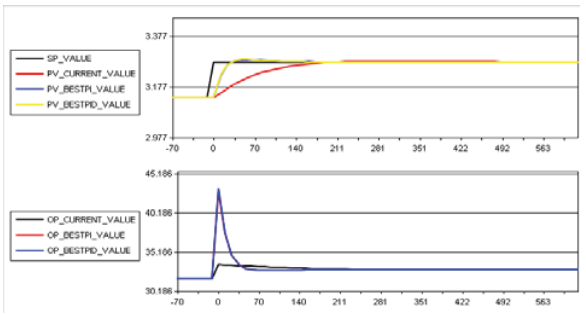


Fig.6: Screenshot for a loop with incorrect tuning

5) In the case of a loop affected by valve stiction, the trends of loop variables, the value of a Stiction Quantification Index (shown in a separate page) allow to evaluate at a glance the situation of the loop (Figure 7). This can be further confirmed by watching the PV(OP) movie (see Figure 11).

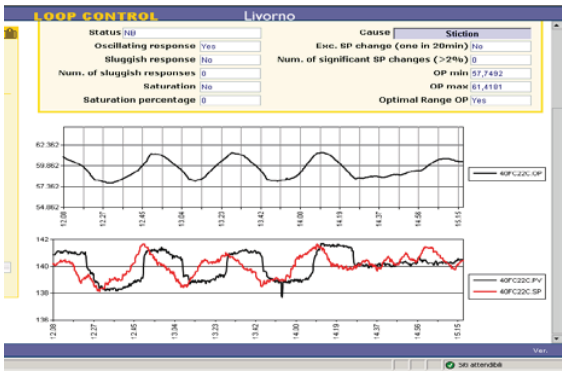


Fig.7: Screenshot for a case of valve stiction: time trends of OP, SP, PV.

7) The loops history can be easily tracked: in the case of a confirmation of issued verdicts, indications for proper actions on the loop can be decided (tuning, valve maintenance, upstream action); an example for a case of

confirmed verdict with increasing stiction in the last four data acquisition is reported in Figure 8.

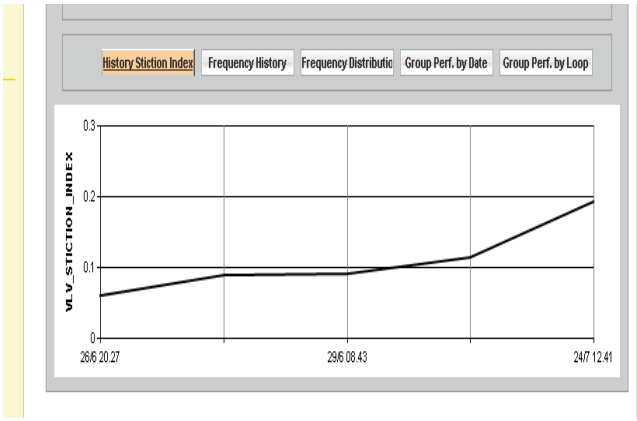


Fig. 8: Loop history for a case valve stiction

4) Many other features allow easy access to more information; for instance: frequency analysis of different oscillating loops and the comparison of dominant frequencies allows to focus on loops possibly affected by the same root cause of oscillations. Other auxiliary performance indexes are evaluated and a large variety of reports about loops statistics can be obtained right way.

5. FIELD VALIDATION

Field validation is the key step of the monitoring system implementation. As first, it allows a direct confirmation of verdicts emitted after loop data analysis, while all indications illustrated in previous section must be considered mere "predictions", i.e. based on identified models, techniques results and values of performance indices. As second point, this is the step where the operator can give indications for the final calibration of threshold values of performance indices and get confidence about the reliability of verdicts issued by the system. Few illustrative examples are reported in the sequel; more than 600 loop acquisitions were checked during the monitoring system validation.

1) Loop xxFC01 (PI control, Constant Set Point). The verdict from AIM and I&RM modules is NG, indicating as cause: sluggish controller. The identification is successful and the old settings (K_c=*1*, T_i=*0.65*), should be changed to new ones: K_c=*0.49*, T_i=*0.13*. An increase of integral action is then suggested; the upgrade index based on the model (see appendix) is: Φ=*0.476*. In this case (single FB loop), it is possible to check directly the predicted improvements: a moderate increase to SP has been given by the operator with old settings, followed by a decrease with new settings. The improvement is evident from Figure 9 (the small amplitude and high frequency oscillation represents an unidentified perturbation present in the plant) and the upgrade index evaluated from plant data is Φ'=*0.573*. This application suggested to reduce the threshold for the index Φ^0 to *0.40* (initially Φ^0 was assumed equal to *0.50*).

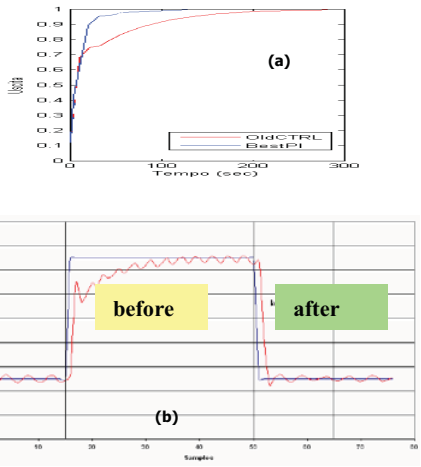


Fig.9: Response to a SP change: model prediction (a), field validation before and after retuning (b)

Loop xxFC02 (PI control, Variable Set Point). Also in this case the verdict from AIM and I&RM modules is NG, owing to sluggish controller; old settings: K_c=*0.8*, T_i=*0.7*; new settings: K_c=*2.6*, T_i=*0.72*. A strong increase of the proportional constant is proposed in this case, while the integral time constant does not change much; the upgrade index based on the model is now Φ=*0.487*. Being a secondary loop under cascade control, in this case it is not possible to give arbitrary set point changes during plant exercise. Moreover, being suggested a large increase of the gain K_c, the operator applied a gradual increase of gain: *0.8, 1.2, 1.6, 2.6*. The corresponding improvement of response in set-point following is evident from Figure 10; the upgrade index evaluated from plant data for increasing value of K_c, is now evaluated by the index (see appendix) *IQI = 0.038, 0.78, 0.85, 0.94*, to confirm the performance improvement.

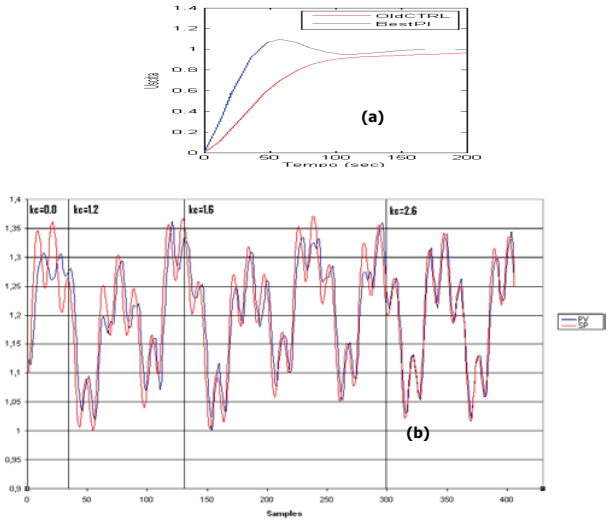


Fig.10: Response to a SP change: model prediction (a), field validation for increasing values of the gain K_c (b)

Loop xxFC08 (PI control, slow varying Set Point). This loop has been repeatedly indicated as affected by stiction in

several analysis. The values of the Stiction Quantification Index increased from 0.07 to 0.195 in about one month. The presence of stiction is clearly recognizable by the PV and OP shapes (close to square waves and triangles, respectively in Figure 11a). Moreover, the plot of PV(OP), which can be seen as a movie on the screen, shows evident stiction characteristics (Figure 11b); because in this case (FC loop), the controlled variable PV is proportional to the valve opening MV. Valve maintenance brought to an improvement of performance and a sharp decrease of the stiction index (Figure 12).

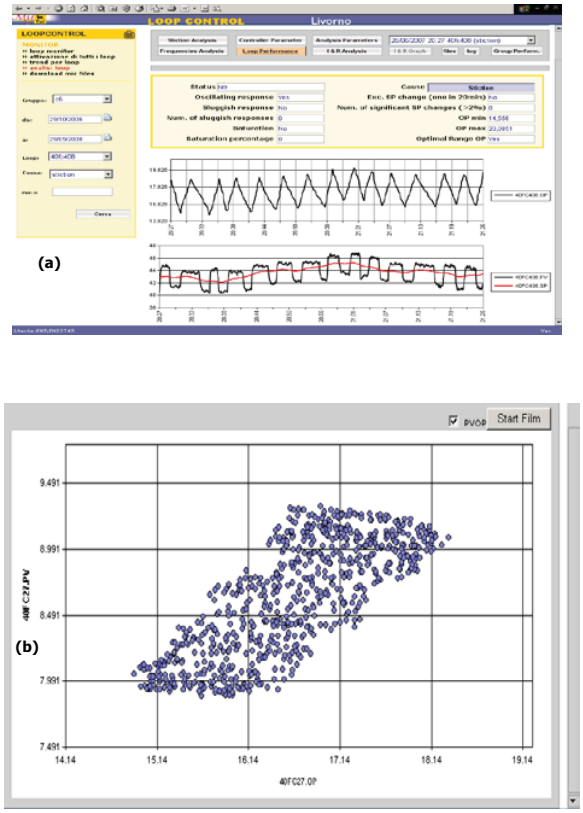


Fig.11: Validation of a loop affected by valve stiction: (a) SP, PV, OP trends; b) PV(OP) movie (FC loop)

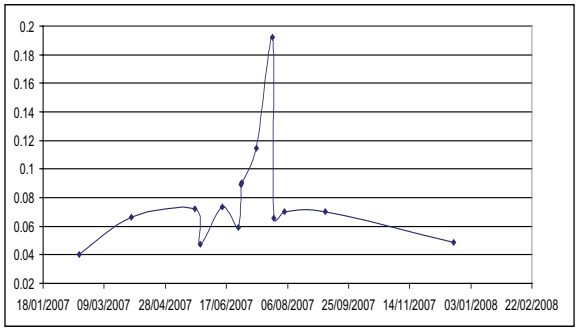


Fig.12: Time trend of the Stiction Quantification Index before and after valve maintenance

6. CONCLUSIONS

The Closed Loop Performance Monitoring system described in the paper has been developed and built with tight cooperation between university and plant personnel. The role of plant operators has been crucial for the success of the implementation, mainly in the calibration of threshold values for key performance indices and in the definition of specifications of the Data Base query system for an efficient analysis of loop performance. This fact has brought to a final version of the system "customized" on user requirements.

The design specifications for a "completely unattended" system forced to the adoption of conservative (default) values for key performance indices and, as a consequence, verdicts are emitted only in case of strong evidence, leaving a certain number of uncertain/unclassified cases. In the stage of assistance to the project, loop analysis was repeated for these cases by changing threshold values, allowing to explain many of them, thus confirming the advantages of a deeper involvement of plant personnel.

The flexibility of the system is an important feature, allowing different levels of interaction with the operator: from the lowest (analysis of periodical performance reports issued by the system) to the highest (actions on loops labelled as poorly performing). The inspection of these loops allows to focus on anomalous situations, both in the case of complete verdicts (cause indicated), and in the case of incomplete diagnosis for a deeper analysis based on process knowledge.

After implementation on a selected plant (about twenty loops), followed by field validation, it has been applied on about fifteen plants, featuring several hundreds loops. A further validation is carried on with the scope of a systematic evaluation of obtained benefits in the perspective of implementation on other group refineries.

Appendix A. ADDITIONAL DETAILS ABOUT PCU

The PCU (Plant Check Up) is the engine of the performance monitoring system and accomplishes an analysis of loop data in order to evaluate performance and to diagnose causes. Some more details are given here about techniques for stiction diagnosis, identification retuning and performance improvement evaluation. Necessarily, only a synthetic illustration is reported here; full details can be found in the references.

A.1 Stiction Diagnosis

Two techniques are mainly used for this scope: the Relay fitting of PV values (*Rel*; Rossi and Scali, 2005) and the improved qualitative shape analysis (Ya^{+}; Scali and Ghelardoni, 2008).

The *Rel* technique consists in the fitting of significant half cycles of the recorded oscillation by means of three different models: a sine wave, a triangular wave and the output response of a first order plus time delay under relay control. The last one is able to approximate the square waves shapes generated by the presence of stiction and modified by the

process dynamics (Figure 13). Relay and triangular shapes are associated with the presence of stiction, while a sinusoidal shape with the presence of external perturbations. By comparing the error between real and fitted data, an evaluation of the accuracy of approximation and then an indication of the underlying phenomenon can be obtained. Once approximations have been performed, a Stiction Identification Index (S_I) can be defined. Being E_S the minimum square error obtained by the sinusoidal approximation and E_{RT} the one obtained by the better approximation between the relay and the triangular waves, S_I is defined as:

$$S_I = (E_S - E_{RT})/(E_S + E_{RT}) \qquad (1)$$

Here E_S, E_{RT} indicate average values of error over the number of examined cycles. S_I varies in the range [-1; +1]: negative values indicate a better approximation by means of sinusoids, positive values by means of relay or triangular approximations. Values close to zero indicate that the two approximations have similar errors and the procedure gives an uncertain answer: the uncertainty region is defined by $|S_I|<0.21$.

The technique presents some analogies with the Curve Fitting Method proposed by He et al (2007): in this case, assuming that stiction is associated to a square wave shape in MV, a triangular wave shape is looked for as distinctive feature of stiction after the first integrator element of the loop. This means in OP signal (for self regulating processes – no integrators) or in the PV signal (for integrating processes).

The relay method always analyses the PV signal and uses the relay shape as additional primitive. The global fitting procedure is more complex and time consuming, but in all cases the elaboration time is absolutely negligible compared with the duration of data acquisition. Finally, the method can also put into evidence the presence of asymmetric stiction, by comparing S_I values on positive and negative half cycles.

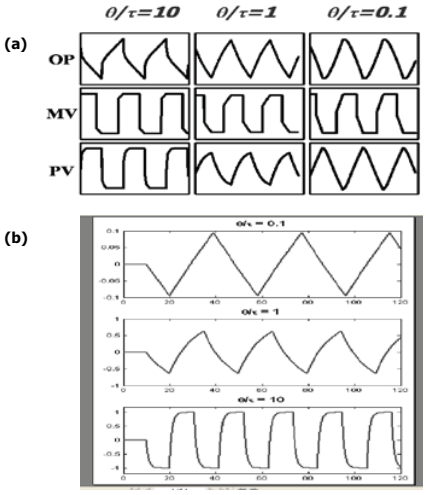


Fig.13: (a) Wave shape in a loop affected by stiction as modified by process dynamics (ratio θ/τ) for a FOPTD process. (b) Different wave shapes generated by a relay feedback controller on a FOPTD process by varying θ/τ

The Ya^+ technique is an extension of the technique originally proposed by Yamashita (2006), which is based on the analysis of trends MV(OP), that is valve output as function of the control action. Its applicability would seem low, because usually only PV and OP are recorded on industrial DCS; in the case of flow control (FC), the controlled variable can be considered known, being proportional to the valve opening. As FC loops are a large majority of base control loops (about 2/3 in the application presented here), the applicability of this technique is large. It is much larger for newly designed plants (for instance power plants with redundant instrumentation) and it will increase in a next future with the diffusion of field bus communication systems and related available information.

In the presence of stiction, the trend changes from linear to a typical parallelogram shape (Fig.14): the horizontal part indicates that the valve opening does not change for increasing controller output.

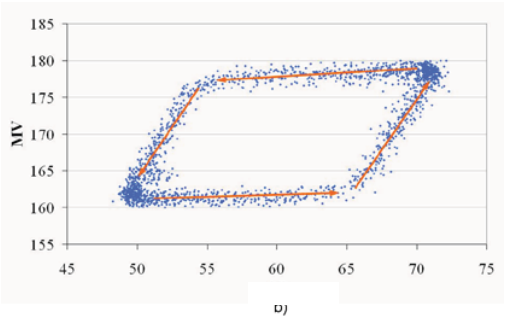


Fig. 14: Valve position (MV) as a function of the controller output (OP) in the presence of stiction (industrial data)

Following Yamashita (2006), the pattern can be approximated by means of three simple symbols: increasing (I), decreasing (D), and steady (S). Possible combinations of symbols for the stiction pattern reported above are represented in Figure 15, as: ID, IS, II; SD, SS, SI; DD, DS, DI.

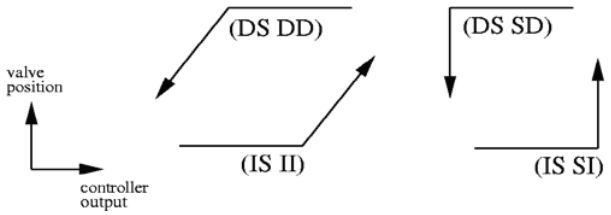


Fig. 15: Qualitative shapes observed in sticky valves.

By combining the symbols for OP and MV signals, a representation of the development in an (OP, MV) plot over time can be obtained. Based on these considerations and counting the duration of time sequences, a stiction index ρ can be defined; values of $\rho > \rho^{\circ}=0.25$ (threshold value for a random signal) are indication of possible stiction in the valve. These considerations have been extended to include different stiction patterns observed in industrial data, for instance the one reported in Fig.16.

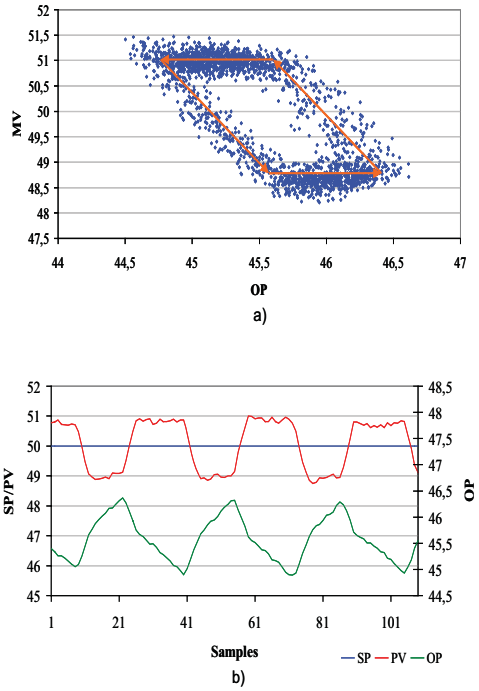


Fig. 16: a) A different MV(OP) pattern observed in industrial data, in the presence of evident stiction (b)

Other patterns are possible depending on valve type (direct/inverse action) and on DCS configuration, as reported in Figure 17.

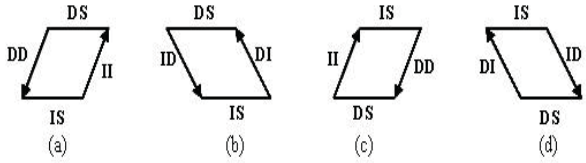


Fig. 17: Additional stiction patterns

They can be explained by the presence of a (even small) delay between OP and MV is present, caused for instance by the combined action of different factors such as: backslash phenomena, valve positioner dynamics, signal quantizer and so on. They can be reproduced by simulation by means of a widely used stiction model, with suitable modifications (Choudury et al, 2005).

Different stiction indices have been defined to be able to identify their presence in industrial data, namely: ρ_B, ρ_C, ρ_D, (in addition to a $\rho_A=\rho$), accounting for the appropriate coupled sequences of I, S, D primitives (further details in Scali and Ghelardoni, 2008).
For a set of 52 data acquisitions, 11 additional loops were indicated as sticky, according to the new index ρ_B, while would not be indicated by ρ_A, as summarized in table 1 (threshold value is 0.25). During the implementation and field validation of this project, only A and B stiction patterns were encountered, owing to some practical constraints adopted in the DCS configuration; C and D patterns may be found in the most general case.

Table 1: Details for the additional 11 sticky loops

Loop #	ρ_A	ρ_B
xxFC1	0.2459	0.4146
xxFC2	0.1941	0.2648
xxFC3	0.2238	0.3444
xxFC4	0.2303	0.2817
xxFC5	0.1889	0.2961
xxFC6	0.2071	0.4882
xxFC7	0.1352	0.2596
yyFC1	0.1200	0.3621
yyFC2	0.1797	0.3712
yyFC3	0.1614	0.3407

A.2 Identification

The Identification Module receives form the AIM module loops with constant SP labelled as NG (No Good) caused by improper tuning and loops labelled as NC (Not Classified) with variable SP.

In the case of constant SP, the recorded lops dynamics refer to a poor performing response caused by the presence of an external perturbation (Figure 18). A Simplex based search technique has been adopted for the solution of this problem, with some modifications to the original algorithm (Nelder and Mead, 1964), in order to overcome the problem of getting stuck in local minima and of managing the presence of constraints. Further details are reported in Rossi (2007) and Rossi and Scali (2009).

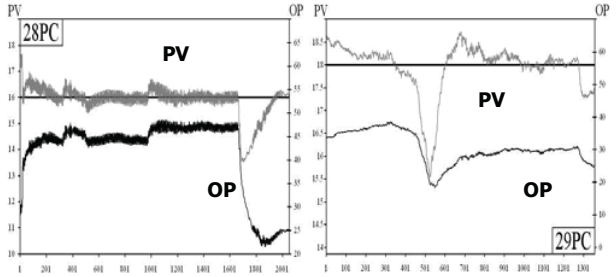


Fig. 18: Poor loop responses caused by external perturbation

Both process and disturbance dynamics are modelled as second order plus time delay systems, with parameters K, K_d (gains); θ, θ_d (delays); τ, τ_d (time constants); ξ, ξ_d (damping factors):

$$P(s)=\frac{K\cdot e^{-\theta s}}{\tau^2 s^2+2\,\xi\,\tau\,s+1};\; P_d(s)=\frac{K_d\cdot e^{-\theta_d s}}{\tau_d{}^2 s^2+2\,\xi_d\,\tau_d\,s+1} \qquad (2)$$

The identification problem can be stated as the minimization over the model parameters vector V, of MSE between recorded and computed values (N is the number of samples):

$$\min_V (MSE);\quad MSE=\sum_{i=1}^{N}(Y_i-\hat{Y}_i)^2 \qquad (3)$$

In the case of variable SP, an ARX process model is identified (Ljung, 1999); in discrete form:

$$y_k + a_1 y_{k-1} + \cdots + a_n y_{k-n} = b_1 u_{k-L-1} + \cdots + b_m u_{k-L-m} + e_k \quad (3)$$

where: y is the output (PV), u is the input (OP), L is the time delay, n and m are model order. From past values of y and u, it is possible to define the output predictor as:

$$\begin{cases} \hat{y}_k = \varphi_k^T \theta, \\ \varphi_k = [-y_{k-1} \quad \cdots \quad -y_{k-n} \quad u_{k-L-1} \quad \cdots \quad u_{k-L-m}]^T \\ \theta = [a_1 \quad \cdots \quad a_n \quad b_1 \quad \cdots \quad b_m]^T \end{cases} \quad (4)$$

Once a time window has been fixed of length equal to N sampling times, (details are given below), a suitable quadratic function of the error between predicted and recorded values:

$$V_N(\theta) = \frac{1}{N}\sum_{k=1}^{N}(y_k - \hat{y}_k)^2 = \frac{1}{N}\sum_{k=1}^{N}(y_k{}^2 + \theta^T \varphi_k \varphi_k^T \theta - 2\theta^T \varphi_k y_k) \quad (5)$$

and, by minimizing V_N with respect to model parameters:

$$\theta^* = \left[\sum_{k=1}^{N} \varphi_k \varphi_k^T\right]^{-1} \left(\sum_{k=1}^{N} \varphi_k y_k\right) \quad (6)$$

The value of the delay L should be known; this limitation can be overcome by repeating the computation of θ^* for different values of L (from *0* to L_{max}) and choosing the delay θ^* corresponding to minimum values of V_N. The initialization of the predictor requires the knowledge of data for $N_0=max(n,m+ L_{max})$ sampling times, before $k=1$.

Several criteria can be defined to evaluate the accuracy of identification. Here, a closed loop index has been adopted, as the scope of identification is modelling for control purpose. Given the SP sequence in the examined window, values of the output variable vector $y'_1, ..y'_N$, for the actual controller and the identified model are computed, originating the index:

$$EV_{CL} = 1 - \frac{\sum_{k=1}^{N}(y_k - y'_k)^2}{\sum_{k=1}^{N}(y_k - y_{mean})^2} \quad (7)$$

where: y_{mean} represents average value for the output. EV_{CL} represents an explained variance, with values less than 1 (and generally not less than 0).

The application of the procedure is different according to the number of SP changes, for instance primary loops or loops under cascade / advanced control.

In the case of only one SP change, only one time window is selected and the procedure is applied as described above. The starting point is fixed n_0 sampling times before the time of SP change, while the final point is taken when the response has settled within a 5% of the output value. The identification is considered successful if $EV_{CL} \geq 0.80$.

For cases of variable SP, time windows of about 20 minutes for FC loops (about one hour for other type of loops) are chosen and the identification procedure is applied on each window. The step response of the two models having larger value of EV_{CL} are compared, as:

$$MD = 1 - \frac{\frac{1}{M}\sum_{k=1}^{M}(y_k^{SR1} - y_k^{SR2})^2}{\sum_{k=1}^{M}\left(\frac{y_k^{SR1} + y_k^{SR2}}{2}\right)^2} \quad (8)$$

where: y_k^{SR1}, y_k^{SR2} are *k-th* step response coefficient of model 1 and 2. The identification is considered successful for $MD \geq 0.95$ (and $EV_{CL} \geq 0.80$). Threshold values have been assumed after intensive simulations and applications on loop data. Referring to Figure 19 and Table 2, acquired data are divided in 8 time windows, identification is heavily wrong for windows 2 and 4 ($EV_{CL} < 0$), is not considered reliable enough for windows 1,3,5,6 ($EV_{CL} < 0.80$), is considered successful for windows 7 and 8 ($EV_{CL} \geq 0.80$); for these two windows $MD = \geq 0.986$ and then the identified model is accepted.

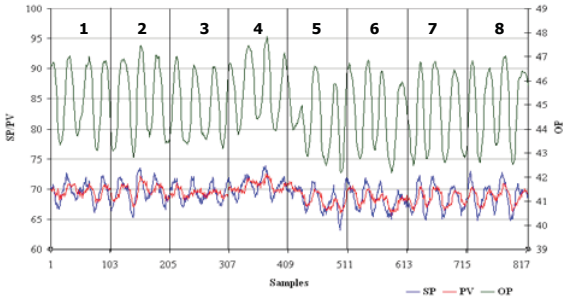


Fig. 19: Variable SP loop and time windows divisions

Table 2: Values of EV_{CL} in the 8 time windows

t.w.	1	2	3	4	5	6	7	8
EV_{CL}	*0.51*	*-40*	*0.66*	*-29*	*0.58*	*0.74*	*0.89*	*0.80*

Identification may be not successful for several reasons, for instance: non linearity of real process or not optimal choice of ARX model order (*n, m*). Nevertheless, main causes of failure are to be found in the presence of valve stiction or external disturbances.

In the first case, a failure is "desirable" (a linear model being not reliable in this case) to avoid the adoption of an incorrect model and a wrong suggestion about cause (retuning instead of stiction).

In the second, it may be possible to find few time windows not heavily affected by disturbances: this is the logical behind the choice of comparing models identified in two different time windows and requiring large values of *MD* and EV_{CL}; more details in Mervi (2007).

A.3 Upgrading Indices

Once the identification has been successfully carried out and a process model is available, the optimal tuning is evaluated according to different available techniques, selected at the configuration stage.

The performance improvement predicted on the basis of the identified model, is evaluated by means of an upgrading index Φ:

$$\Phi = \frac{IAE_{Act} - IAE_{Best}}{IAE_{Act} - IAE_{Min}} \quad (9)$$

where: IAE is the Integral of Absolute Error of the response for the actual reguator (*Act*), for the best controller having PI/PID structure (*Best*) and for the optimal one for the identified model (*Min*). For $\Phi \rightarrow 0$, the proposed controller is closet to the optimal one; for any $\Phi > 0$ there are improvements, but a threshold has been assumed to implement the new retuning (the proposed threshold $\Phi^{\circ} = 0.50$, has been decreased to *0.40*, after field validation).

Other indices allow to evaluate the real performance improvements on the plant, before and after retuning, in the two cases of primary loops (with rare SP changes, mainly step-wise) and secondary loops (with frequent SP changes, imposed by the primary loop acting on them).

For primary loops a new index Φ' is defined, having the same expression (9), with controllers tagged as *Act* and *Best* substituted by *Old* and *New* (to indicate before and after retuning) (Figure 20a).

For secondary loops, the IQI (Improvement Quantification Index) is defined, to evaluate the error between recorded SP and PV values before and after retuning (Figure 20b):

$$IQI = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\left(SP_i - PV_i\right)^2}{\frac{1}{N}\sum_{i=1}^{N}\left(SP_i - SP_{medio}\right)^2} \quad (10)$$

where: N: is the number of sampling times where the tuning is maintained as constant, SP_i: , PV_i , i-th value of SP and PV, and, SP_{ave}, is the average SP value in the time range where tuning parameters are left constant.

Values of IQI close to 1 indicate perfect control, while small or negative values indicate scarce performance.

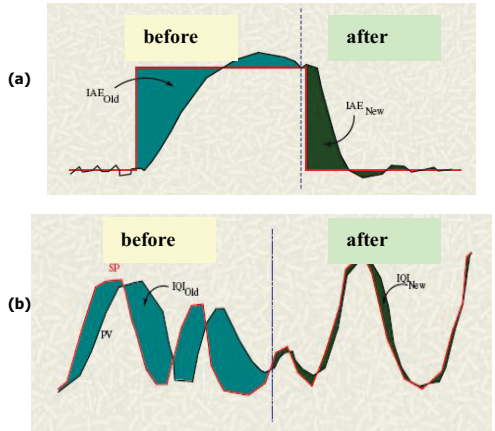


Fig. 19: Representation of the upgrading indices for: constant (a) and variable SP loops (b)

REFERENCES

Choudhury M.A.A.S, Shah S.L., and Thornhill N.F. (2004): "Diagnosis of Poor Control Loop Performance using High Order Statistics", *Automatica*, **40**(10), 1719-1728.

Choudhury M.A.A.S., Shah S.L., Thornhill N.F. (2005): "Modelling Valve Stiction", *Control Engineering Practice*, Vol. 13, pp. 641-658.

Hägglund T. (1995): "A Control Loop Performance Monitor", *Control Eng. Practice*, **3**, 1543-1551.

Hägglund T. (1999): "Automatic Detection of Sluggish Control Loops", *Control Eng. Practice*, **7**, 1505-1511.

He Q. P., Wang J., Pottman M., Qin S. J. (2007): "A Curve Fitting Method for Detecting Stiction in Oscillating Control Loops"; Ind. Eng. Chem. Res., Vol.46 (13), 4549-4560.

Horch A. (1999): "A Simple Method for Detection of Stiction in Control Valves", *Control Eng. Practice*, **7**, 1221-1231.

Horch A. (2006): "Benchmarking Control Loops with Oscillation and Stiction"; In: *Process Control Performance Assessment*, Ed.s: A. Ordys, D. Uduehi and M. A. Johnson; Springer, London –UK.

Jelali M, Scali C. (2009): "Comparative Study of Valve Stiction Detection Methods"; In: *Detection and Diagnosis of Stiction in Control Loops*, Ed.s: M. Jelali and B. Huang; Springer, London –UK.

Ljung L. (1999): "*System Identification"*, Prentice Hall, Upper Saddle River, N.J. (USA), (1999).

Mervi A (2007): "MS Thesis" (in Italian), Chem. Eng. Dept. – University of Pisa (I).

Nelder G.A., Mead R. (1964): "A Simplex Method for Function Minimization", The Computer Journal, Vol.7, No. 4, pp. 308-313.

Qin S. J. (1998): "Control performance monitoring: a review and assessment", *Comp. & Chem. Engineerin*g, **23**, 173-186.

Rossi M., Scali C. (2005): "A Comparison of Techniques for Automatic Detection of Stiction: Simulation and Application to Industrial Data", *J. Proc. Control*, **15** (5), 505-514.

Rossi M. (2007): "Development of Techniques for Closed Loop Monitoring of Chemical Processes", *PhD Thesis*; Chem. Eng. Dept. – University of Pisa (I).

Scali C. and Rossi M. (2009): " Modeling Disturbance Dynamics to Improve Controller Performance in Industrial Loops**";** *Chemical Product and Process Modeling*: Vol. 4: Iss. 2, Article 8.

Scali C., Ghelardoni C. (2008): "An Improved Qualitative Shape Analysis Technique for Automatic Detection of Valve Stiction in Flow Control loops", *Control Engineering Practice*, 16, pp. 1501-1508.

Thornhill N. F., Horch A. (2007): "Advances and New Directions in Plant-Wide Disturbance Detection and Diagnosis"; *Control Engineering Practice*, Vol. 15(10), pp. 1196-1206.

Yamashita Y. (2006): "An Automatic Method for Detection of Valve Stiction in Process Control loops"; *Control Engineering Practice*, Vol. 14, pp. 503-510.

MPC: Current Practice and Challenges

Mark L. Darby*, Michael Harmse, Michael Nikolaou*****

**CMiD Solutions, Houston, TX*
USA (Tel: 713-477-7791; e-mail: darbymark@sbcglobal,net).
***IPCOSAptitude Ltd., Cambridge, United Kingdom (e-mail:Michael.Harmse@ipcosaptitude.com)*
**** Chemical and Biomolecular Engineering, Houston, TX (e-mail:nikolaou@uh.edu)*

Abstract: Linear Model Predictive Control (MPC) continues to be the technology of choice for constrained multivariable control applications in the process industry. Successful deployment of MPC requires "getting right" multiple aspects of the problem. This includes the design of the regulatory controls that receive setpoints from MPC, design of the multivariable controller(s) themselves, test design for model identification, model development, and dealing with nonlinearities. In the following, we highlight approaches and techniques that are successfully applied in practice and provide an overview of recent technological enhancements that are being made to MPC. While significant progress has been made in both the technology and practice, there are challenges with MPC, mostly related to the effort required to develop an application and to ensure adequate performance over time. Suggestions for addressing these issues are included as possible research directions.

Keywords: model predictive control, model-based control, constraints, control system design, modeling, process identification.

1. INTRODUCTION

Model predictive control (MPC) is a mature technology. It is the standard approach for implementing constrained, multivariable control in the process industries today. MPC provides an integrated solution for controlling interacting systems with complex dynamics and constraints. A key aspect of MPC is its ability to deal with degrees of freedom, that may arise when there are more or fewer inputs (manipulated variables) than outputs (controlled variables), or when zone limits for controlled variables are used, which is the typical situation in practice. Broadly defined, MPC refers to a control algorithm that explicitly incorporates a process model to predict the future response of the controlled plant. While the model may be linear or nonlinear, we consider linear MPC as it is used in the majority of industrial applications in the refining and petrochemical industries today (and increasing, in other industries). For these applications, the plant model is identified based on data generated from a dedicated plant test. Today, there are a number of technology vendors which provide MPC solutions, including software to facilitate the development of MPC applications and monitoring of the performance of these applications over time. The last 10-15 years has seen significant efforts by technology suppliers to improve the usability of MPC products.

While the "science" of MPC has advanced and the technology is now easier to apply, there is still an "art" aspect to the application of MPC that largely comes from experience. The success of an MPC application depends on the multiple technical decisions that are made by the control engineer in the course of an implementation. In addition, there are both technical and organizational issues that are critical to ensuring that MPC benefits are sustained in the longer term once an MPC is commissioned (Darby and Teeter, 2005). Based on our experience, we find that the success rate of MPC across the industry is uneven. Some companies are consistently successful in deploying MPC, whereas others are not. In the following, our main emphasis concerns the technical aspects of MPC that arise in the course of an implementation.

MPC is positioned above a regulatory control level as shown in Figure 1. The manipulated variables for the MPC are typically setpoints of PID controllers, executed in a distributed control system (DCS). The MPC may also directly manipulate valve position signals rather than, e.g., flow.

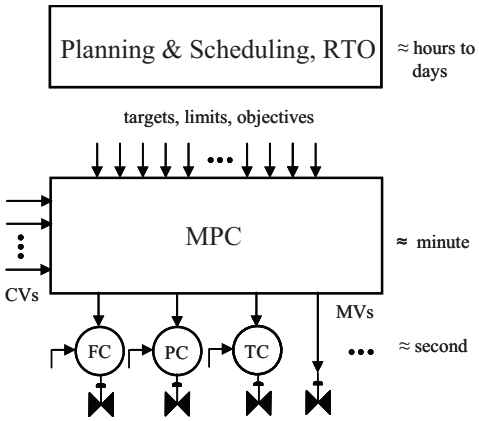


Figure 1. Control hierarchy

The DCS executes a at a higher sampling rate than the MPC, typically sub–second to multi-second sample time, compared to a 30 sec to 2 min execution period for MPC.

Certain targets and objectives for MPC come from higher level functions such as planning and scheduling, typically communicated to the operator in an open-loop fashion, or from a real-time optimizer, if present. Note that there it not necessarily a one-to one-translation of decisions from upper level functions to targets and limits in the MPC. Economic objectives and priorities may also be involved. Examples include gasoline vs. diesel objectives (winter vs. summer) in a refinery and the priority of feed stocks in an ethylene plant. In addition, there are day-today issues that may arise such as a late shipment, or a product tank becoming full.

Part of the challenge in implementing MPC is that the regulatory control layer is not a given (or should not be taken as a given). The design problem is really one of deciding on the best overall structure for the regulatory level *and* MPC, given the control objectives, expected constraints, at least qualitative knowledge of the expected disturbances, and robustness considerations. Similarly, the selection of the controlled variables for MPC is not one of simply deciding which subset of available measurements should be selected. It may be that available measurements are insufficient and additional sensors are needed. In addition, not all variables that need to be controlled may be available on a frequent-enough basis; therefore, we have the problem of inferring qualities from secondary measurements. The above decisions are by no means trivial and represent key aspects of the controller synthesis problem that have attracted significant attention over the past four decades (Buckley, 1964; Weber and Brosilow, 1972; Morari et al., 1980; Larsson and Skogestad, 2000; Stephanopoulos and Ng, 2000).

Once the regulatory level is decided upon, the remaining decisions relate to how to structure the MPC layer: Should one controller or multiple MPC controllers be used? For each controller, there is the issue of deciding on the manipulated variables, the controlled variables, and the feedforward variables. Non-linearities are other issues that must also be addressed, if significant in an application. Note that the techniques discussed here are based on approaches that retain a linear(ized) dynamic model at the core of the MPC engine.

The typical MPC project sequence is as follows:

Pretest and Preliminary MPC design.

Plant Testing.

Model and Controller Development.

Commissioning and Training.

In the pretest phase of work, the key activity is one of determining the base level regulatory controls for MPC, tuning of these controls, and determining if current instrumentation is adequate. The outcome of this phase is a list of issues that must be fixed or resolved before plant testing can proceed. Typical problems that are identified are valve issues (sizing and excessive valve stiction), faulty instruments, and sensor location. The other task that begins in this phase is one of learning the process and understanding the operational challenges and expected constraints. In addition, a preliminary design for the MPC is typically performed, i.e., identification of controlled and manipulated, and number of MPCs.

Plant testing consists of generating plant data for model identification. Additional process knowledge and insight comes from this phase of work. Testing requires moving all inputs that may be manipulated variables for the MPC. Testing may be performed manually or automatically. During this phase of work, frequent lab measurements are collected, if an inferential model of product qualities is required.

In the next phase or work, modeling of the plant is completed, including any required inferential s and non-linear compensators. It is here that the models are analyzed for consistency. The final design for the controller or controllers is completed and simulations performed to test the model and tune the controller.

Commissioning involves turning on the controller and observing its performance on the plant and making tuning adjustments as needed to obtain a properly functioning controller. Training of operations staff on the live controller is begun in this phase.

In the following, we provide a high level description of MPC, MPC, without much emphasis on the particular theoretical properties of the MPC algorithm, for which there is already a substantial body of work (Mayne et al., 2000.). Subsequently, we present a detailed discussion of the key tasks and decisions that are made in the course of an implementation. Where appropriate, current practice is highlighted and guidelines are given. The impact of recent technological enhancements that have appeared are discussed. Lastly we suggest areas where improvements may be made.

2. MPC OVERVIEW

A simplified block diagram of the typical MPC is shown in Figure 1. Key functionality of the components shown in the figure are described below.

Target Selection: Target selection determines the best feasible, steady-state operating point, $\mathbf{x}_k^s, \mathbf{u}_k^s$ based on steady-state gains of the model. It can be implemented on the basis of minimizing deviations from desired steady-state "resting values" or as the result of an economic-based steady-state optimization, typically either a linear program (LP) or a quadratic program (QP).

Controller: The controller determines optimal, feasible future inputs to minimize predicted future errors, over a moving horizon, from targets determined by target selection. Tuning parameters (e.g., weights) are used to establish the dynamic

objectives and trade-offs. A QP is typically used to perform the controller optimization.

Estimator: The estimator updates the model estimate to account for unmeasured disturbances and model errors. It includes a deterministic part that models the effect of controller-manipulated process inputs (and other measured process inputs) on the process outputs, and a stochastic part (which may only be implicit) that models the effect of unmeasured disturbances on the process outputs. The simplest form for the estimator is the original MPC output correction (and still widely used today), where the current offset between the measurement and the model prediction is used to bias future model predictions. A state space model represents a more general and flexible approach to modeling unmeasured disturbances in the estimator.

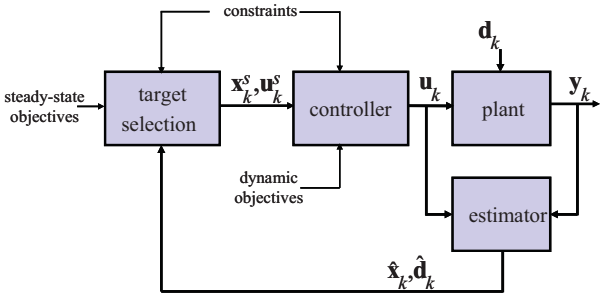


Figure 2. Simplified MPC block diagram

Various model forms are used in the various MPC products available today. Most common are the finite step response (FSR) or finite impulse response (FIR), but state space model formulations are also found. Recent controller products suggest a trend towards increased use of state space model formulations, because of the flexibility they offer to represent stable, integrating, and unstable processes in a single structure.

Our intent is not to delve into differences between the formulation and options of the various products. The interested reader is referred to Maciejowski (2002) and Qin and Badgwell (2003) Suffice it to say that differences exist among the products as to the approaches taken, but that they address important features such as prioritization of constraints, economic objectives and tuning parameters to influence CV vs. MV variance trade-offs. Most MPC controllers today force consistency between the sequence of input moves generated by the controller and the steady-state solution determined by the target selection. This consistency, which is equivalent to the imposition of a terminal constraint, provides nominal and robust stability (Genceli and Nikolaou, 1993; Rawlings and Muske, 1993; Ying and Joseph, 1999)

1. DCS STRATEGY

In deciding upon an appropriate DCS strategy for the MPC, there are several factors that need to be considered and balanced. Major factors are disturbance rejection, process interaction, robustness to model errors, and constraint considerations. Another factor is the influence of the DCS strategy on the settling time of the system, which affects the control horizon in MPC.

Fortunately, when implementing MPC, an existing DCS strategy is in place that can be evaluated and changed, if necessary. We are aware that some practitioners choose to use existing DCS schemes "as is" as opposed to modifying or pairing the PID loops in a different way. However, such modifications can have a significant impact on both MPC control performance and the ease of implementation (e.g., testing). Note that with modern DCS systems a different DCS strategy ("fall-back") may be used when MPC is switched off or fails.

A typical decision concerns whether to incorporate a cascade, such as temperature to flow cascade on a distillation column, or a temperature to pressure cascade on a direct-fired heater. As we have discussed, the DCS typically operates at a higher sample frequency than the MPC; therefore an existing cascade, if tuned well, will likely have much better disturbance rejection capability than the MPC. An additional advantage is that a cascade may help to linearize important CVs controlled by the MPC (because of the linearizing effect of feedback in the inner loop in a cascade scheme). This can be advantageous in providing acceptable control over a wider range of, e.g., plant feed rates.

The thinking with respect to cascades with MPC has clearly evolved over the years. In earlier days of MPC, it was often thought preferable to "break" an existing TC cascade and design the MPC to manipulate flow controllers. The motivation was that this would lead to simpler (overdamped) models and allow the interaction to be addressed by the MPC. What was missed with this approach was the rejection capability of the DCS via the higher sampling frequency, and the robustness that results from incorporating a TC into the MPC strategy. Consider the case of the two-by two subsystem associated with the product purities of a binary distillation column, controlled in the reflux-boilup configuration (so called *L-V* configuration). Consider two cases: 1), MPC control of compositions via *L* and *V* and 2) MPC control of the compositions via L and a stripping section TC controller that manipulates boilup. We assume that the controlled temperature correlates well with the bottoms product composition. The model relationships for these two cases are

$$\begin{bmatrix} y \\ x \end{bmatrix} = \begin{cases} L-V: & \begin{bmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{bmatrix} \begin{bmatrix} L_{sp} \\ V_{sp} \end{bmatrix} \\ L-TC: & \begin{bmatrix} g_{11}^{TC} & g_{12}^{TC} \\ g_{21}^{TC} \approx 0 & g_{22}^{TC} \end{bmatrix} \begin{bmatrix} L_{sp} \\ TC_{sp} \end{bmatrix} \end{cases} \quad (1)$$

Due to its lower triangular structure, *L-TC* is a more robust formulation compared to a full decoupling strategy with manipulated variable *L* and *V*, especially if the process is ill-conditioned, or more accurately, has large RGA elements (Skogestad and Morari, 1987). In most cases a temperature

cascade would be retained if it performs well. In a distillation column, it may be necessary to select another tray temperature if the existing one does not correlated well with product quality. Note that dual-ended temperature controls would normally be avoided because of interactions and the potential for the controllers to wind up (i.e., saturate) , if a section of the temperature profile shifts to a region of insensitivity (e.g., due to a feed composition change).

Another cascade decision concerns level to flow cascades, associated with feed drums, reflux accumulator drums and distillation column sumps. The questions is: should a flow be controlled directly by the MPC (with the associated level controlled by the MPC)? A motivation for doing so is to obtain a direct handle on inflows, without the dynamics of the level controller. Such an approach is useful when a plant capacity constraint exists, such as column flooding, and unit feed rate is also manipulated by the MPC. By directly manipulating column feed, tighter control of a plant capacity constraint can be achieved by taking advantage of liquid holdup in intermediate drums. Additional justification is to shorten system settling time by removing the dynamics of level controllers.

A disadvantage of including levels in the MPC is that levels, which are integrating variables with respect to flow, are that they harder to keep in bounds during an open-loop plant test. Levels are affected by both material and energy balance effects. While material balance effects may be straightforward to model, energy balance contributions affects must also be modeled, which tend to contribute over a longer time frame. Part of the challenge with integrating variables is related to the identification problem, as it is common to identify the first difference of an integrating CV, which decreases the signal-to-noise content. Note: in some FIR-based ID methods a double difference is used - one difference for the integrator and an additional difference (for both inputs and outputs) to remove integrating or slow disturbance effects. An additional challenge is that it is common for an MPC controller to contain logic to switch off if an integrating variable cannot be balanced (zero difference) at steady state., thus making integrating variables more sensitive to measurement spikes.

An alternative to controlling the level in MPC is to keep the level cascade in the DCS and manipulate the level setpoint to influence the corresponding flow rate (taking advantage of buffering capacity). In this case, the level measurement could also be brought into the MPC as a CV (and controlled within bounds). For this situation, model relationship between level setpoint and flow is zero gain (i.e., dynamic response only, zero steady-state gain). We should note that practitioners are divided on what is the best approach, although most of the experience is with FIR- or FSR-based MPC. Examples of step response models for these two cases are shown in Figure 3.

We should note that the theory and experience-to-date indicate that integrating variables are more easily handled within a state-space formulation, as it allows more flexibility in the unmeasured disturbance model – i.e., selection of the disturbance channels and incorporation of additional output measurements (Qin and Badgwell, 2003; Froisy, 2009).

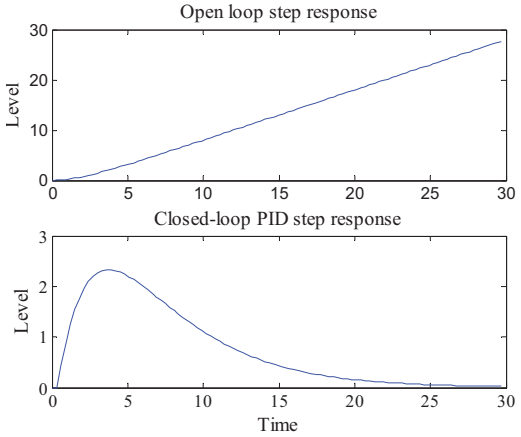


Figure 3. Step response models for integrating level: open-loop vs. closed-loop with PID

An important issue concerns valve positions of PID loops that are directly manipulated by MPC or that are affected by other manipulated variables of the MPC. For example, manipulating an FC controller setpoint will affect the valve position associated with the FC as well as the valve position associated with a downstream pressure controller. When a valve approaches a saturated state (either fully open or closed), not only is PID control lost for its associated controlled variable, but model mismatch (and nonlinearity) is introduced to all MPC-controlled variables that depend on the PID controller response. As a result, the MPC needs to keep PID controller outputs in a controllable range. This can be achieved by bringing the PID controller output into the MPC as a controlled variable. This approach is illustrated in Figure 3a. In this case, MPC manipulates the setpoint of the PID controller setpoint as necessary to keep the controller output in range. How well the PID output can be controlled and how close to saturation the MPC limit can be placed depends on: PID tuning, disturbances characteristics, and the degree of nonlinearity. It may be necessary to retune the PID loop based on the response of the controller output (a smooth response in the valve, without significant proportional "kick" is desirable).

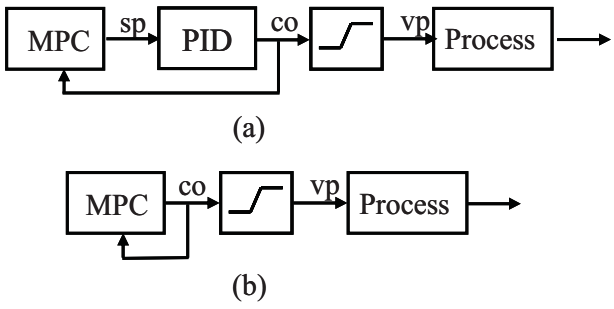


Figure 4. Alternate MPC strategies for maintaining valve positions in controllable range – (a) MPC to PID and (b) direct output to valve.

If a valve associated with a PID controller saturates more than 25% of the time, or if economics dictate operation at a fully open or closed-loop state, it may be preferable to directly manipulate the controller output directly, as shown in Figure 3b. In this way, the valve limit can be strictly enforced, resulting in control closer to the true valve limit. In this situation, additional disturbances may result from opening the PID loop that need to be addressed by the MPC

Regardless of the strategy, valves issues often arise in a project. Significant valve stiction (if greater than say 2%) must be corrected. In addition, valve nonlinearities may require compensation as part of the MPC strategy.

Example To illustrate how the various considerations discussed previously influence the MPC design, consider the two-column configuration shown in Figure 5, which is to be part of an MPC application that maximizes plant feed rate (not shown).

The following convention is used: ZC.sp denotes the setpoint of a PID loop to control Z; ZC.pv denotes the process variable or feedback variable for loop ZC; and ZC.op represents the output signal sent to the valve position.

For this example, assume it is known that the second column is susceptible to flooding, as indicated by a high value in DP1.pv, and that PC2op often saturates fully open. Because flooding is a constraint for column two, we would consider breaking the LC2 cascade and directly manipulating flow FC3.sp in the MPC. Due to the saturation potential of PC2, we would also consider directly manipulating its valve via PC2.op and controlling pressure within the MPC. If both temperature controllers perform well and the associated temperatures are good indicators of composition, they would be retained. These considerations then lead to an MPC with the following manipulated variables:

FC2.sp - column 1 reflux flow controller setpoint.

TC1.sp - column 1 temperature controller setpoint.

PC1.sp - column 1 pressure controller setpoint.

FC3.sp - column 2 feed flow controller setpoint.

FC5.sp - column 2 reflux flow controller setpoint.

TC2.sp - column 2 temperature controller setpoint.

PC2.op - column 2 pressure controller output.

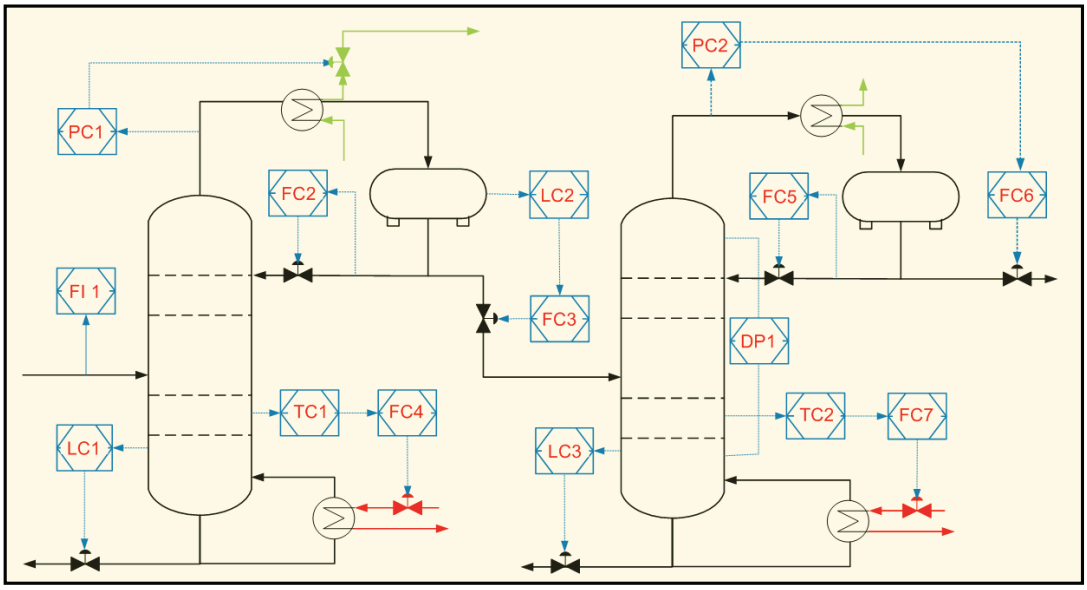


Figure 5. Example process to be controlled by MPC

2. PLANT TESTING

The plant test and subsequent model identification are the most important steps in an MPC project, and incur the most time, often representing more than 50% of the total project time. The importance of the accuracy of the plant model for MPC cannot be overstated. One cannot simply tune an MPC controller to compensate for a poor model. Further the effort involved in testing and identifying an MPC model is not a one-time event. To ensure adequate performance of an MPC application and sustain its benefits over time, it is necessary to re-perform plant testing to update the MPC model (all or in part) when control performance deteriorates due to a process change such as a process revamp.

Until the mid 90's, it was typical practice to conduct manual, open-loop tests, concentrating on the testing of one manipulated variable at a time, but moving other process inputs as necessary to maintain process operation in a desired region. Automatic testing via uncorrelated binary sequences

such as PRBS or GBN increased in popularity in the mid to late 1990's, and closed-loop testing approaches started appearing in the early 2000's. Today we are witnessing increased use of multivariable closed-loop testing methods in the industry as a means to reduced costs (human effort and time) and improved model accuracy due to richer data sets. Of course, an initial model must be available to perform a closed-loop test. An initial model may be available from an existing controller; otherwise, an initial model may need to be developed (e.g., from pretest data).

All of the above testing methods continue to be used today. Some MPC engineers continue to advocate manual testing methods, arguing that it is more conducive to developing process knowledge. While this is indeed an important step, we believe that sufficient process knowledge can come from the pretesting phase and the early stages of an automatic or closed-loop test, where the testing may start with just a few inputs.

Regardless of the testing approach, it is important to generate data in the frequency range of interest. This requires varying the pulse widths of the input signals, e.g., from 10% to 125% of the estimated settling time. A typical guideline is to achieve an average pulse width of an (uncorrelated) input signal equal to $1/3^{rd}$ of the open loop settling time of the process. Automatic signals can easily be generated to achieve a desired average pulse width. Input amplitudes are selected to keep process inputs and process outputs within desired ranges, but should be large enough to overcome valve stiction limits. Larger amplitude moves are preferred as long as the process responses remain within a linear range (unless linearizing transformations are used). A goal is to obtain a signal-to-noise ratio of at least 6-to-1.

The closed-loop testing approaches that have been developed for MPC also utilize uncorrelated binary signals. In Zhu (2001), generalized binary signals (GBN) (Tulleken, 1990) are applied to selected manipulated variables as dithers (added to MPC-generated manipulated variables) and to certain MPC-controlled variable setpoints. In Kalafatis et al. (2006), a closed-loop testing approach is described in which GBN binary signals are generated within the multivariable controller to maximize MV amplitude while keeping predicted CVs within preset constraints. Control action is only applied when predicted CVs exceed their limits.

Important quantities not measured online may require development of an inferential model. Generating data for inferential model development represents a much better approach than using only historical data, which typically has insufficient excitation and feedback effects. To ensure adequate data for model development, the process is moved to different steady-state operating values during the course of the plant test. Note that it is important to get accurate time stamps of the lab samples so that the data can be properly synchronized with measured plant test data for model identification. Due to the importance of the plant model, it is important not to stop a plant test prematurely. As a result, it is good practice to perform model identification throughout the testing phase until model quality is deemed adequate.

3. IDENTIFICATION METHODS

Dynamic Modeling. Various model structures are routinely used in the identification of models for MPC. Low order, parametric techniques continue to find application; however, these are nonlinear approaches, which require specification of model order (which is not straight forward). Processes with heat integration, recycle and/or embedded PID loops typically require higher order models to capture the resulting complex input-output behavior. As a result, we continue to find that finite impulse response (FIR) and high-order ARX (auto-regressive with exogenous input) models remain popular in MPC applications, both of which can be identified with linear least squares methods. For the FIR structure, smoothing techniques are used to reduce parameter variance (e.g., Dayal and MacGregor (1996)). Model reduction techniques are typically used with high-order ARX models to reduce parameter variance (see, e.g., Zhu (1998)).

We have witnessed increased use of subspace identification methods in industrial MPC applications over the past 10 years. This follows the development of these algorithms in the 90's (Larimore, 1983; Larimore, 1990; Overschee and De Moor, 1994). A key advantage of a subspace method is that it directly yields a multivariable state space model, which is an advantage for a state-space controllers. However, even for FIR- or FSR-based MPC, a subspace method offers advantages as it considers the correlation of the output measurements in the identification, thus leading to a potentially more accurate and robust model. Industrial experience with a subspace identification method has been discussed in Zhao et al. (2006). Their experience has shown that complex relationships can be accurately modeled with a state space model of relatively modest order (range of 5 to 15), which captures both slow and fast dynamics. Advantages compared to a parametric technique are that the model order selection can be automated and only linear methods are required. Compared to FIR models, their experience has shown that subspace leads to more accurate estimates of gain and gain ratios, which are critical to capturing the true degrees of freedom in the MPC and ensuring reliable LP performance.

For the closed-loop situation, traditional subspace methods are biased; thus, special treatment is required. Modifications can be made to subspace methods that lead to consistent estimates (as summarized in (Qin, 2006)), although in theory, prediction error methods (e.g., ARX) lead to estimates with lower parameter variance. A challenge with closed-loop identification (using a direct approach) is the importance of obtaining an accurate noise model, which is problematic in practice, since typical process disturbances cannot be captured by white noise, passed thru a linear filter. In practice, one can attempt to minimize the bias by "overwhelming" noise feedback in the frequency range of interest (Jorgensen and Lee, 2002).

Important decision made during the model identification step relate to the following:

Data slicing Determining the sections of data should be included/excluded in the identification.

Data pre-processing Includes such option as spike removal, offset correction, prefiltering/detrending options, and shifting data based on known delays.

Selection of input and outputs – inputs include both candidate manipulated variables and measured disturbances.

Model Structure This includes decisions such as FIR model length, model orders of ARX or subspace, integrating variable or not

Nonlinearities Do nonlinearities warrant additional modeling?

Each of the above steps are typically iterative. With data slicing, the important issue is removing data that would otherwise lead to model bias. This includes time periods with significant unmeasured disturbances or plant upsets, such as pump shutdown, or where valve saturation occurs with PID loops. Prefiltering/detrending, can significantly impact the identification results. It is important to pre-filter/detrend to suppress slow drifts and minimize their contribution to model bias. In some MPC identification packages user options for prefiltering/detrending are not provided. Data differencing is often used, but since it suppresses low frequency information can lead to model gain errors.

In the selection of inputs and outputs, one will have a good idea of which are the manipulated and controlled variables, but it may not be as clear as to which other inputs should be selected as disturbance variables. Note it may be desirable to include a disturbance variable simply as part of the identification step to improve the quality of the models to the key manipulated variables, and not use it as a feedforward variable in the controller. With a subspace identification method, due to the fact that it explicitly considers the correlation of the outputs, the proper selection of output variables can improve the model accuracy of a given input-output channel, regardless of whether the additional outputs are used in the controller.

An aspect of model structure selection is whether to model a controlled variable as an integrating variable. Many times, process knowledge will guide this decision (such as liquid level to flow). However, slow responding stable variables (slower than the controller prediction horizon of the controller) often lead to improved control if modeled as an integrator, especially if they are subjected to input-type of disturbances.

Nonlinearities are typically handled with a static linearizing transformation on inputs and/or outputs. This is the familiar Hammerstein and Weiner model structures, as shown in Figure 6. In typical MPC practice, these static nonlinear functions are SISO (one-to-one) as opposed to MIMO. This is because a MIMO structure would be problematic when constraints are imposed. With physical insight, one may have knowledge as the functional forms such as valve-flow relationships or logarithm of distillation product impurity.

Figure 6. General Hammerstein-Wiener model structure; f_H is the Hammerstein static nonlinear transformation, f_W is the Weiner static nonlinear transformation.

For the general case, when a specific nonlinear transformation is unknown, a piece-wise linear relationship can be empirically derived, assuming testing is over a range wide enough to capture the nonlinearity. An example is shown in Figure 7 for the case of a valve position (controller output) and an associated measurement (e.g., flow). This transformation could be used with either of the valve position scenarios shown in Figure 4.

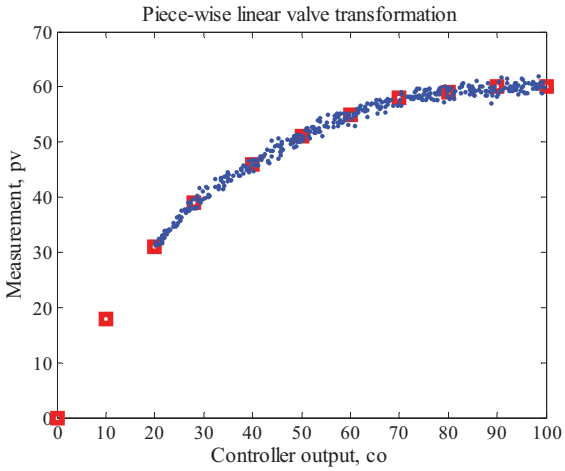


Figure 7.Example piece wise linear transformation

Many of the commercial MPC include the necessary pre- and post-processing capability to handle Hammerstein and Weiner transformations. To deal with dynamic nonlinearity one can use multiple models and "schedule" these based on knowledge of the operating point. Although this would be an easy thing to do, it is not commonly done with empirical models. An example of where multi models *are* routinely used is in ethylene applications, where there is a different furnace model for each major feed type.

Inferential modeling. For the situation where an inferential model must be developed for product qualities that are not measured online (measured infrequently by lab), a couple of approaches can be used.

The most common is to develop a regression model of the quality from directly measured variables such as flow, temperatures, and pressure. It is common for the multiple measurements (for example temperatures) used as inputs to the regression to be correlated. This requires multivariate regression techniques such as principal component regression (PCR), principal component analysis (PCA) and partial least squares (PLS). The key idea is to project the measurement values into a reduced number of important directions

(number of directions less than the number of measurements) to avoid problems associated with correlation/ill-conditioning. Improved regression modeling is possible if a steady-state simulation model is available. In this case, measurements can be selected to minimize steady-state offset in the primary variables (lab measured) for expected disturbances and setpoint changes (Pannocchia and Brombilla, 2003; Hori et al., 2005).

The other approach to inferential modeling utilizes a simplified, fundamental (nonlinear) model of the processes, where parameters in the model are tuned (or optimized) to best fit the model to lab samples. In this case, the nonlinear model provides feedback to the MPC. The advantage of this approach (assuming the model is adequate) is that less process testing is required to fit the inferential model and the model can be expected to operate satisfactorily over a wider range of operation compared to a purely regression model. See, e.g., Friedman (2001), where a static nonlinear model is used for prediction of distillation product compositions.

4. CONTROLLER DEVELOPMENT

An MPC application is typically applied to a unit such as a fluid catalytic cracking unit (FCCU) or ethylene unit. A single MPC or multiple MPC controllers may be applied, depending on the unit objective and constraints. Consider as an example the FCCU shown in Figure 8. If the unit objective is to maximize unit feed and downstream throughput constraints exist, such as DC4 flooding, one would consider a single controller. If there are no throughput constraints in the downstream columns, one would consider two MPC controllers: one for the reactor/regenerator/ main fractionator/wet gas compressor, and one for the all of the downstream distillation columns.

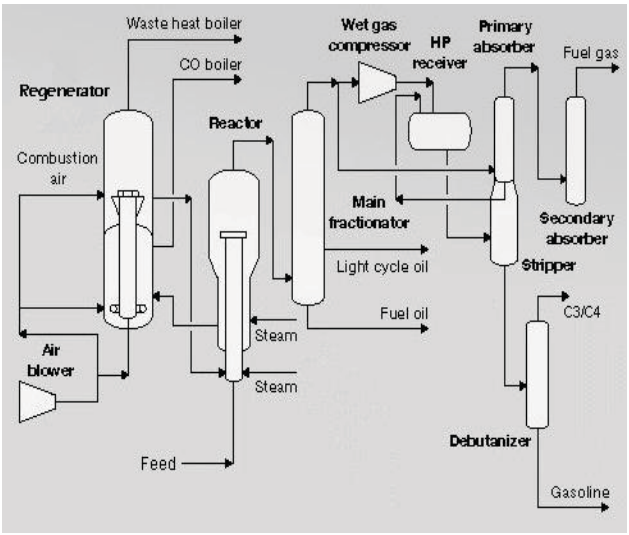


Figure 8. Fluid catalytic cracking unit.

A single unit controller is harder to implement and maintain, and if not implemented properly, or if sufficient engineering expertise is not available onsite, the result may be low MPC service factors or a controller that does not meet economic objectives. When a controller is not performing correctly or is not understood by operations, operators will typically "pinch" manipulated variables (set upper and lower limits close to each other) to overly constrain the MPC in order keep control within a region that the operator feels comfortable. When limited resources available are available, an alternative would be to first implement distributed controllers and later consolidate controllers after experience and confidence is gained.

It is good practice to develop models on an individual equipment basis. For example model reactors and distillation columns separately and build up the overall model from the various sub-models. Thus, the modeling should not be treated as one black-box, linking all inputs to all outputs (Haarsma and Mutha, 2006). If the modeling of the individual equipment is done properly, the key manipulated and controller variables have been identified and modeled and the manipulated variables for the overall model is the super set of MVs and CVs for the sub-models. Note that feedforward variables in the sub-models need to be truly independent variables from the viewpoint of the assembled model for the MPC.

With the above approach, it is typical to develop a sub-model based on its feed measurements (e.g., the feed rates to the primary absorber / stripper in Figure 8), but the overall MPC may require a model expressed in terms of unit feed. In this case, one can develop the required model from a convolution of the primary absorber sub-model and a model from unit feed to primary absorber feed. Note that the model prediction errors in the predicted feed to primary absorber feed can be used as a feedforward variable to the primary absorber / striper and DC4 with this arrangement. This is sometimes known as a prediction error feedforward in MPC jargon.

It is good modeling practice to ensure that the MPC model satisfy material balances (delta flows in equal delta flows out). When levels are controlled in MPC, the material balance consistency implies that the rate of change of levels *and* flows equal zero at steady state. Another area of consistency is where embedded PID loops imply a unity or zero gain.

As we have mentioned previously, the accuracy of the steady-state gains is critically important as they determine the steady-state operating point (target selection layer in Figure 2). This, in turn, can have a significant effect on the control layer as both target selection and dynamic control are executed at the same frequency. The challenge is that gains from an empirical model may not represent the true degrees of freedom that exist in the plant. As a result, the target selection layer may exploit fictitious degrees of freedom, a problem that tends to get .worse with problem size (due to the increased number of possible submatrices).

Consider the case that an LP is used as the target selection. At each execution , the LP will invert a square sub-matrix of the overall gain matrix. If the sub-matrix is ill-conditioned, the resulting changes to the plant may be excessive, possibly leading to cycling or instability. This normally becomes an issue when key manipulated variable handles are constrained (and therefore unavailable) and weaker manipulated variables

must be used. Note that a degree of freedom can be removed from the LP by fixing gain ratios (forcing exact colinearity). A key modeling issues is deciding whether a degree of freedom exists or not. This decision can be guided by the models themselves and their uncertainty) or from engineering insight. Two approaches are used in practice to help with this problem. One approach is to analyze various sub-matrix combinations of the gain matrix in terms of singular value decomposition (SVD or the relative gain array. Sub-matrices with high condition number or large RGA elements become candidates for forcing a collinear relationship, particularly when expected gain errors suggest a co-linearity. Another approach is an online method that automatically disregards small singular values in the sub-matrix inverse, based on user defined tolerances (Qin and Badgwell, 2003). In the authors' opinion, neither approach is completely satisfactory. Analyzing sub-matrices can be a time consuming task and tuning with singular value tolerances can lead to unexpected effects.

During the controller development phase, initial controller tuning is performed. This relates to establishing criteria for utilizing available degrees of freedom and control variable priorities. In addition, initial tuning values are established for the dynamic control. Steady state responses corresponding to expected constraint scenarios are analyzed to determine if they behave as expected, especially with respect to the steady-state changes in the manipulated variables. This step may force additional analysis and treatment of gains and gain ratios.

5. COMMISSIONING

One reason we want to execute the various project steps well is to minimize rework in the commissioning phase. In the best case, commissioning of the controller involves simply making tuning adjustments and observation of the controller under different constraint situations and plant disturbances. In the worst case, control performance is unacceptable and the control engineer is forced to revisit earlier decisions such a base level regulatory strategy or plant model quality. Both of these can lead to retesting and re identification of at least portions of the plant model, resulting in delays and possible cost overruns.

During commissioning it is typical to revisit model decisions and assumptions, and switch out certain models, or modify gains, to obtain acceptable control. Typically, 50-70% of the commissioning effort deals with models.. Commissioning typically takes place over a 2-3 week period. In reality, commissioning is an ongoing effort, although the subsequent effort is normally treated as controller support and maintenance. During the commissioning phase there are only so many different constraint and operating scenarios that can be considered. Certain operating scenarios and constraint sets can *only* be observed certain times of the year due to seasonal effects. It is therefore important that the operating company have in-house expertise that can be used to answer questions ("whys is the controller doing that?"), troubleshoot, and resolve problems that arise over time.

Once a controller is commissioned, it is important to monitor controller performance to ensure benefits are maintained. Unfortunately, multiple factors can contribute to controller performance deterioration. A change in the operating point or a plant modification may invalidate portions of the plant model. Performance degradation of other control systems (PIDs and MPCs) can lead to poor performance. For example, a PID loop associated with, or upstream of, an MPC may develop a cycle resulting from valve stiction. While technology can help with the diagnosis, ultimately expertise must be brought to bear to resolve and correct the problem. Left uncorrected such problems lead to low service factors, or worse, an MPC being permanently switched off.

6. TRENDS AND SUGGESTED RESEARCH DIRECTIONS

The impact of faster and multi-core processors are being seen in MPC products. Increased processing speed is allowing an increased number of future moves to be calculated over the control horizon and also allows for much faster controller execution. In Barham (2006), an MPC approach is described in which all manipulated variables are valve positions. It is applied to an entire FCCU, and executes on a six-second interval. Transformations are used to linearize the relationship between valves and controlled variables. In Froisy (2006), a new state space controller is described that is based on an infinite horizon move plan, Notable features include model assembly of smaller submodels into one large overall MIMO state space model, and an automation feature that simplifies the configuration and tuning of disturbance estimators within a dynamic Kalman Filter framework. We are also seeing increased offerings of MPC at the DCS level where it can execute at a 1 second interval. However, unit wide or multi-unit MPC implementations are still most often implemented in a separate, dedicated computer.

In the remainder we provide suggestions for areas of improvement, including ideas for how this might be accomplished. General themes are of facilitating improvements at the various steps in an MPC implementation, maximizing the use of data and a priori knowledge, and minimizing the impact of changing or modifying key design decisions.

6.1 DCS Strategy

As we have discussed, decisions related to structure selection of the combined MPC-DCS system are multifaceted. Fortunately, there is experience with many of the major refinery and chemical units that can guide these decisions, although specific experience may not always be sufficient for a particular plant (due to idiosyncrasies of the particular plant). This is of course problematic for processes or industries where MPC has not been previously applied. As a result, the path to an acceptable MPC controller may involve iteration. It is therefore advantageous if rework can be minimized in light of design changes. It is also clear that methods that rely on systematic design rather than trial and error only would be valuable.

In recent years, techniques and products have been developed which apply multivariable identification methods to develop models that are in turn used to tune PID loops. Such approaches can be used to improve the performance of PID loops associated with an MPC system with reduced engineering effort (Zhu, 2003; Harmse et al., 2009). Note that once a multivariable model is available (relating the effect of PID controller outputs on PID controlled variables), one could use standard techniques such as relative gain array (RGA) or block relative gain (BRG) (as a function of frequency) to focus on the most promising PID loop pairings and simulate various suggested pairing possibilities. Experience has shown that testing and developing a multivariable model for the typical loops found at the DCS level can often be completed within a day for the typical loops that are found at the DCS level (Darby and Harmse, 2009).

Changing the PID loop pairings or tuning parameters (if behavior is significantly different) requires a change to the affected models in an MPC system. Historically, such changes have required plant retesting. However, with completed knowledge of the models at the PID level (including the PID controllers themselves), it is theoretically possible to convert the MPC models to reflect a different PID configuration or tuning), potentially avoiding an expensive retest. Such an approach is described in Rejek et al. (2004). One major claim for this approach is that one could perform the plant test in one DCS configuration, but implement the MPC in another configuration as in Barham et al. (2006). To our knowledge, various options for solving this problem have not been investigated. Open questions concern accuracy and robustness issues as to how best to perform this model conversion.

6.2 Plant Testing

It is well known that that independent binary input test signals are generally inadequate (inefficient at best) for the identification of ill-conditioned systems. The reason is that the weak process directions (e.g., separation changes in a distillation column) are poorly identified in the presence of noise. The solution is to use correlated inputs, which can be generated in open or closed loop (Anderson and Kummel, 1992; Koung and MacGregor, 1994; Li and Lee, 1996). For example, in a distillation column, large changes in both reflux and reboil flow rates are required to adjust separation in any significant way. As discussed previously, ill-conditioning is often found in the MPC steady-state gain matrix. Properly designed input sequences can be expected to improve estimation of ill-conditioned sub-matrices. Recent results show that independent binary signals can also be inferior for well-conditioned systems, depending on the active constraints. In Darby and Nikolaou (2008a, b), using a criterion which maximizes the likelihood of satisfying integral controllability, optimal inputs (both amplitudes and covariance of the inputs) are shown to depend on both the system's conditioning and the specific active constraints. While correlated inputs can be achieved with independent perturbations of controller CV setpoints or limits, such an approach may translate into ineffective input perturbations due to the influence of the target selection layer. A closed-loop experimental design approach for MPC would be desirable, although treated rigorously, would require knowledge of the feedback law. This would be problematic for MPC as each constraint set represents a different control law. A possible approach is to replace the binary test signals that are currently used in closed-loop MPC with a traditional or control-relevant experimental design. An experimental design could be performed consistent within a feasible region established for the target selection layer and implemented through the controller to ensure constraint satisfaction. This might be done in a manner similar to that used in Sagias (2004) for PRBS signals, where the dynamic objective function is modified to the allow trade-off of control and test objectives. The other aspect of experimental design concerns frequency content. This aspect would need to be investigated as well. We note that with current MPC practice, the frequency content is specified indirectly based on the type of binary signal chosen and the specified average pulse width. Extending this concept further, if basis functions with desired frequency content were pre-specified, this might allow the experimental design to be expressed in terms of input amplitudes and covariances.

6.3 Identification

As mentioned earlier, there are multiple consistency relationships (e.g. gains and gain ratios) that should be enforced in the constructed MPC model. Instead of imposing these conditions by altering the identified model as a post-processing step, it would be better to incorporate these as constraints in the identification. Within the context of least squares, imposition of linear constraints results in parameter estimators with smaller variance (Seber and Lee, 2003).

Other consistency relationships could be incorporated. Material balance consistency has been discussed, but consistency can be extended to component balances. For example, for binary distillation columns, relationships can be derived which link the steady-state gains associated with top and bottom purities for a given regulatory control structure (Häggbloom and Waller, 1988).

Another area that should be exploited is a priori information available in the form of physical models. Such an approach was discussed for a steady-state inferential predictor. The basic idea is to combine available model information and data in a grey-box identification problem. The key motivation is one of getting better models with less data and effort, not necessarily one of capturing the nonlinear behavior. A linearized model from a fitted nonlinear model may well be adequate. However, it the nonlinearity were significant, the nonlinear model could be used to update models in the MPC. We should mention that in our view (for the foreseeable future) a full nonlinear model is not needed or feasible for the majority of control problems common found in industry (polymer applications and batch applications being notable exceptions). Tools to empirically determine the Hammerstein and Wiener static compensators would be useful (such as described in Zhu (2000) for the case of cubic splines). One

could also consider combing the dynamic identification step with piece-wise linear transformations in a single identification problem. We might expect that nonlinear models could be developed for certain submodels of an MPC, if the improvements or costs saving to develop the model are significant. The online implementation might be posed as a constrained estimation problems using techniques such as found in Rawlings and Bakshi (2006), where the estimator would provide feedback (and possibly future predictions) to MPC. An example of a common situation where nonlinear affects are often encountered is variable liquid hold-up (e.g., reflux accumulator), which causes a variable dynamic response in downstream composition signals.

6.3 Improved Disturbance Estimators

A key advantage in utilizing a general state space formulation is improved (unmeasured) disturbance modeling. For example, it is well known that the output bias approach, traditionally used for the MPC model update step, can lead to sluggish response to an input disturbance (Shinskey, 1994). A properly designed estimator overcomes this limitation (Muske and Badgwell, 2002; Pannocchia and Rawlings, 2003). An additional advantage of state space disturbance model is that of incorporating additional output measurements. A typical example is shown in Figure 9. The variable u represents the MPC manipulated variable and y is the MPC controller variable; yI is an intermediate variable. Examples include:

> Case 1: u is plant feed rate, y_I a downstream flow measurement and y is a downstream controlled variable.
>
> Case 2: u is column reflux flow rate, y_I a tray temperature, and y a product analyzer.
>
> Case 3: u is a PID setpoint, y_I is the PID error and, y is an MPC controlled variable..

Case 1 represents the example considered earlier (cf section 6). Case 2 provides structure similar to a traditional cascade control (Froisy, 2006). Case 3 models the behavior of the base level PID, e.g., a time series model of the PID error, which has the advantage that unnecessary moves are not generated by MPC when the base layer is capable of rejecting the unmeasured disturbance (Haarsma and Mutha, 2006). Improved disturbance modeling could also be applied to the situation where MPC is directly manipulating a valve (and local flow, temperature, and/or pressure measurements are available).

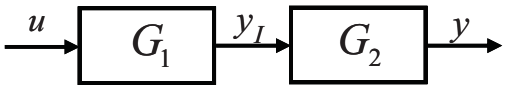


Figure 9. Plant model with intermediate variable y_I.

While it is possible to replicate such approaches within traditional FIR- and FSR-based MPC, they involve ad-hoc solutions. An interesting possibility is the use of improved disturbance estimators within traditional FIR- or FSR-based MPC (Badgwell, 2009). What is unclear is the extent to which these improved estimators are actually being used within existing state space controllers. Anecdotal evidence suggests a gap between these capabilities and actual usage by MPC engineers. Part of the challenge in developing these more general estimators is that it requires linking disturbances to particular model channels. Tools and techniques to facilitate or simplify this step would be helpful.

6.4 Robustness

Model errors impact performance of MPC at both the target selection layer and the dynamic controller, although, as we have seen, the problem is more acute with the steady-state target selection layer. While the target selection layer offers advantages in terms of constraint control, economic optimization, and dealing with non-square systems, it represents a source of challenges for an MPC implementation, ones that grow with the size of the controller. The goal is to prevent the optimizer from exploiting fictitious degrees of freedom, and from exploiting true degrees of freedom that may exist, but lead to large steady-state moves for only small economic improvement. Another challenge is minimizing the impact of effects that can lead to chatter in the steady-state targets (Shah et al., 2001; Kozub, 2002). This includes high frequency noise associated with controlled variables, unmeasured disturbances and/or model error. An approach that has been used industrially to minimize change from the optimizer layer to the dynamic layer is based on a minimum-movement criterion (to achieve all control objectives) in the dynamic layer and to invoke a QP optimization once all predicted controlled variables are within a pre-defined funnel (Lu, 2003).

Given the importance of the steady-state gain matrix in the optimizer, Kassman et al. (2000) proposed a robust LP formulation based on ellipsoidal uncertainty of the gain matrix. An advantage of their approach is that the resulting optimization problem is convex. An open question is how well their approach addresses problems with ill-conditioning. Note that we have avoided mentioning worse case formulations due to their tendency to provide overly cautious control for the average case.

The challenges outlined above could benefit from additional research. Pertinent questions are whether it is possible to avoid inverting the gain matrix for the entire plant and whether techniques could be used to avoid exploiting uncertain (and undesirable) degrees of freedom. These issues might be considered with the general problem of how to coordinate multiple MPCs, which is currently receiving increased research attention. .We note that when the plant optimum is consistent with maximum throughput, a simplified coordinator can be used (Aske et al., 2008). Such an approach explicitly limits the degree of freedom that are used in the plant wide control scheme.

7. CONCLUSION

The MPC algorithm is a mature technology and there is good understanding of the algorithm's properties and behavior. But as discussed, there are facets of the technology that could be improved. As one would expect, the performance of MPC systems does not depend only on the "control law" (MPC tuning) but on successful completion of all of the following steps: articulation of control objectives, selection of measurements and manipulations, configuration of controller structure (i.e. interconnections among MVs and CVs), and, finally, design of the control law (Stephanopoulos, 1984). Even though the control law can be designed in a fairly systematic way, completion of the design steps above it is less systematic, and offers a margin of creativity. Process understanding remains indispensable for these steps. Improving the ability to systematically complete these steps would certainly contribute towards designing better MPC systems. Industry and academia can continue collaboration towards this end, with full understanding of the need for sanitized academic solutions to bear industrial relevance and that common practice may not necessarily be best practice.

REFERENCES

Anderson, H. W. and Kummel, M. (1992). Evaluating estimation of gain directionality part 2: A case study of binary distillation. *Journal of Process Control*, **2**, 67-86.

Aske, E. M. B., Strand, S. and Skogestad, S. (2008). Coordinator MPC for maximizing plant throughtput. *Computers & Chemical Engineering*, **32**, 195-204.

Badgwell, T. A. (2009). Personal communication with T. A. Badgwell.

Barham, M., Watts, J., Bailey, M., Carroll, T., Garcia, L., Cutler, C., Hetzel, M., Szuhay, J. and Boyden, S. (2006). Adaptive dynamic matrix control on a fluidized catalytic cracking unit. *ISA 2006*, Houston.

Buckley, P. S. (1964). *Techniques of Process Control.* John Wiley & Sons.

Darby, M. L. and Harmse, M. (2009). Is it time for a new approach to PID tuners. *AIChE 2009 Spring Meeting*, Tampa.

Darby, M. L. and Nikolaou, M. (2008a). Multivariable system identification for integral controllability. *Automatica*, submitted.

Darby, M. L. and Nikolaou, M. (2008b). Design of experiments for identification of multivariable models satisfying integral
controllability: The dynamic model case. *AIChE Annual Meeting*, Philadelphia.

Darby, M. L. and Teeter, E. (2005). Best practices of MPC/RTO in ethylene. *AIChE Spring Meeting*, Atlanta.

Dayal, B. S. and MacGregor, J. F. (1996). Identification of finite impulse response models: Methods and robustness issues. *Industrial & Engineering Chemistry Research*, **35**, 4078-4090.

Friedman, Y. Z. (2001). Experience with GDS, a first principles inference model for distillation columns. *ERTC Computer Conference*, Paris.

Froisy, J. B. (2006). Model predictive control - Building a bridge between theory and practice. *Computers & Chemical Engineering*, **30**, 1426-1435.

Froisy, J. B. (2009). Personnel communication with J.B. Froisy.

Genceli, H. and Nikolaou, M. (1993). Robust stability analysis of l1- norm model predictive control. *AIChE Journal*, **39**, 1954-1965.

Haarsma, G. and Mutha, R. (2006). Implementation of unit wide advanced process control in ethylene plant. *AIChE Spring Meeting*, New Orleans.

Häggbloom, K. E. and Waller, K. V. (1988). Transformations and consistency relations of distillation control structures. *AIChE Journal*, **34**, 1634-1648.

Harmse, M., Hughes, R., Dittmar, R., Singh, H. and Gill, S. (2009). Robust optimization-based multiloop PID controller tuning: A new tool and an industrial example. *Advanced Control of Chemical Processes, ADCHEM 2009*, Istanbul.

Hori, E. S., Skogestad, S. and Alstad, V. (2005). Perfect steady-state indirect control. *Industrial & Engineering Chemistry Research*, **44**, 863-867.

Jorgensen, S. B. and Lee, J. H. (2002). Recent advances and challenges in process identification. *Proceedings of the Sixth International Conference on Chemical Process Control*, Tucson, 55-74.

Kalafatis, K., Patel, K., Harmse, M., Zheng, Q. and Craig, M. (2006). Multivariable step testing for MPC projects reduces crude unit testing time. *Hydrocarbon Processing*, 93-99.

Kassman, D. E., Badgwell, T. A. and Hawkins, R. B. (2000). Robust steady-state target calculation for model predictive control. *AIChE Journal*, **46**, 1007-1024.

Koung, C. W. and MacGregor, J. F. (1994). Identification for robust multivariable control - The design of experiments. *Automatica*, **30**, 1541-1554.

Kozub, D. J. (2002). Controller performance monitoring and diagnosis. Industrial perspective. *Fifteenth Triennial World Congress of the International Federation of Automatic Control*, Spain.

Larimore, W. E. (1983). System identification. reduced-order filtering and modeling via canonical variate analysis. *Proceedings of the 1983 American Control Conference*, San Francisco, 445-451.

Larimore, W. E. (1990). Canonical variate analysis in identification, filtering and adaptive control. *Proceedings of the 29th Conference on Decision and Control*, Honolulu, 596-604.

Larsson, T. L. and Skogestad, S. (2000). Plantwide Control - A review and a new design procedure. *Modeling, Identification and Control*, **21**, 209-240.

Li, W. and Lee, J. (1996). Frequency domain closed-loop identification of multivariable systems for feedback control. *AIChE Journal*, **42**, 2813-2827.

Lu, J. Z. (2003). Challenging control problems and emerging technologies in enterprise optimization. *Control Engineering Practice*, **11**, 847-858.

Maciejowski, J. (2002). *Predictive Control with Constraints.* Englewood Cliffs: Prentice Hall.

Mayne, D. Q., Rawlings, J. B., Rao, C. V. and Scokaert, a. P. O. M. (2000.). Constrained model predictive control: stability and optimality. *Automatica*, **36**, 789-814.

Morari, M., Arkun, Y. and Stephanopoulos, G. (1980). Studies in the synthesis of control structures for chemical processes Parts I, II, III. *AIChE Journal*, **26**, 220-260.

Muske, K. and Badgwell, T. A. (2002). Disturbance modeling for offset-free linear model predictive control. *Journal of Process Control*, **12**, 617-632.

Overschee, P. v. and De Moor, B. (1994). N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems. *Automatica*, **31**, 1853-1864.

Pannocchia, G. and Brombilla, A. (2003). Consistency of property estimators in multicomponent distillation control. *Industrial & Engineering Chemistry Research*, **42**, 4452-4460.

Pannocchia, G. and Rawlings, J. B. (2003). Disturbance models for offset-free model-predictive control. *AIChE Journal*, **49**, 426-437.

Qin, J. Q. and Badgwell, T. A. (2003). A survey of industrial model predictive control. *Control Engineering Practice*, **11**, 733-764.

Qin, S. J. (2006). An overview of subspace identification. *Computers & Chemical Engineering*, **30**, 1502-1513.

Rawlings, J. B. and Bakshi, B. R. (2006). Particle filtering and moving horizon estimation. *Computers & Chemical Engineering*, **30**, 1529–1541.

Rawlings, J. B. and Muske, K. R. (1993). The stability of constrained receding horizon control. *IEEE Trans. Automatic Control*, **AC-38**, 1512-1516.

Sagias (2004). *Closed loop identification of a process controlled by MPC.* Chemical Engineering, University of Houston, Houston, TX.

Seber, G. A. F. and Lee, A. J. (2003). *Linear Regression Analysis.* 2nd Ed. Hoboken, NJ, John Wiley.

Shah, S. L., Patwardhan, R. and Huang, B. (2001). Multivariate controller performance monitoring: Methods, applications and challenges. *Chemical Process Control VI*, Tuscon.

Shinskey, F. G. (1994). *Feedback Controllers for the Process Industries.* New York, McGraw-Hill.

Skogestad, S. and Morari, M. (1987). Implications of large RGA elements on control performance. *Industrial & Engineering Chemistry Research*, **26**, 2323-2330.

Stephanopoulos, G. (1984). *Chemical Process Control.* New York, Prentice Hall.

Stephanopoulos, G. and Ng, C. (2000). Perspectives on the synthesis of plant-wide control structures. *Journal Of Process Control*, **10**, 97-111.

Tulleken, H. J. A. F. (1990). Generalized binary noise test-signal concept for improved-experiment design. *Automatica*, **28**, 37-49.

Weber, R. and Brosilow, C. (1972). The use of secondary measurements to improve control. *AIChE Journal*, **18**, 614-623.

Ying, C. M. and Joseph, B. (1999). Performance and stability analysis of LP-MPC-QP-MPC cascade control systems. *AIChE Journal*, **45**, 1521-1534.

Zhao, H., Harmse, M., Guiver, J. and Canney, W. (2006). Subspace identification in industrial applications - A review of recent progress and industrial experience. *14th IFAC Symposium on System ID*, Newcastle.

Zhu, Y. (1998). Multivariable process identification for MPC: The asymptotic method and its applications. *Journal Of Process Control*, **8**, 101-115.

Zhu, Y. (2000). *Multivariable System Identification for Process Control.* Amsterdam, Pergamon.

Zhu, Y. (2001). Progress in MPC identification: A case study on totally closed-loop plant test. *ERTC Computing Conference*, Milan.

Zhu, Y. (2003). Robust PID tuning using closed-loop identification. *Advanced Control of Chemical Processes, ADCHEM 2004*, Hong Kong, 165-170.

Power-Shaping Control of an Exothermic Continuous Stirred Tank Reactor (CSTR) ⋆

Audrey Favache * **Denis Dochain** **

* *Université catholique de Louvain, Unité d'ingéniérie des matériaux et procédés, B-1348 Louvain-la-neuve; Université catholique de Louvain, CESAME, B-1348 Louvain-la-neuve (e-mail: audrey.favache@uclouvain.be).*
** *Université catholique de Louvain, CESAME, B-1348 Louvain-la-neuve (e-mail: denis.dochain@uclouvain.be).*

Abstract: The exothermic continuous stirred tank reactor (CSTR) is a classical yet complex case study of nonlinear dynamical systems. Power-shaping control is a recent approach for the control of nonlinear systems based on the physics of the dynamical system. In this paper we present a general methodology to apply the power-shaping control approach to the exothermic CSTR study case. It results in a global Lyapunov function for the exothermic CSTR. This Lyapunov function is then reshaped by the means of a controller in order to stabilize the process at a desired temperature. Some considerations on the local and global convergence to the desired state are presented.

Keywords: Nonlinear control; Power-Shaping Control; exothermic CSTR

1. INTRODUCTION

Thermodynamic systems, and among them chemical reaction systems, are usually nonlinear dynamical systems. They can therefore have a complex behaviour and be difficult to analyze and to control. Stability analysis of nonlinear systems requires the use of abstract mathematical tools such as the two Lyapunov methods or the passivity theory. Over the past years, several works have combined those abstract concepts with the underlying physical phenomena giving rise to the dynamical behaviour of the system. These works include for instance the study of port-Hamiltonian systems (Dalsmo and van der Schaft (1998), Maschke and van der Schaft (2005), Eberard et al. (2006)), energy-balancing passivity based control (PBC) (Ortega et al. (2001), Jeltsema et al. (2004)) or the introduction of the contact formalism for expressing the dynamics of systems in which irreversible phenomena arise (Eberard et al. (2005), Eberard (2006), Favache et al. (2007)). The exothermic continuous stirred tank reactor (CSTR) is a classical study case of nonlinear systems. Indeed, the dynamical behaviour shows complex features, such as multiple equilibrium points. Up to now no exact physical interpretation of the complex behaviour of the exothermic reactor has been found (Favache and Dochain (2009)).

Power-shaping control (Ortega et al. (2003)) has been developed in the past years as an extension of energy-balancing passivity-based control (Ortega et al. (2001), Jeltsema et al. (2004)). In energy-balancing passivity based control, the controller reshapes the energy function of the system so that it has a minimum at the desired equilibrium point. The controller provides to the system a finite amount of energy so as to drive the system to the desired state. This concept has been widely applied to electro-mechanical systems (Ortega et al. (1999), Maschke et al. (2000), Ortega et al. (2002)) but also to thermodynamic systems where the storage function is the entropy instead of the energy (Alonso et al. (2002), Otero-Muras et al. (2006)). Nevertheless energy-balancing passivity-based control can only be applied to systems without pervasive dissipation, i.e. systems where the power provided by the controller is equal to zero at the desired equilibrium point. To overcome this difficulty the concept of power shaping control was introduced firstly for the stabilization of nonlinear RLC circuits (Ortega et al. (2003)). Contrary to energy-balancing passivity-based control, the storage function used for the control is related to the power and not to the energy. Power-shaping control has subsequently been applied to the control of mechanical and electro-mechanical systems (Garcia-Canseco et al. (2008)). Power-shaping control is based on a particular formulation of the system dynamics, namely the Brayton-Moser equations (Brayton and Moser (1964a), Brayton and Moser (1964b)). Although the first systems which have been described using this formalism are electrical circuits, it is shown in Jeltsema and Scherpen (2003), Jeltsema and Scherpen (2007) and Garcia-Canseco et al. (2008) that mechanical systems can also been expressed in this form.

As the work of Alonso, Ydstie and coworkers (see e.g. Alonso et al. (2002), Antelo et al. (2007), Farschman et al. (1998)), the present research is basically motivated by the objective to connect thermodynamics with process control design (see also Favache and Dochain (2009)). In this paper we apply power-shaping control to the exothermic CSTR case study with the aim of bringing more physical insight in its dynamical behaviour. After a brief presentation of

⋆ The work of A. Favache is funded by a grant of aspirant of Fonds National de la Recherche Scientifique (Belgium).

the main principles of power-shaping control in Section 2 and of the CSTR case study in Section 3, we shall apply the power-shaping control theory to our example. First we shall use the power-shaping approach to analyze the open-loop behaviour (Section 4) and then to design a control action (Section 5). Finally Section 6 presents some general comments on the possibility of extending the power-shaping approach to more complex systems, namely systems with more than one reactant, and/or more than one reaction.

2. POWER-SHAPING CONTROL

In this section, we briefly explain the principles of power-shaping control. The statements are given without any proof. For more details, the reader can refer to Jeltsema and Scherpen (2003), Ortega et al. (2003), Jeltsema and Scherpen (2007), Garcia-Canseco et al. (2008).

2.1 The Brayton-Moser formulation

Let us consider a dynamical system of dimension n with m inputs. The state of the system is given by the vector $x \in \mathbb{R}^n$ and the input is given by vector $u \in \mathbb{R}^m$. The power-shaping control theory is based on the Brayton-Moser formulation of the system dynamics (Brayton and Moser (1964a), Brayton and Moser (1964b)). In this formulation the system dynamics are of the following form:

$$Q(x)\frac{dx}{dt} = \nabla P(x) + G(x)u \tag{1}$$

where $Q(x) : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$ is a non-singular square matrix, $P(x) : \mathbb{R}^n \to \mathbb{R}$ is a scalar function of the state and $G(x) : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$. Additionally the symmetric part of the matrix $Q(x)$ is negative semi-definite, i.e.:

$$Q(x) + Q^t(x) \preceq 0 \tag{2}$$

The function $P(x)$ is called the potential function. In electrical and mechanical systems, the potential function has the units of power. In electrical systems it is related to the so-called content and co-content of the resistances (Ortega et al. (2003), Jeltsema and Scherpen (2007)) while it is related to the Rayleigh dissipation function (Jeltsema and Scherpen (2003)) in mechanical systems. In both cases, the potential function $P(x)$ is related to the dissipated power in the system.

Let us now assume that the system dynamics are given by the following relation:

$$\frac{dx}{dt} = f(x) + g(x)u \tag{3}$$

where $f(x) : \mathbb{R}^n \to \mathbb{R}^n$ and $g(x) : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^m$. The system (3) can be written in the form (1) if there exists a non-singular matrix $Q(x)$ fulfilling (2) and that solves following partial differential equation:

$$\nabla(Q(x)f(x)) = \nabla^t(Q(x)f(x)) \tag{4}$$

This condition is equivalent to the existence of the potential function $P(x)$ (i.e. if the potential function $P(x)$ exists, its Jacobian matrix must be symmetric). This one is the solution of the following partial differential equation system:

$$\nabla P(x) = Q(x)f(x) \tag{5}$$

Finally the function $G(x)$ is given by the following relation:

$$G(x) = Q(x)g(x)$$

2.2 Power-shaping control

Let us assume that the system dynamics can be expressed using the Brayton-Moser equations presented above. The desired equilibrium state is denoted by x^*. The rationale of power-shaping control is to choose the input $u(x)$ such that in closed loop the system dynamics are given by the following relation:

$$Q(x)\frac{dx}{dt} = \nabla P_d(x)$$

where $P_d(x) : \mathbb{R}^n \to \mathbb{R}$ is the reshaped potential function. The desired equilibrium point x^* must be a local minimum of the potential function $P_d(x)$ in order to be locally asymptotically stable. The function $P_d(x)$ cannot be arbitrarily chosen since the following relation has to be fulfilled:

$$g^\perp(x)Q^{-1}(x)\nabla P_a(x) = 0 \tag{6}$$

where $g^\perp(x) : \mathbb{R}^n \to \mathbb{R}^{n-m} \times \mathbb{R}^n$ is a full-rank left annihilator of $g(x)$ (i.e. $g^\perp(x)g(x) = 0$ with $\operatorname{rank}\left(g^\perp(x)\right) = n - m$) and $P_a(x) = P_d(x) - P(x)$. The condition (6) ensures the existence of a function $u(x)$ such that:

$$Q(x)\left(f(x) + g(x)u(x)\right) = \nabla P(x) + \nabla P_a(x)$$

Under these conditions, the control input $u(x)$ that achieves to reshape $P(x)$ into $P_d(x)$ is the following one:

$$u(x) = \left(G^t(x)G(x)\right)^{-1} G^t(x)\nabla P_a(x) \tag{7}$$

where $G(x) = Q(x)g(x)$

3. THE EXOTHERMIC CONTINUOUS STIRRED TANK REACTOR (CSTR)

In our research, we have applied the power-shaping methodology to a classical process control case study: the exothermic continuous stirred tank reactor (CSTR), illustrated in Figure 1. The reaction that is taking place is $A \to B$. In order to simplify the model, the following assumptions have been considered:

- the reactor is liquid phase and the volume V is constant.
- the density ρ and the specific heat c_p of the mixture are constant (i.e. independent of the temperature or of the composition).
- the reaction heat is independent of the temperature.
- the reaction is irreversible.
- the reaction kinetics obey to the mass action law, i.e. $r = k(T)n_A$ where $k(T)$ is the kinetic constant, depending only on the temperature T and n_A is the number of moles of component A. The function $k(T)$ is assumed to be monotonically increasing. Moreover let us assume that [1] $\lim_{T\to 0} k(T) = 0$, $\lim_{T\to\infty} k(T) = k_0$ and
$$\lim_{T\to 0}\frac{dk}{dT} = \lim_{T\to\infty}\frac{dk}{dT} = 0$$
- the dynamics of the jacket can be neglected.

[1] These assumptions on $k(T)$ are for instance fulfilled by the commonly used Arrhenius law.

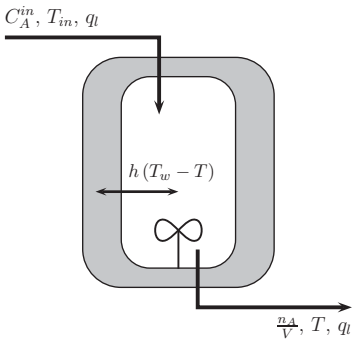


Fig. 1. Schematic view of the CSTR

- the heat exchange between the reactor and the jacket is proportional to the temperature difference between them, with h the heat exchange coefficient.
- the system is controlled by the cooling fluid flow rate. It acts directly on the heat transfer coefficient h. Therefore we shall consider in the sequel that the control input is the quantity $\frac{h}{\rho c_p V}$.

Under these assumptions the dynamic model of the system is given by following equations:

$$\begin{cases} \dfrac{dn_A}{dt} = \dfrac{q_l}{V}\left(C_A^{in}V - n_A\right) - k\left(T\right)n_A \\ \dfrac{dT}{dt} = \dfrac{q_l}{V}\left(T_{in} - T\right) + \dfrac{\left(-\Delta_r H\right)}{\rho c_p V}k\left(T\right)n_A \\ \qquad\qquad + \dfrac{h}{\rho c_p V}\left(T_w - T\right) \end{cases} \tag{8}$$

where q_l is the volumetric inlet and outlet flow rate, C_A^{in} is the inlet concentration of A, T_{in} is the inlet temperature, $(-\Delta_r H)$ is the reaction heat and T_w is the temperature of the cooling fluid. Using the notations of (3), we have:

$$f\left(x\right) = \begin{pmatrix} \delta\left(C_A^{in}V - n_A\right) - k\left(T\right)n_A \\ \delta\left(T_{in} - T\right) + \gamma k\left(T\right)n_A \end{pmatrix}$$

$$g\left(x\right) = \begin{pmatrix} 0 \\ T_w - T \end{pmatrix}$$

with $x = [n_A, T]^t$, $u = \frac{h}{\rho c_p V}$, $\delta = \frac{q_l}{V}$ and $\gamma = \frac{(-\Delta_r H)}{\rho c_p V}$.

It can be shown that this system can have up to three equilibrium points in open loop (i.e. for $u = 0$), depending on the numerical values of the parameters (see e.g. Aris and Amundson (1958), Uppal et al. (1974)). Here we shall consider the case with three equilibrium states, two being stable and one being unstable.

4. THE OPEN LOOP BEHAVIOUR

The methodology described in Section 2.1 has been applied to the system described in Section 3. The first step was to find the square matrix $Q\left(x\right)$ that meets the required properties (2) and (4). The partial differential equation (4) for the open-loop CSTR is written as follows:

$$\begin{aligned} &- q_{11}n_A\frac{dk}{dT} - q_{12}\delta + q_{12}\gamma n_A\frac{dk}{dT} \\ &+ \frac{\partial q_{11}}{\partial T}\left(\delta\left(C_A^{in}V - n_A\right) - kn_A\right) \\ &+ \frac{\partial q_{12}}{\partial T}\left(\delta\left(T_{in} - T\right) + \gamma k n_A\right) \\ &= -q_{21}\delta - q_{21}k + q_{22}\gamma k + \frac{\partial q_{21}}{\partial n_A}\left(\delta\left(C_A^{in}V - n_A\right) - kn_A\right) \\ &+ \frac{\partial q_{22}}{\partial n_A}\left(\delta\left(T_{in} - T\right) + \gamma k n_A\right) \end{aligned} \tag{9}$$

where q_{ij} is the entry in position (i, j) of the matrix $Q\left(n_A, T\right)$. We first transformed (9) into an algebraic equation by restricting ourselves to a subset of possible matrices $Q\left(n_A, T\right)$. This algebraic equation then has been solved and a possible matrix has been found [2]. In our case, the symmetric part of matrix $Q\left(x\right)$ was found to be definite negative.

Remark 1. Indeed a family of possible matrices $Q\left(n_A, T\right)$ has been found that both satisfy (2) and (4). But since they all have a similar form apart from a constant parameter, we shall treat them in the sequel as one unique matrix.

Next the potential function is found by integrating (5). The general form of the potential form is given by the following expression:

$$\begin{aligned} P\left(n_A, T\right) = &\int p\left(T\right)dT \\ &+ \omega\left[\gamma\left(C_A^{in}V - n_A\right) + \left(T^{in} - T\right)\right]^2 \end{aligned} \tag{10}$$

where ω is a positive constant and $p\left(T\right) : \mathbb{R} \to \mathbb{R}$ is a non-linear function of $k\left(T\right)$ and T. The quadratic term of (10) is clearly linked to the convection phenomena, whereas the integral term is related to the reaction kinetics.

Remark 2. Actually the function $p\left(T\right)$ is not unique. It depends directly on the matrix $Q\left(x\right)$. Since we have found a family of matrices $Q\left(x\right)$, there is a corresponding family of functions $p\left(T\right)$ that are similar apart from a constant parameter.

Let us now consider the equilibrium points $\bar{x} = \left(\bar{n}_A, \bar{T}\right)$ of the open-loop CSTR:

$$Q\left(\bar{x}\right)\left.\frac{dx}{dt}\right|_{\bar{x}} = \nabla P\left(\bar{x}\right) = 0$$

Since the matrix $Q\left(x\right)$ is non-singular, the equilibrium points $\bar{x}$ are also critical points of the potential function $P\left(x\right)$ (i.e. $\nabla P\left(\bar{x}\right) = 0$) and conversely. The analysis of the Hessian matrix of the obtained potential function (10) at each of the equilibrium points shows that the stable ones are local minima of the function $P\left(x\right)$ whereas the unstable one is a saddle point. The level curves of the function $P\left(x\right)$ are given in Figure 2.

The variation of function $P\left(x\right)$ along the trajectories of the system are given by the following relation:

[2] Calculation details can be found in Favache and Dochain (2008).

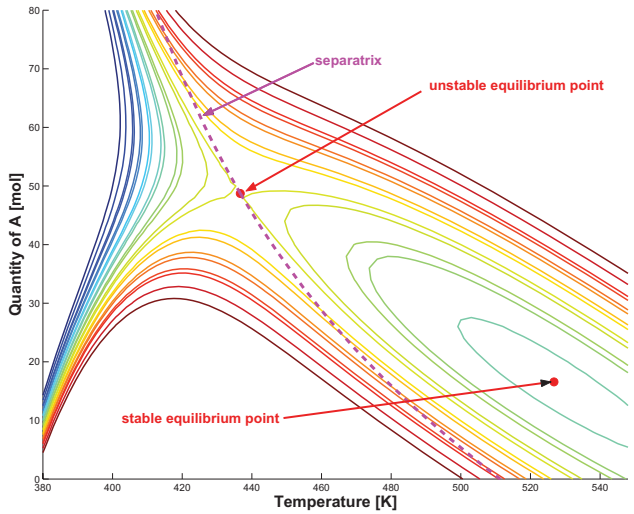


Fig. 2. Level curves of the potential function

$$\begin{aligned}\frac{dP}{dt} &= \nabla^t P(x) f(x) \\ &= \nabla^t P(x) Q^{-1}(x) \nabla P(x) \\ &= \frac{1}{2}\nabla^t P(x) \left(Q^{-1}(x) + \left(Q^t(x)\right)^{-1}\right) \nabla P(x)\end{aligned}$$

Since the symmetric part of $Q(n_A, T)$ is negative definite, we have $\frac{dP}{dt} \leq 0$ where the equality holds only for $\nabla P(x) = 0$, i.e. for the equilibrium points. As a consequence $P(n_A, T)$ is decreasing along the trajectories. Moreover $P(x)$ is radially unbounded. Therefore the potential function $P(n_A, T)$ is a global Lyapunov function for the system (Khalil (2002)).

5. POWER-SHAPING CONTROL OF THE CSTR

5.1 Controller design

Let us denote by $x^* = [n_A^*, T^*]^t$ the desired equilibrium state. Since the input has an effect only on the temperature dynamics, the values of n_A^* and T^* cannot be chosen independently. This is stated in the following lemma:

Lemma 1. Let us consider some control input $u(x)$ such that the closed loop system has at least one equilibrium point. The equilibrium points of the closed loop system are contained in the following set:

$$(n_A^*, T^*) \in \left\{ (n_A, T) \,\middle|\, \begin{array}{c} n_A = \dfrac{\delta C_A^{in} V}{k(T) + \delta} \\ T \in]0, +\infty[\setminus \{T_w\} \end{array} \right\} \tag{11}$$

Proof. Let $x^* = (n_A^*, T^*)$ be an equilibrium point of the closed loop system for the control input $u(x)$. By definition of an equilibrium point, we have:

$$\left\{ \begin{array}{rl} 0 = & \dfrac{q_l}{V}\left(C_A^{in} V - n_A^*\right) - k(T^*) n_A^* \\ 0 = & \dfrac{q_l}{V}(T_{in} - T^*) + \dfrac{(-\Delta_r H)}{\rho c_p V} k(T^*) n_A^* \\ & + u(x^*)(T_w - T^*) \end{array} \right. \tag{12}$$

The first equation can be directly rewritten as follows:

$$n_A^* = \frac{\delta C_A^{in} V}{k(T^*) + \delta}$$

Let us now assume that $T^* = T_w$ is a possible equilibrium of the open loop system. If we replace in the second equation of (12), we find the following relation:

$$0 = \frac{q_l}{V}(T_{in} - T_w) + \frac{(-\Delta_r H)}{\rho c_p V} k(T_w) \frac{\delta C_A^{in} V}{k(T_w) + \delta}$$

This relation implies that T_w is an equilibrium point of the open-loop system. But, except an unlikely particular case, there is no reason that the temperature of the cooling fluid is exactly equal to an equilibrium temperature of the open-loop CSTR. Hence T_w cannot be an equilibrium temperature of the closed-loop system. □

Lemma 2. (11) is a necessary condition for $x^* = (n_A^*, T^*)$ to be a local minimum of $P_d(x)$.

Proof. A necessary condition for x^* to be a local minimum of $P_d(x)$ is the following one:

$$\nabla P_d(x^*) = \nabla P(x^*) + \nabla P_a(x^*) = 0$$

By replacing $\nabla P(x)$ and $\nabla P_a(x)$ by their expressions, the following relation is obtained:

$$\begin{aligned} -\nabla P(x^*) &= -Q(x^*) f(x^*) \\ &= \nabla P_a(x^*) = Q(x^*) g(x^*) u(x^*) \end{aligned}$$

Since $Q(x^*)$ is invertible, this can be rewritten as :

$$f(x^*) + g(x^*) u(x^*) = 0 \tag{13}$$

When replacing $f(x)$ and $g(x)$ by their respective expression, (12) is obtained. Thus the rest of the proof of Lemma 1 also applies here. □

The control action is then found by applying the methodology described in Section 2.2. First we solve (6) using the previously found expression for the matrix $Q(x)$. The left annihilators of $g(x)$ are given as follows:

$$g^{\perp}(x) = [\varphi \quad 0]$$

with $\varphi \in \mathbb{R}^*$. (6) is thus written as follows:

$$\frac{\varphi}{\det Q(x)} \left(q_{22}(n_A, T) \frac{\partial P_a}{\partial n_A} - q_{12}(n_A, T) \frac{\partial P_a}{\partial T} \right) = 0$$

where $q_{12}(n_A, T)$ and $q_{22}(n_A, T)$ are the elements of $Q(x)$ in position $(1,2)$ and $(2,2)$, respectively. With our matrix $Q(x)$ the solution of this partial differential equation is given as follows:

$$P_a(n_A, T) = f_a\left(n_A + \int w(T)\, dT\right)$$

where $w(k(T))$ is a rational function of the kinetic coefficient $k(T)$. $f_a(z) : \mathbb{R} \to \mathbb{R}$ can be any smooth real-valued function. In our case we have chosen $f_a(z)$ to be a second order polynomial. Using now (7), the control action of the following form is obtained for stabilizing the desired equilibrium point:

$$u(n_A, T) = \frac{-\mu(n_A - n_A^* + W_{T^*}(T)) + u^*(T_w - T^*)}{T_w - T} \tag{14}$$

where $W_{T^*}(T) : \mathbb{R}^{*+} \to \mathbb{R}$ is given by the following expression:

$$W_{T^*}(T) = \int_{T^*}^{T} w(k(\tau))\, d\tau$$

Obviously $W_{T^*}(T^*) = 0$. $\mu \in \mathbb{R}$ is a parameter of the controller and u^* is the value of the input at the equilibrium state:

$$u^*(T_w - T^*) = -\delta(T_{in} - T) - \gamma\delta\left(C_A^{in}V - n_A^*\right)$$

In order to ensure that x^* is a local minimum of $P_d(x^*)$, μ has to be lower bounded. The bounds are obtained by imposing that the Jacobian matrix of $P_d(x^*)$ is positive definite at the desired closed-loop equilibrium point. The control action $u(x)$ acts in (8) via the term

$$g(x)\,u(x) = \begin{pmatrix} 0 \\ -\mu\left(n_A - n_A^* + W_{T^*}(T)\right) + u^*(T_w - T^*) \end{pmatrix}$$

which does not depend on $(T_w - T)$ anymore. This means that the actual control input is the transferred heat which is equal to:

$$-\mu\left(n_A - n_A^* + W_{T^*}(T)\right) + u^*(T_w - T^*)$$

5.2 Considerations on local and global convergence

The controller designed in the previous section only guarantees a local convergence to the desired set point, i.e. the system will converge to desired point if and only if the initial conditions are close enough to it. Global convergence is ensured if and only if the shaped potential function $P_d(x)$ does not have other local minima, i.e. if and only if x^* is a global minimum of the function $P_d(x)$.

Let us assume that the point $x^\# = \left[n_A^\#, T^\#\right]^t$ is another local minimum of the function $P_d(x)$. For the same reasons as stated before, the following relation has to be fulfilled:

$$n_A^\# = \frac{\delta C_A^{in} V}{k(T^\#) + \delta} \tag{15}$$

Moreover, for $x^\#$ to be a closed loop equilibrium of (8), the following relation has to be fulfilled:

$$\delta\left(T_{in} - T^\#\right) + \delta\gamma\left(C_A^{in}V - n_A^\#\right) + u^*(T_w - T^*) - \mu\left(n_A^\# - n_A^* + W_{T^*}\left(T^\#\right)\right) = 0 \tag{16}$$

By analogy with (15), let us define the function $\tilde{T}^\#(n_A)$ by the following implicit relation:

$$k\left(\tilde{T}^\#(n_A)\right) = \frac{\delta C_A^{in} V}{n_A} - \delta$$

Consider the following function:

$$\Delta(n_A) = \delta\left(T_{in} - \tilde{T}^\#\right) + \delta\gamma\left(C_A^{in}V - n_A\right) + u^*(T_w - T^*) - \mu\left(n_A - n_A^* + W_{T^*}\left(\tilde{T}^\#\right)\right)$$

where the dependence of $\tilde{T}^\#(n_A)$ on n_A has been omitted for sake of clarity. The equilibrium points x^* and $x^\#$ are zeros of the function $\Delta(n_A)$. Therefore x^* is the unique convergence point if it is the unique zero of $\Delta(n_A)$.

The function $\Delta(n_A)$ is linear in the parameter μ. Therefore it can be written as follows:

$$\Delta(n_A) = \Delta_0(n_A) + \mu\Delta_\mu(n_A)$$

Remark 3. In Favache and Dochain (2009), we have presented several old and new results that aim at linking the thermodynamics and the system theory concepts via the CSTR study case[3] In this paper we have introduced a function of the state, denoted $\Delta_{eq}(T)$. By integrating the notations and assumptions of the present paper, this function is given by the following expression:

$$\Delta_{eq}(T) = k(T)\,n_A V(-\Delta_r H) - \rho c_p V\left[u(n_A, T)(T - T_w) + \delta\left(T - T^{in}\right)\right]$$

It can be seen directly that $\Delta_{eq}(T)$ and $\Delta(n_A)$ are linked by the following relation:

$$\Delta(n_A) = \frac{\Delta_{eq}\left(\tilde{T}^\#(n_A)\right)}{\rho c_p V}$$

It can be shown that the assumptions on the form of $k(T)$ described in Section 3 and the existence of three open loop equilibria implies that $\Delta_0(n_A)$ is increasing for low and high temperatures, but decreasing on one determined interval[4] As a consequence, $\Delta(n_A)$ can have several zeros, depending on the term $\mu\Delta_\mu(n_A)$.

If the function $w(T)$ has been adequately chosen, then $\Delta_\mu(n_A) > 0$ for $n_A > n_A^*$ and $\Delta_\mu(n_A) < 0$ for $n_A < n_A^*$. This means that if μ is chosen sufficiently large and positive, then the term $\mu\Delta_\mu(n_A)$ reshapes the initial function $\Delta_0(n_A)$ so as to make the two undesired zeros to vanish (see Figure 3). As a conclusion, there is a lower bound on the parameter μ in order to ensure global convergence.

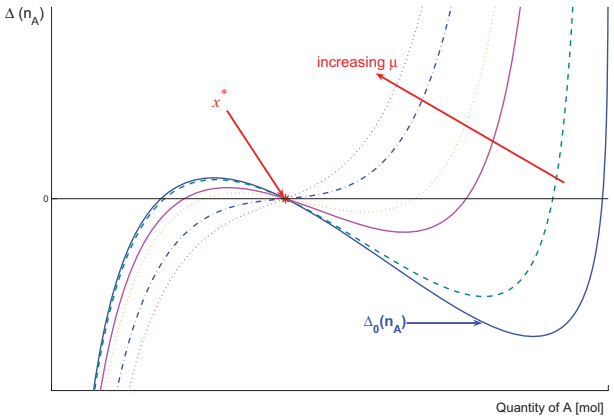


Fig. 3. Influence of μ on $\Delta(n_A)$

5.3 Simulation results

In this section we present some simulation results of the controlled system. First the performance of the controller for reference tracking is shown in Figure 4. Then Figures 5 and 7 show the cases where only local convergence and global convergence, respectively, to the desired equilibrium point is ensured. These figures show the temperature evolution for different initial conditions. The corresponding level curves of the potential function $P_d(x)$ are shown for

[3] In Favache and Dochain (2009) we have considered a CSTR with a reversible reaction, but the results can be applied directly for an irreversible reaction by setting the kinetic reaction coefficient of the reverse reaction equal to zero.

[4] This can be deduced from the form of $\Delta_{eq}(T)$ in Favache and Dochain (2009) using the fact that $\tilde{T}^\#(n_A)$ is a strictly decreasing function.

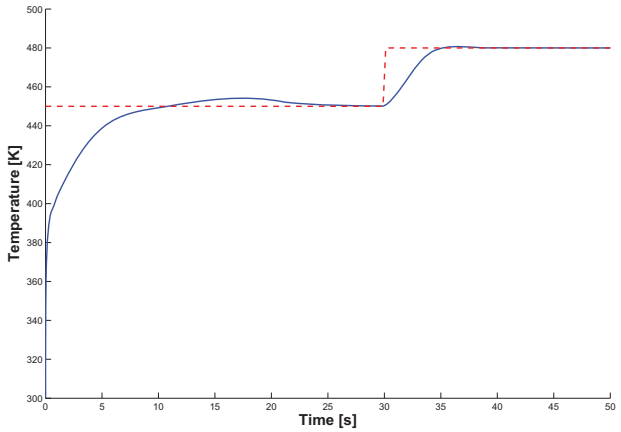


Fig. 4. Reference tracking

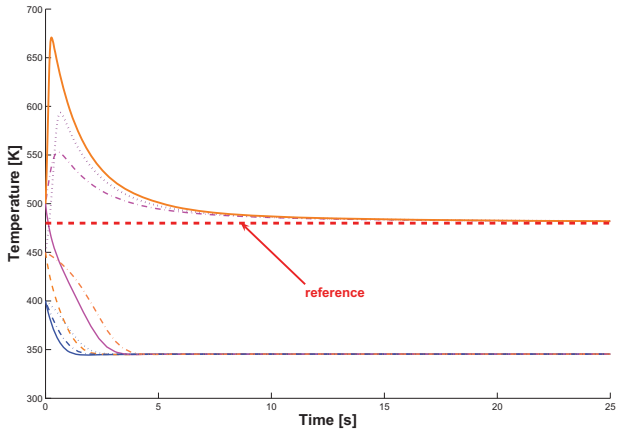


Fig. 5. Temperature evolution for different initial conditions (local convergence)

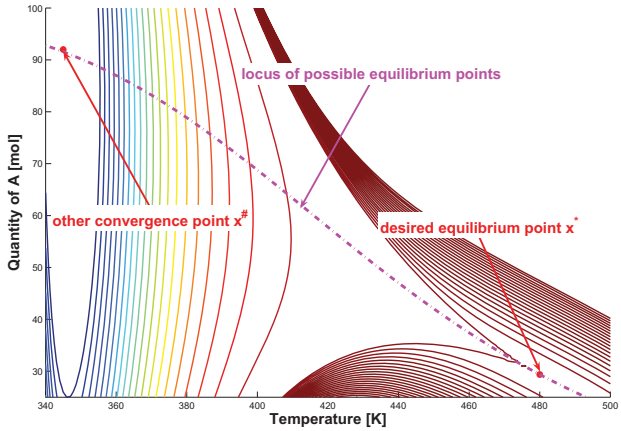


Fig. 6. Level curves of $P_d(x)$ (local convergence)

both cases in Figures 6 and 8, respectively. It can be clearly seen on Figure 5 that there exist two convergence points, depending on the initial condition. This is confirmed by the level curves of the function $P_d(x)$ in Figure 6 where two local minima can be distinguished.

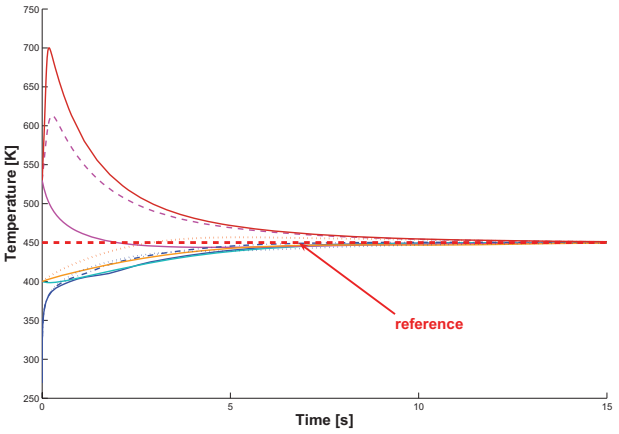


Fig. 7. Temperature evolution for different initial conditions (global convergence)

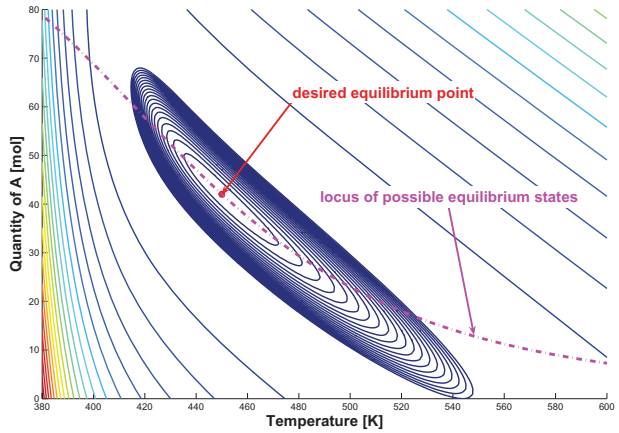


Fig. 8. Level curves of $P_d(x)$ (global convergence)

5.4 Robustness analysis

The control law given in (14) requires the complete state feedback. Moreover it also requires the knowledge of the temperature dependence of the kinetic coefficient $k(T)$ that appears in the expression of $W_{T^*}(T)$. In most practical applications, the kinetic coefficient is determined experimentally and the on-line measurement of the concentration is not always achievable. In this section we shall analyze the effect on the closed-loop convergence and stability of modeling errors in the kinetic coefficient $k(T)$.

Let us assume that there is a modeling error on the kinetic coefficient. The aim is to stabilize the state (n_A^*, T^*). The controller is designed using the function $\hat{k}(T)$ instead of the real kinetic coefficient $k(T)$:

$$\hat{k}(T) = (1 + \xi(T))\, k(T) \tag{17}$$

with $\xi(T) > -1$.

Assumption 1. Despite the error on $k(T)$, the equilibrium value of n_A is known:

$$n_A^* = n_A^{eq}(T^*) = \delta \frac{C_A^{in} V}{\delta + k(T)} \neq \delta \frac{C_A^{in} V}{\delta + \hat{k}(T)}$$

Assumption 2. Despite the error on $k(T)$, the equilibrium value of the control input u^* is known:

$$u_c^* = -\delta\left(T_{in} - T^*\right) - \gamma k\left(T^*\right) n_A^*$$
$$\neq -\delta\left(T_{in} - T^*\right) - \gamma \hat{k}\left(T^*\right) n_A^*$$

Assumption 3. The control input has been designed such that, based on the estimated value of the kinetic coefficient, the desired equilibrium is asymptotically stable.

From (14), the control input applied to the system is given by the following expression:

$$u_c\left(T_w - T\right) = \mu\left(n_A^* - n_A - \int_{T^*}^{T} \hat{w}\left(\tau\right) d\tau\right) + u_c^*\left(T_w - T^*\right)$$

where we have introduced the following notation, for the sake of clarity: $\hat{w}\left(T\right) = w\left(\hat{k}\left(T\right)\right)$. Assumptions 1 and 2 imply that $\left(n_A^*, T^*\right)$ is still an equilibrium of the closed-loop system. Assumption 3 implies that the function $w\left(y\right)$ and the parameter μ have been chosen such that the following matrix is negative definite:

$$\hat{\Lambda} = \begin{pmatrix} -\left(\delta + \hat{k}^*\right) & -n_A^* \left.\dfrac{d\hat{k}}{dT}\right|_{T^*} \\ \gamma\hat{k}^* - \mu & -\delta + \gamma n_A^* \left.\dfrac{d\hat{k}}{dT}\right|_{T^*} - \mu\hat{w}\left(T\right) \end{pmatrix}$$

$\hat{\Lambda}$ is the matrix of the linearized system around the desired equilibrium state if the kinetics was indeed equal to $\hat{k}\left(T\right)$. As a consequence we have:

$$\operatorname{tr}\hat{\Lambda} < 0 \quad \text{and} \quad \det\hat{\Lambda} > 0$$

The actual matrix of the linearized system around $\left(n_A^*, T^*\right)$ is written as follows:

$$\Lambda = \begin{pmatrix} -\left(\delta + k^*\right) & -n_A^* \left.\dfrac{dk}{dT}\right|_{T^*} \\ \gamma k^* - \mu & -\delta + \gamma n_A^* \left.\dfrac{dk}{dT}\right|_{T^*} - \mu\hat{w}\left(T\right) \end{pmatrix}$$

The trace and the determinant of Λ are given by the following relations:

$$\operatorname{tr}\Lambda = \Psi^* - \delta - \mu\hat{w}\left(T^*\right)$$
$$\det\Lambda = -\delta\Psi^* + \mu\left[\left(\delta + k^*\right)\hat{w}\left(T^*\right) - n_A \left.\frac{dk}{dT}\right|_{T^*}\right]$$

with

$$\Psi^* = -\left(k^* + \delta\right) + \left.\frac{dk}{dT}\right|_{T^*} \gamma n_A^* \tag{18}$$

Using (17), we can define the following quantity $\hat{\Psi}^*$ by analogy with (18):

$$\hat{\Psi}^* = \gamma n_A^* \left.\frac{d\hat{k}}{dT}\right|_{T^*} - \left(\hat{k}^* + \delta\right)$$
$$= \Psi^*\left(1 + \xi^*\right) + \gamma n_A^* k^* \left.\frac{d\xi}{dT}\right|_{T^*} + \delta\xi^*$$

where $\xi^* = \xi\left(T^*\right)$. As a consequence the trace and the determinant of $\Lambda^{(cl)}$ can be rewritten as follows:

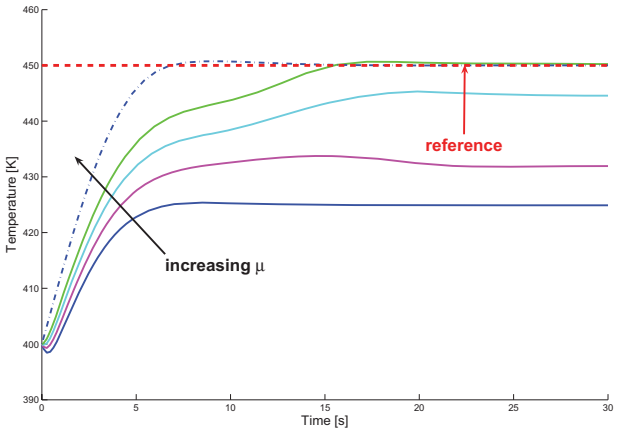


Fig. 9. Influence of the modelling error on $k\left(T\right)$ for different values of μ

$$\left(1 + \xi^*\right)\operatorname{tr}\Lambda = \underbrace{\hat{\Psi}^* - \delta - \mu\hat{w}\left(T^*\right)}_{\operatorname{tr}\hat{\Lambda}^{(cl)} < 0} - \left[\left.\frac{d\xi}{dT}\right|_{T^*} \gamma k^* n_A^* + 2\xi^*\delta + \mu\xi^*\hat{w}\left(T^*\right)\right]$$

By applying the same development on the determinant, the following expression is obtained:

$$\left(1 + \xi^*\right)\det\Lambda = \det\hat{\Lambda} + \xi^*\delta\left(\delta + \mu\hat{w}\left(T^*\right)\right) + n_A^* k^* \left.\frac{d\xi}{dT}\right|_{T*}\left(\gamma\delta + \mu\right)$$

The closed-loop equilibrium is asymptotically stable if and only if the trace is strictly negative and the determinant is strictly positive. Using the inequalities of Assumption 3, this means that if $\xi^* > 0$ and $\left.\frac{d\xi}{dT}\right|_{T^*} \geq 0$, the closed-loop equilibrium is asymptotically stable. But if these two conditions are not fulfilled, then the closed-loop equilibrium can become unstable. Nevertheless, if the function $w\left(y\right)$ has been taken sufficiently large such that:

$$\left(\delta + k^*\right)\hat{w}\left(T^*\right) - n_A \left.\frac{dk}{dT}\right|_{T^*} > 0$$

then $\det\Lambda^{(cl)}$ is increasing with μ such that there is a lower bound on μ that ensures the local asymptotic convergence (see Figure 9).

6. EXTENSION TO MORE COMPLEX SYSTEMS

The power-shaping approach has given interesting results on the simplified CSTR case study. It is therefore of major interest to see if this approach can be extended to more complex systems, and more particularly to systems with more than one reacting chemical species and/or with multiple reactions.

We keep the same assumptions as before, but we consider more complex kinetics. Under these assumptions, the general form of the dynamics of a non-isothermal CSTR can be deduced from (8).

Let us now consider a CSTR with N_r independent reactions. From all the chemical species in the mixture, it is only necessary to consider those that intervene in the

kinetics of the reactions. Let N_c be the number of chemical species which intervene in the expression of the reaction kinetics. For each species i the time evolution is given by the following differential equation (with $i = 1, \ldots, N_c$):

$$\frac{dn_i}{dt} = \delta\left(C_i^{in}V - n_i\right) + \sum_{l=1}^{N_r} \Gamma_{il} r_l\left(T, n\right) \tag{19}$$

where Γ_{il} is the stoichiometric coefficient of species i in the reaction l. The temperature dynamical equation becomes as follows:

$$\frac{dT}{dt} = \delta\left(T^{in} - T\right) + \sum_{l=1}^{N_r} \gamma_l r_l\left(T, n\right)$$

where $\gamma_l = \frac{(-\Delta_r H)_l}{\rho c_p V}$ and $(-\Delta_r H)_l$ is the reaction heat of the reaction l. The time evolution equations of n_A and T have a similar form and the dynamics of the system can be written as follows (see Dochain et al. (1992)):

$$\frac{dx}{dt} = \delta\left(x^{in} - x\right) + \Gamma r\left(x\right) \tag{20}$$

where $x = [n_A, \ldots, n_{N_c}, T]^t$, . $\Gamma \in \mathbb{R}^{N_c+1} \times \mathbb{R}^{N_r}$ is a matrix that contains the stoichiometric coefficient of species i in the reaction l in position (i, l) if $1 \leq i \leq N_c$ and that contains γ_l in the l^{th} column if $i = N_c + 1$.

We shall now apply the same approach as in Section 4, i.e. we shall first look for a matrix $Q\left(x\right)$ that fulfills the conditions (2) and (4). By using (20), we have the following relations $(i, j = 1, \ldots, N_c + 1)$:

$$\left(Q\left(x\right) f\left(x\right)\right)_i = \sum_{k=1}^{N_c+1} q_{ik}\left(\delta\left(x_k^{in} - x_k\right) + \Gamma_{kl} r_l\left(x\right)\right)$$

and thus:

$$\begin{aligned}\left(\nabla\left(Q\left(x\right) f\left(x\right)\right)\right)_{ij} = -\delta q_{ij} &+ \sum_{k=1}^{N_c+1}\left[q_{ik}\sum_{l=1}^{N_r}\Gamma_{kl}\frac{\partial r_l}{\partial x_j}\right] \\ &+ \sum_{k=1}^{N_c+1}\frac{\partial q_{ik}}{\partial x_j}\left(\delta\left(x_k^{in} - x_k\right) + \Gamma_{kl} r_l\left(x\right)\right)\end{aligned}$$

Consequently, condition (4) can be rewritten as follows:

$$\begin{aligned}&-\delta q_{ij} + \sum_{k=1}^{N_c+1}\left[q_{ik}\sum_{l=1}^{N_r}\Gamma_{kl}\frac{\partial r_l}{\partial x_j}\right] \\ &+ \sum_{k=1}^{N_c+1}\frac{\partial q_{ik}}{\partial x_j}\left(\delta\left(x_k^{in} - x_k\right) + \Gamma_{kl} r_l\left(x\right)\right) \\ &= -\delta q_{ji} + \sum_{k=1}^{N_c+1}\left[q_{jk}\sum_{l=1}^{N_r}\Gamma_{kl}\frac{\partial r_l}{\partial x_i}\right] \\ &+ \sum_{k=1}^{N_c+1}\frac{\partial q_{jk}}{\partial x_i}\left(\delta\left(x_k^{in} - x_k\right) + \Gamma_{kl} r_l\left(x\right)\right)\end{aligned} \tag{21}$$

for all $i, j = 1, \ldots, N_c + 1$.

In the simplified CSTR case, we have first restricted the set of possible matrices $Q(x)$ to find a solution for (4) by transforming the partial differential equation into an algebraic one which is simpler to solve. The same can be done in the more complex case by adequately restricting the set of possible solutions.

A general solution of the algebraic form of (21) has not been found yet. Despite of this, some characteristics of the solution (if it exists) have already been derived in Favache and Dochain (2008). We shall now look at some particular cases of more complex reactions to get a first intuition of the existence of non-singular and negative semi-definite solutions of the algebraic form of (21). The detailed solution for the three particular cases (namely parallel reactions, reactions with two reactants and consecutive reactions) is given in Favache and Dochain (2008). In the three cases it has been assumed that the kinetic functions can be expressed as powers of the concentration of the reactants, i.e.:

$$r_l\left(n, T\right) = k_l\left(T\right)\prod_{i\in\Theta_l}\left(n_i\right)^{\zeta_{il}}$$

where Θ_l is the set of reactants of reaction l and ζ_{il} is some positive constant.

6.1 Parallel reactions and reactions with two reactants

In this case a solution has been found, but the obtained matrix has not a negative semi-definite symmetric part. Thus the corresponding potential function cannot be used as a Lyapunov function for the open-loop system because they are not decreasing along the system trajectories.

As shown in Garcia-Canseco et al. (2008), given a matrix $Q\left(x\right)$ that fulfills (4), other solutions to (4) can be built departing from the first one. Applying this methodology on the matrices that have been found could lead to another matrix $Q\left(x\right)$ that would be negative-definite and hence give a Lyapunov function for the open-loop system.

Remark 4. In the particular case of parallel reactions with first order kinetics, another solution for the matrix $Q\left(x\right)$ has been found. This solution does not exist for higher order kinetics. This particular solution has a very similar form of that of the matrix proposed for the case with one reaction. This seems to indicate that the matrix $Q\left(n_A, T\right)$ that has been used in Section 4 is a particular solution of (9) that exists only in the case of first-order kinetics.

6.2 Consecutive reactions

In this case the algebraic form of (21) has no non-singular solution. This does not mean that the dynamics cannot be put into the Brayton-Moser form. But if the Brayton-Moser form exists, then the partial differential equation (21) has to be solved.

7. CONCLUSION

A general description of the power-shaping control approach of the CSTR has been given in this paper. The main results that were obtained by this approach have been presented and illustrated by some simulation results. A detailed mathematical analysis is provided in Favache and Dochain (2008). Contrary to previous works and approaches, a global Lyapunov function for the exothermic CSTR has been found using the power-shaping approach. This Lyapunov function could then be used to design a controller for stabilizing the reactor at a desired temperature. Some results about the local and global convergence of the controller have also been shown.

The Lyapunov function that was found is the potential function of the Brayton-Moser formulation of the CSTR

dynamics. Although the potential function has a physical meaning for the Brayton-Moser formulation of electrical or mechanical systems, a precise physical interpretation of the potential function of the CSTR has still to be found. This interpretation should give more physical insight on the reasons of the existence of multiple open loop equilibria, and also on the action of the controller.

The controller obtained by the power-shaping control approach depends on the on-line measurements of the concentration and of the knowledge on the reaction kinetics. Both quantities are usually not exactly known. We have shown, that if the parameters of the controller are adequately chosen, the control action is robust with respect to modeling errors on the kinetics. But a robustness study on the influence of the concentration measurement errors should also be of great interest before applying it to a real reactor.

Finally we studied the possibility of extending the previous work to more complex systems, and more precisely to CSTRs with multiple reactions and/or multiple reactants. The extension seems to be rather complex, even for simple cases such as two parallel reactions or two consecutive reactions with mass action law kinetics. Indeed, in the simple CSTR case with a single first-order kinetics reaction, the solution of the partial differential equation needed to write the system dynamics in the Brayton-Moser form has been found by transforming it into an algebraic one. For the three considered particular cases the corresponding algebraic equation system has either no non-singular solution (consecutive reactions) or an indefinite solution (parallel reactions, reaction with two reactants). The study presented here is only embryonic since it does not imply that the power-shaping control approach cannot been applied. But the Brayton-Moser form of the dynamics (if it exists) actually needs the solution of the partial differential equation, and not of its simplified version which is the algebraic equation. Nevertheless in the cases where an indefinite matrix has been found, the work presented in Garcia-Canseco et al. (2008) offers the possibility of finding an alternative negative semi-definite matrix without having to solve the partial differential equation. Also in this study we have only considered a particular form of the kinetics, namely we have assumed that the kinetics can be written as a product of powers of the concentration (i.e. a more general form of the mass action law) and a kinetic term. This is only a restricted class of the possible kinetic laws. Indeed it could also be interesting to apply the power-shaping control to other forms of kinetics such as the Monod kinetics (biological systems) or the Michaelis-Menten kinetics (enzymatic reactions), for instance.

ACKNOWLEDGEMENTS

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

REFERENCES

Alonso, A., Ydstie, B., and Banga, J. (2002). From irreversible thermodynamics to a robust control theory for distributed process systems. *Journal of Process Control*, 12, 507–517.

Antelo, L.T., Otero-Muras, I., Banga, J.R., and Alonso, A.A. (2007). A systematic approach to plant-wide control based on thermodynamics. *Computers & Chemical Engineering*, 31(5-6), 677–691.

Aris, R. and Amundson, N.R. (1958). An analysis of chemical reactor stability and control–I : The possibility of local control, with perfect or imperfect control mechanisms. *Chemical Engineering Science*, 7(3), 121–131.

Brayton, R. and Moser, J. (1964a). A theory of nonlinear networks I. *Quaterly of applied mathematics*, 22, 1–33.

Brayton, R. and Moser, J. (1964b). A theory of nonlinear networks II. *Quaterly of applied mathematics*, 22, 81–104.

Dalsmo, M. and van der Schaft, A. (1998). On representations and integrability of mathematical structures in energy-conserving physical systems. *Journal on Control and Optimization*, 37(1), 54–91.

Dochain, D., Perrier, M., and Ydstie, B. (1992). Asymptotic observers for stirred tank reactors. *Chemical Engineering Science*, 47, 4167–4178.

Eberard, D. (2006). *Extension des systèmes hamiltoniens ports aux systèmes irréversibles: une approche par la géométrie de contact*. Ph.D. thesis, Université Claude Bernard, Lyon 1.

Eberard, D., Maschke, B., and van der Schaft, A. (2005). Port contact systems for irreversible thermodynamical systems. In *Proc. 44^{th} IEEE Conference on Decision and Control and European Control Conference*. Sevilla (Spain).

Eberard, D., Maschke, B., and van der Schaft, A. (2006). Energy-conserving formulation of rlc-circuits with linear resistors. In *Proc. 17^{th} International Symposium on Mathematical Theory of Networks and systems*. Kyoto (Japan).

Farschman, C.A., Viswanath, K.P., and Ydstie, B.E. (1998). Process systems and inventory control. *AICHE Journal*, 44(8), 1841–1857.

Favache, A., Maschke, B., and Dochain, D. (2007). Contact structures: application to interconnected thermodynamical systems. In *Proc. European Control Conference 2007*. Kos (Greece).

Favache, A. and Dochain, D. (2008). Stability analysis of a CSTR via power-shaping control. Technical report, CESAME, Université catholique de Louvain, Belgium.

Favache, A. and Dochain, D. (2009). Thermodynamics and chemical systems stability: The CSTR case study revisited. *Journal of Process Control*, 19(3), 371–379.

Garcia-Canseco, E., Jeltsema, D., Scherpen, J., and Ortega, R. (2008). Power-based control of physical systems: two case studies. In *Proc. 17^{th} IFAC World Congress*. Seoul (Korea).

Jeltsema, D., Ortega, R., and M.A. Scherpen, J. (2004). An energy-balancing perspective of interconnection and damping assignment control of nonlinear systems. *Automatica*, 40(9), 1643–1646.

Jeltsema, D. and Scherpen, J.M. (2003). On mechanical mixed potential, content and co-content. In *Proc. European Control Conference*, 73–78.

Jeltsema, D. and Scherpen, J.M. (2007). A power-based description of standard mechanical systems. *Systems & Control Letters*, 56(5), 349–356.

Khalil, H. (2002). *Nonlinear Systems*. Prentice Hall, Upper Saddle River, 3^{rd} edition.

Maschke, B., Ortega, R., and van der Schaft, A. (2000). Energy-based Lyapunov functions for forced Hamiltonian systems with dissipation. *IEEE Transactions on Automatic Control*, 45(8), 1498–1502.

Maschke, B. and van der Schaft, A. (2005). From conservation laws to port-Hamiltonian formulation of distributed parameter systems. In *Proc. 16^{th} IFAC world congress*. Prague (Czech Republic).

Ortega, R., Jeltsema, D., and Scherpen, J. (2003). Power shaping: A new paradigm for stabilization of nonlinear RLC-circuits. *IEEE Transactions on Automatic Control*, 48(10), 1762–1767.

Ortega, R., van der Schaft, A., Maschke, B., and Escobar, G. (1999). Energy-shaping of port-controlled Hamiltonian systems by interconnection. In *Proc. 38^{th} IEEE Conference on Decision and Control*, 1646–1651. Phoenix (USA).

Ortega, R., Van Der Schaft, A., Mareels, I., and Maschke, B. (2001). Putting energy back in control. *IEEE Control Systems Magazine*, 21(2), 18–33.

Ortega, R., van der Schaft, A., Maschke, B., and G.Escobar (2002). Interconnection and damping assignment: passivity-based control of port-controlled Hamiltonian systems. *Automatica*, 38, 585–596.

Otero-Muras, I., Szederknyi, G., Alonso, A., and Hangos, K. (2006). Dynamic analysis and control of chemical and biochemical reaction networks. In *Proc. International Symposium on Advanced Control of Chemical Processes*, 165–170. Gramado (Brazil).

Uppal, A., Ray, W.H., and Poore, A.B. (1974). On the dynamic behavior of continuous stirred tank reactors. *Chemical Engineering Science*, 29(4), 967–985.

Treatment Planning of Cancer Dendritic Cell Therapy Using Multi-Objective Optimization

K. Lakshmi Kiran*, S. Lakshminarayanan*

**Department of Chemical and Biomolecular Engineering, National University of Singapore, 117576 Singapore (Tel: 0065-65168484; e-mail: chels@ nus.edu.sg).*

Abstract: Cancer immunotherapy is one of the emerging therapies for cancer treatment where immune cells are guided to fight against cancer. Clinical immunologists are proposing different ideas to stimulate the immune cells and dendritic cell therapy is one among them. Like, other treatment modalities, the challenge in dendritic cell therapy is when and how much dendritic cells should be administered. In this work we use a mathematical model which elucidates the activation of the helper T-cells and cytotoxic T-cells by the intervention of dendritic cells. The objective is to minimize the tumor cells for a given input of dendritic cells. Then multi-objective optimization is applied on the model to design the treatment planning in order to achieve the objective.

Keywords: Cancer, Immune system, Dendritic cell therapy, Mathematical model, Multi-objective optimization

1. INTRODUCTION

Cancer stands next only to heart disease in the list of most fatal diseases in the world. From Fig.1, it is obvious that the decrease in death rate for cancer patients over the years 1950-2003 has been minimal as compared to other major diseases. Cancer related deaths have been escalating meteorically - according to World Health Organization, 7.6 million people died of cancer (out of 58 million deaths overall) in 2005. They speculate that cancer deaths will increase to 18% and 50% by 2015 and 2030 respectively. Recently, the American Cancer Society reported that around 1.5 million new cancer cases and 0.6 million cancer death cases occurred in the US in 2007. According to another report on worldwide cancer rates by the WHO's International Agency for Research on Cancer (IARC) (Paola Pisani, 2002), North America leads the world in the rate of cancers diagnosed in adults, followed closely by Western Europe, Australia and New Zealand. In 1994, in Britain, almost one in three were expected to develop the disease over their survival period and it is estimated to increase to one in two by 2010 with reference to the trends at that time (Imperial Cancer Research Fund). Another publication from the Australian Institute of Health and Welfare (1999) projects that, based on the incidence rates existing in 1999, one in three men and one in four women would be directly affected by cancer in the first 75 years of life. Moreover, a loss of 254,000 potential years of life to the community each year was estimated as a result of people dying of cancer before the age of 75. As a whole, cancer is currently responsible for 29% of male deaths and 25% of female deaths in Australia. In Singapore, the proportion of cancer deaths among all causes of death rose steadily from 14.8% (in the years 1968-1972) to 27.1% (in the period 1998-2002) reflecting a worldwide trend. The above mentioned figures are alarming and have drawn the attention of researchers to understand the mechanism of cancer and come out with better therapies.

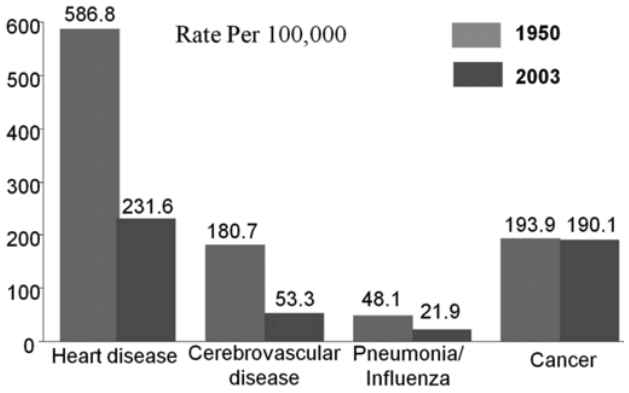


Fig.1. Change in death rates of different diseases in US from 1950 to 2003

The main characteristics of the cancer cell are its uncontrolled and unregulated growth (Hanahan and Weinberg, 2000; Martins et al., 2007). It is caused by the external factors such as UV radiations, carcinogenic chemicals as well as transfer of the cancer prone genes from parents. When a normal cell interacts with these external factors, its information system (DNA) gets damaged and the normal cell transforms to a cancer cell. Initially, the clump of the cancer cells is confined to particular location and it is regarded as being benign. If the cancer is not diagnosed and treated in the benign stage, it will change into a malignant form, and the cancer cells could migrate to different parts of the body and ultimately may lead to the death of the patient. So, it is better for the cancer patients to be provided suitable

treatment at the early stages itself so as to prolong and enhance the quality of their life.

1.1 Cancer treatment modalities

Over the past 50 years, many cancer treatment modalities have been discovered. The most prominent of these are surgery, chemotherapy, and radiation therapy. Some of the emerging therapies are immunotherapy and viral therapy. However, a specific therapy for all types of cancer is still missing and the available therapies have their own advantages and shortcomings. Surgical techniques to remove tumors have been in practice even thousand years ago. Usually, surgery is favoured to remove the tumors diagnosed at the very early stages ensuring almost complete cure. Surgery can be very difficult if the tumor location is near critical locations in the body (e.g. brain); furthermore, surgery is not a preferred option if the cancer had already metastasized by the time of diagnosis. In any case, complete clearance of the tumor cells is not assured with surgery. Radiation therapy is an alternative to surgery in order to kill the localized tumor cells. In radiation therapy, high energy radiations are used and its dosage is determined such that the normal cells near the tumor are spared. Thus, radiation therapy also depends on the location of the tumor and sensitivity of the tumor cells to the radiation.

As a result, surgery is followed by chemotherapy or radiation therapy to suppress further tumor growth. In the case of chemotherapy and radiation therapy, precise care should be taken to avert the damage to normal tissues. Among these two therapies, chemotherapy is preferred because it is a systemic therapy. In systemic therapy, the drug flows throughout the body and destroys the migrated cancer cells along with residual cancer cells near the surgery location. Chemotherapy is always given as a course in cycles based on the patient health status rather than as a one-time treatment. This is done so as to maintain the drug concentration within the dosage limits in the body and kill the remaining cancer cells in the subsequent treatments (Dua et al., 2008). However, the side effects of the chemotherapy are significant and sometimes they become serious than the disease itself. In contrast to chemotherapy and radiation therapy, immunotherapy has fewer side effects, because, typically, the patient's own cells are modified and used as therapeutic agents. The prime objectives of any therapy are to keep the number of cancer cells below a lethal level and avoid the side effects caused by the therapeutic agents. This can be achieved by optimal scheduling and the optimal administration of the therapeutic agents if the dynamical characteristics of the system are known.

2. CANCER IMMUNOTHERAPY

The role of the immune system in cancer treatment was first observed by William Coley, a New York surgeon. Cancer immunotherapy is the stimulation of immune cells to fight the tumor cells. The main function of the immune system is to fight against the abnormal changes in the body, and the successful functioning of it lies in its ability to distinguish the "self" and "non-self" based on the self-marking molecules. The immune system recognizes the abnormality with the help of antigens presented by the injured or abnormal cells. If the immune system exhibits a response based on antigen recognition, then antigens are called immunogenic. However, not all antigens are immunogenic. For example, if the tumor is larger in size, immune cells may not respond to it. Before studying the tumor-immune interactions, it is informative to take a brief look at the mechanisms of the immune system. In this regard, the review paper by Adam et al., (2003) provides a comprehensive discussion on the immune response in cancer.

In our body, the immune action is carried out by specialized cells called lymphocytes which are mostly present in the blood. The common lymphocytes are macrophages, dendritic cells, natural killer (NK) cells, lymphokine activated killer cells, B-cells and T-cells. Immune response is categorized as natural immunity, humoral immunity and cellular immunity based on the lymphocytes. Macrophages, dendritic cells, and natural killer cells are responsible for natural immunity, in which these cells directly attack the infected cells (cancer cell) and act as antigen presenting cells (APC). Antigen is an agent which can easily be recognized by immune cells. Thus, APC highlights the infected cells and alerts the T-cells for further action against the infected cells. In humoral immunity, antibodies produced by B-cells encounters the infected cells. Each B-cell has a specific antibody of a particular shape. The concept of antibody-antigen interaction resembles the mechanism of lock and key. When the shape of an antibody of a B-cell matches exactly with the shape of the antigen corresponding to the infected cells, B-cell proliferates and produces plasma cells which actively secrete the antibodies. These antibodies neutralize the activity of the infected cell by inhibiting their cell division process, by producing a lethal group of enzymes called complement and by opsonization. In opsonization, antibodies coat the infected cells in order to make them easily recognizable by the killer lymphocytes. This process is known as antibody dependent cell-mediated cytotoxicity. In cellular immunity, the key players are T-cells which are further classified as helper T-cells ($CD4^+$) and cytotoxic T-cells ($CD8^+$). Helper T-cells gets activated by the natural immune cells and regulates the production of the cytokines. Cytokines are the enzymes which keep the momentum of all the immune cells as per their requirement. Interleukins and interferons are regarded as the important cytokines to fulfil the immune action. Cytotoxic T-cells directly attack the infected cells after its activation by the cytokines.

From the above discussion, we can infer that key elements for immunotherapy are antibodies, cytokines, and the natural immune cells. On this basis, immunotherapy is classified into three main schemes, monoclonal antibody therapy (MAT), adoptive cellular immunotherapy (ACI) and vaccines (Adam et al., 2003). MAT involves the introduction of externally developed tumor specific antibodies into the patient's body using hybridoma technology. ACI constitutes the

modification of the lymphocytes (helper T-cells, cytotoxic T-cells, NK cells) by using cytokines like interleukin (Kirschner and Panetta, 1998). This ultimately ameliorates the anti-tumor activity of the lymphocytes. This is done in two ways namely, lymphokine-activated killer cell (LAK) therapy and tumor infiltrating lymphocyte (TIL) therapy. In both therapies, the lymphocytes are activated externally and later they are injected back at the tumor site. In LAK therapy, the lymphocytes are obtained from the *in vitro* culturing with high concentration of IL-2 which is extracted from the patient's blood. Conversely, in TIL therapy, lymphocytes are taken from the patient tumor sites and cultured with the high concentrations of interleukin *in vitro*.

Generally, cancer vaccines are in the form of cells, molecules or micro-organisms. They facilitate the quick recognition of tumor cells by the immune cells. In other words, tumor associated antigens (TAA) are made tangible to the immune cells. Consequently, tumor-immune interactions are enhanced thereby achieving the goal of cancer immunotherapy.

3. TUMOR-IMMUNE INTERACTION MODELS

Tumor-immune interaction models explain interactions between different types of immune cells and the tumor. There are many tumor-immune interaction models and a few of them will now be described. In the model proposed by De Boer and Hogeweg (1986), interactions between macrophages, T-lymphocytes and tumors are considered. The macrophages and T-lymphocytes are given in different dosages. This model captures the "sneaking through" phenomenon i.e. when lower dosages of tumor is introduced, the immune system may not recognize it and then the tumor grows to a bigger size; however, the immune cells reject the tumor when they are given in higher quantities. Another successful model (Kuznetsov et al., 1994) was developed based on the studies of the B-lymphoma BCL_1 in the spleen of mice. The model is very simple, and considers only two states (effector and tumor cells). Effector cells represent any of the killer immune cells. In this model, the parameters were estimated using *in vivo* data. Then, bifurcation analysis was performed to find the critical parameters for sneak through phenomenon. This model was further extended and modified (Kirschner and Panetta, 1998) by including the dynamics of the interleukin. Based on bifurcation analysis, this work emphasized that tumor immunogenicity is an important parameter. Their work also discussed about the effects of immunotherapy (adoptive cellular therapy and interleukin) through mathematical analysis. Models such as those by (Castiglione and Piccoli, 2007; de Pillis et al., 2005; de Pillis et al., 2006), consider the natural killer cells and cytotoxic $CD8^+$T cells as different states rather than considering them under the same family as effector cells. In de Pillis et al. (2005), the authors focus on the impact of NK cells and $CD8^+$ T cells on tumor growth. Their model is in the form of a system of ODEs. Parameters of the model were estimated and validated with the published mice and human data. In addition, sensitivity analysis was done on the model. The sensitivity analysis concluded that the variable to which model is sensitive is patient-specific. This model was extended with slight modifications in the functional forms of growth and death terms of the immune cells (de Pillis et al., 2006). The model was also used for understanding the effects of combination therapy (chemo-immuno and vaccine therapy) for different patient parameters. Other models have included the dynamics of NK cells, B cells, helper and cytotoxic T-cells , and LAK cells (Szymanska, 2003). In this work, dendritic cell vaccine (DCV) is considered. DCV is produced by the process called dendritic cell transfection. In this process, some TAA are configured and cultivated with autologous dendritic cells that is extracted from the patient itself (Cappuccio et al., 2007; Castiglione and Piccoli, 2007; Piccoli and Castiglione, 2006). The resulting vaccine is injected back into the patient. Here, we consider the mathematical model proposed by Piccoli and Castiglione (2006) because this is the only model which includes the dynamics of dendritic cells and apply multi-objective optimization using non-dominated sorting genetic algorithm (NSGA) to find the optimal scheduling of dendritic cell vaccine interventions.

4. MULTI-OBJECTIVE OPTIMIZATION

Multi-objective optimization (MOP) is the optimization of two or more conflicting objectives of a system represented in the form of a mathematical model subjected to certain known constraints. Most practical problems such as product and process design, finance, aircraft design, automobile design, and medical applications have multiple objective scenarios. In these problems, an optimal decision needs to be taken in the presence of trade-offs between the conflicting objectives. In MOP, there may be a number of solutions in the feasible region, and the decision maker has to analyze all the solutions based on the prior knowledge of the system before a final solution is adopted (Tamaki et al., 1996).

Suppose there are 'n' decision variables and 'p' objectives. MOP tries to find a point $x = (x_1, \ldots\ldots, x_n)$ which minimizes (or maximizes) the values of the objective functions $f = (f_1, \ldots\ldots, f_p)$ within the feasible region F of x. In contrast to single-objective optimization problems, an exact solution may not exist for the MOP problems because of the trade-off characteristics among the objectives. Hence a concept of the Pareto-optimal set was introduced for MOP problems. Pareto-optimal set is, 'a family of points which is optimal in the sense that no improvement can be achieved in any objective without degradation in others'.

Definition (Tamaki et al., 1996): Let $x^0, x^1, x^2 \in F$

1. x^1 is said to be *dominated* by (or inferior to) x^2, if $f(x^1)$ is partially less than $f(x^2)$, i.e.,

$f_i(x^1) \geq f_i(x^2)$, $\forall_i = 1, \ldots\ldots, p$, and $f_i(x^1) > f_i(x^2)$, $\exists_{i=1,\ldots\ldots p}$.

2. x^0 is the Pareto-optimal (or non-dominated), if there doesn't exist any $x \in F$ such that x dominates x^0

As Pareto-optimal solution is a logical strategy to the MOP, the prime goal of solving the MOP is to obtain a Pareto-optimal set. The Pareto-optimal solutions can be obtained by solving on a one-at-a-time basis using single objective optimization methods like weighted sum method and the ε-constraint method. Unlike the conventional methods, population-based methods (e.g. evolutionary algorithms) such as genetic algorithm, particle swarm optimization, simulated annealing and differential evolution can generate Pareto-optimal set simultaneously. The searching strategy of different evolutionary algorithms is different and they are based imitating some natural processes. However, the common theme of all the evolutionary algorithms is to search the whole hyper-domain of decision variables and find the best possible solution. Population based methods are further subdivided into non-Pareto approaches and Pareto approaches. In non-Pareto approach, the selection/reproduction of the new population in the subsequent generations are based on the objective function values whereas in Pareto-approaches, the new population is generated not only on the basis of the objective value themselves but also on their dominance property. In this work, we used the Pareto-based approach known as non-domination based genetic algorithm (NSGA-II) which was proposed by Deb et al., (2000) because of its elitism and minimal computational complexity. Genetic algorithms (GAs) which imitate the process of natural evolution have shown successful results in many optimization problems which are difficult to solve by the conventional methods of the mathematical programming (Nemhauser et al., 1989).

4.1 Non-domination based genetic algorithm for multi-objective optimization

A brief introduction to the NSGA algorithm (Deb et al., 2000) is provided here. The algorithm is initiated with suitable values for population size and number of generations. The stopping criterion of the algorithm is the maximum number of generations. Broadly, the prime steps involved in the algorithm in each generation are selection, offspring production and recombination. First, the population is initialized randomly within the bounds of the decision variables. Once the population is initialized, they are sorted into separate fronts based on non-domination as discussed earlier. Among these fronts, the first front members completely dominate others in the current population and the second front members are dominated by only the first front members and so on. Each individual in the front is given a rank (fitness) based on the front in which they are present. The first front individuals are assigned a fitness value of 1 and second front individuals are assigned fitness value of 2 and so on. Apart from this, a parameter called crowding distance is calculated for each individual. The interpretation of the crowding distance is the closeness of an individual to its neighbours. A larger crowding distance indicates the diverse nature of the population. Crowding distance is compared only when the individuals belong to the same front. Thus, best N parent individuals are selected from the current population based on the rank and crowding distance, where N is the population size. Then, in the offspring production step, the selected parent individuals are used to generate offspring via the crossover and mutation operators. Finally, in the recombination step, the offspring population is combined with the current generation population and the combined population is used as an initialized population for the next generation. In this way, the procedure is repeated until the maximum number of generations. Thus, the important tuning parameters in this algorithm are the number of generations, indices for crossover and mutation processes. The schematic representation of the algorithm is shown in Fig. 2.

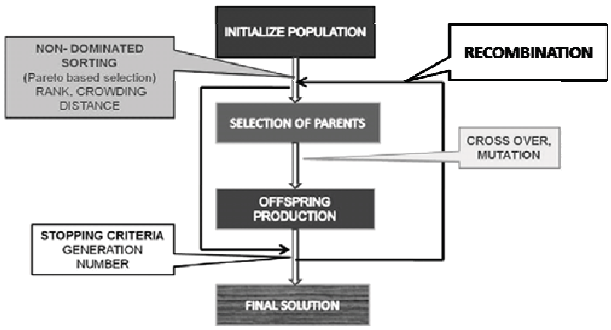


Fig.2. Non-dominated sorting genetic algorithm

5. MATHEMATICAL MODEL

The model taken from Piccoli and Castiglione (2006) describes the interactions among the tumor cells, helper and cytotoxic T-cells, dendritic cells and interleukin. The model assumes that tumor cells are immunogenic and do not metastasize. In other words, tumor cells are recognized by dendritic cells and are presented to cytotoxic T-cells. The interactions between the cytotoxic T lymphocytes and the tumor cells are described by a kinetic scheme and are presented in the form of ordinary differential equations. The states in the system are denoted by

- *H(t)*, helper T-cells (CD4+)
- *C(t)* , cytotoxic T-cells (CD8+)
- *M(t)*, tumor cells
- *D(t)*, dendritic cells
- *I(t)* , interleukin

The pharmacodynamics is represented by the term e_2MC in equation (3) and the pharmacokinetics is captured in equation (4). In equation (4), 'u' is the input rate of the dendritic cells. Once the dendritic cells are injected, the CD4+ cells, CD8+ cells, and interleukins are triggered as shown in equations (1, 2 & 5) respectively by the following terms. ($c_0Dd_0\gamma(H, f_0)$, $c_1I(M+D)d_1\gamma(C, f_1)$, a_4HD). The first two terms in equations (1) and (2) correspond to the natural evolution of

the cytotoxic T-lymphocytes. In the same way, the second and third terms in equation (5) explain the loss of interleukin (IL-2) due to its interactions with CD8+ cells and its natural decay respectively. The model assumes that tumor growth follows the logistic equation with the constants 'd_2' and 'f_2' in the absence of immune interactions. Another assumption in the model is that externally administered DCV is the only source of dendritic cells. The initial value of tumor is $M(0) = 0.1$, the initial level of helper and cytotoxic T-cells are taken to be their equilibrium values $H(0) = a_0/b_0$ and $C(0) = a_1/b_1$. It is assumed that there is no immune response before the treatment and therefore $I(0)$ and $D(0)$ are taken to be zero.

5.1 Model equations

$$\frac{dH}{dt} = a_0 - b_0 H + c_0 D d_0 \gamma(H, f_0), \quad (1)$$

$$\frac{dC}{dt} = a_1 - b_1 C + c_1 I(M + D) d_1 \gamma(C, f_1), \quad (2)$$

$$\frac{dM}{dt} = d_2 \gamma(M, f_2) - e_2 MC, \quad (3)$$

$$\frac{dD}{dt} = -e_3 DC + u, \quad (4)$$

$$\frac{dI}{dt} = a_4 HD - c_4 CI - e_4 I, \quad (5)$$

where $\gamma(x, c) = x(1 - x/c)$

5.2 Problem formulation:

Objective 1: $$\min_{u(t_i)} \int_0^{t_f} M(t, u(t_i))\, dt + M(t_f) \quad (6)$$

Objective 2: $$\min_{u(t_i)} (Max(M(t, u(t_i)))) \quad (7)$$

t_f is the final time i.e. the planning horizon and t_i is the i^{th} time injection of the dendritic cells.

Constraints:

Equations (1) through (5) i.e. the mathematical model and

$$t_{i-1} \le t_i \le t_{i+1} \ , \ i = 1, 2...., 9$$

In this paper, two objectives are considered based on the typical goals sought by oncologists. Objective 1, as given by equation (6), seeks to minimize the summation of the running load and final load of tumor cells. Objective 2 (equation (7)) seeks to minimize the maximum possible value of the tumor cells in the given time horizon. If only objective 1 is taken into consideration, it may so happen that the tumor cells may shoot up to a very high value at a particular time while remaining at lower values at other times. This sudden shoot up of tumor cells to a higher value may lead to later stages of cancer which is very difficult to treat. So, in order to maintain the tumor in the benign stage, objective 2 is also considered. The time horizon considered to implement the multi-objective optimization is 4500 hours (approximately 6 months). In the given time horizon, the plan is to give injections of DCV ten times. Thus, the decision variables of the problem are the time of injections. It is assumed that the duration of injection every time is one hour and total vaccine quantity given in each injection is 0.5. Thus, $u(t_i)$ is equal to 0.5 c mm^{-3} h^{-1}.

Table 1. Parameter values

Parameter	Description	Value	Units (c=cells, h=hours)
a_0	CD4 T birth rate	10^{-4}	c h^{-1} mm^{-3}
b_0	CD4 T death rate	0.005	h^{-1}
c_0	Max. proliferation of CD4 T	10	
d_0	½ saturation constant of CD4 T	10^{-2}	c $^{-1}$ h^{-1} mm^3
f_0	Carrying capacity of CD4 T	1	c mm^{-3}
a_1	CD8 T birth rate	10^{-4}	c h^{-1} mm^3
b_1	CD8 T death rate	0.005	h^{-1}
c_1	Max. proliferation of CD8 T	10	
d_1	½ saturation constant of CD8 T	10^{-2}	h^{-1} $(mm^{-3}/c)^2$
f_1	Carrying capacity of CD8 T	1	c mm^{-3}
d_2	½ saturation constant of tumor	0.02	h^{-1}
e_2	Killing by CD8 of tumor	0.1	c $^{-1}$ h^{-1} mm^3
f_2	Carrying capacity of tumor	1	c mm^{-3}
e_3	CD8 T killing of DC	0.1	c $^{-1}$ h^{-1} mm^3
a_4	IL-2 production by CD4 T	10^{-2}	c $^{-1}$ h^{-1} mm^3
c_4	IL-2 uptake by CD8 T	10^{-7}	c $^{-1}$ h^{-1} mm^3
e_4	IL-2 degradation rate	10^{-2}	h^{-1}

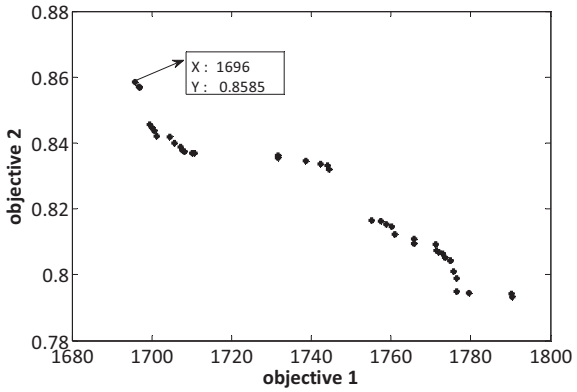


Fig.3. Optimal Pareto solutions for the multi-objective optimization problem

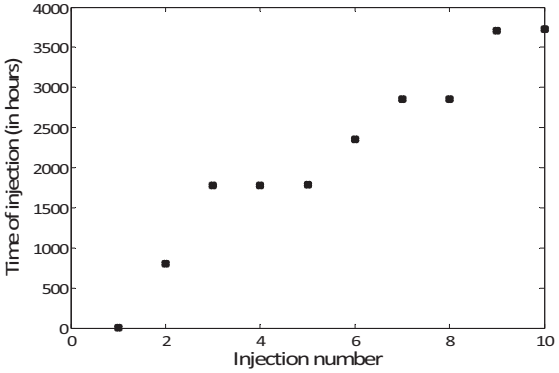


Fig.4. Plot of time of injections of the dendritic cell vaccine for the chosen solution

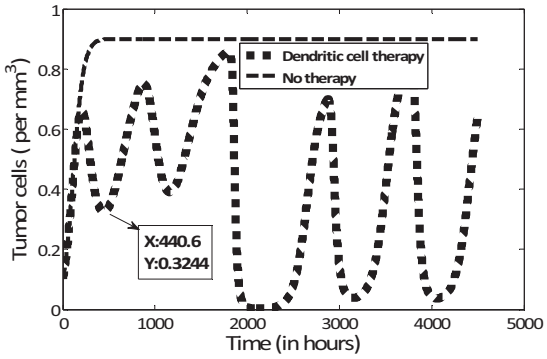


Fig.5. Dynamics of tumor cells for no therapy and dendritic cell therapy

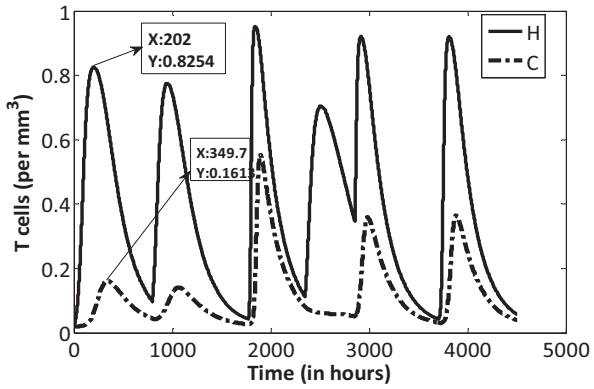


Fig.6. Dynamics of helper T-cells and cytotoxic T-cells

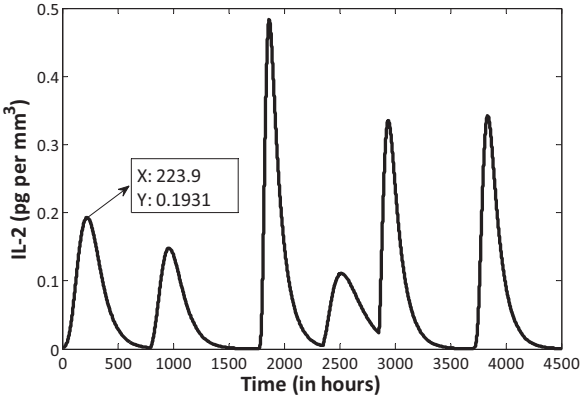


Fig.7. Dynamics of interleukin

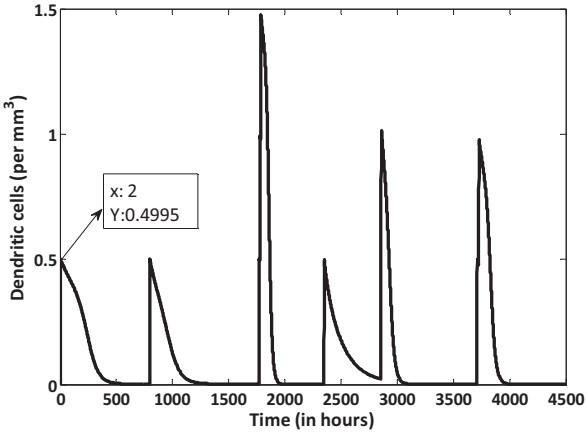


Fig.8. Dynamics of dendritic cells

6. RESULTS AND DISCUSSION

We used MATLAB's implementation of NSGA-II to solve the multi-objective optimization problem as outlined above. The algorithm starts with an initial population of possible solutions within their mentioned bounds and in each generation, the solutions are updated based on the genetic principles. In this work, we considered the population number to be 40 and the maximum number of generations as 40.

Finally, the algorithm provides the best Pareto curve as shown in Fig.3. It can be seen from the Fig.3 that there are gaps in the Pareto curve in spite of the problem being continuous. These gaps can be filled by increasing the population number and the number of generations of the genetic algorithm, but at the cost of higher computational effort. Then, one of the solutions can be chosen from the Pareto curve as per the requirement. Here, solution marked in Fig 3. is chosen because it corresponds to the least value of objective 1 (tumor burden) among all points of the Pareto curve. The time of injections for the chosen point from the Pareto curve is given in Fig 4. According to this treatment plan, DCV is injected and the evolution of the tumor cells, T-cells, interleukin, and dendritic cells are shown in figures 5 through 8 respectively.

From Fig.4, it is observed that injection timings are almost close during (1700-1800 hours, 2850-2900 hours, and 3700-3750 hours). The reason is the increase of tumor cells during these periods. So, in order to bring them down immediately, the frequency of injection times is increased. Thus, the treatment planning seems to be reasonable and logical.

From figures 5 through 8, the first intervention of dendritic cells is considered. This intervention is given in the second hour. Because of this intervention, the helper T-cells, interleukin, and cytotoxic T-cells reach their highest value at around 200 hours (8 days), 221.9 hours (9 days), 349.7 hours (15 days) respectively from the time of intervention of DCV. At the same time, tumor cells reach a lower value at around 440.6 hours (19 days) due to the first intervention. Similarly, tumor cells took around 15 days to reach its lower value after second intervention of DCV. Thus, for the given input of DCV, the time gap between the interventions should not be more than 3 weeks in order to avoid the peak value of the tumor cells. Moreover, it takes more time for the helper T-cells and cytotoxic T-cells to get activated before reacting to the tumor cells. This activation time can be reduced by injecting more DCV (but within the threshold range) or by introducing the interleukin therapy (approved by FDA) where interleukin is also given externally. One of our future plans is to consider a combination of dendritic cell therapy and interleukin therapy.

7. CONCLUSIONS

We have applied multi-objective optimization to find the optimal schedule of dendritic cell therapy for a given input of dendritic cells. The freedom of choice of any one solution as per the requirements is a major advantage of multi-objective optimization strategy. The obtained treatment planning seems to be reasonable in controlling the tumor cells from reaching a higher value. Further analysis suggests that the time gap between the interventions should be less than 3 weeks in order to achieve the objective. Thus, the obtained protocol design can guide caregivers in treating cancer subjects.

ACKNOWLEDGEMENTS

This work was supported by the National University of Singapore in the form of a research scholarship to the first author.

REFERENCES

Adam, J.K., Odhav, B. and Bhoola, K.D. (2003) Immune responses in cancer, *Pharmacology & Therapeutics*, **99**, 113-132.

Cappuccio, A., Castiglione, F. and Piccoli, B. (2007) Determination of the optimal therapeutic protocols in cancer immunotherapy, *Mathematical Biosciences*, **209**, 1-13.

Castiglione, F. and Piccoli, B. (2007) Cancer immunotherapy, mathematical modeling and optimal control, *Journal of Theoretical Biology*, **247**, 723-732.

De Boer, R.J. and Hogeweg, P. (1986) Interactions between macrophages and T-lymphocytes: Tumor sneaking through intrinsic to helper T cell dynamics, *Journal of Theoretical Biology*, **120**, 331-351.

de Pillis, L.G., Gu, W. and Radunskaya, A.E. (2006) Mixed immunotherapy and chemotherapy of tumors: modeling, applications and biological interpretations, *Journal of Theoretical Biology*, **238**, 841-862.

de Pillis, L.G., Radunskaya, A.E. and Wiseman, C.L. (2005) A Validated Mathematical Model of Cell-Mediated Immune Response to Tumor Growth, *Cancer Res*, **65**, 7950-7958.

Deb, K., Agrawal, S., Pratap, A. and Meyarivan, T. (2000) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. Springer-Verlag, Paris, France, 849-858.

Dua, P., Dua, V. and Pistikopoulos, E.N. (2008) Optimal delivery of chemotherapeutic agents in cancer, *Computers & Chemical Engineering*, **32**, 99-107.

Hanahan, D. and Weinberg, R.A. (2000) The Hallmarks of Cancer, *Cell*, **100**, 57-70.

Kirschner, D. and Panetta, J.C. (1998) Modeling immunotherapy of the tumor – immune interaction, *Journal of Mathematical Biology*, **37**, 235-252.

Kuznetsov, V.A., Makalkin, I.A., Taylor, M.A. and Perelson, A.S. (1994) Nonlinear dynamics of immunogenic tumors: Parameter estimation and global bifurcation analysis, *Bulletin of Mathematical Biology*, **56**, 295-321.

Martins, M.L., Ferreira Jr, S.C. and Vilela, M.J. (2007) Multiscale models for the growth of avascular tumors, *Physics of Life Reviews*, **4**, 128-156.

Nemhauser, G.L., Rinnooy Kan, A.H.G. and Todd, M.J. (1989) *Optimization*. North-Holland.

Paola Pisani, F.B., D. Maxwell Parkin, (2002) Estimates of the world-wide prevalence of cancer for 25 sites in the adult population, *International Journal of Cancer*, **97**, 72-81.

Piccoli, B. and Castiglione, F. (2006) Optimal vaccine scheduling in cancer immunotherapy, *Physica a-Statistical Mechanics and Its Applications*, **370**, 672-680.

Szymanska, Z. (2003) Analysis of immunotherapy models in the context of cancer dynamics, *Int.J. Appl. Math. Comput. Sci.*, **13**, 407- 418.

Tamaki, H., Kita, H. and Kobayashi, S. (1996) Multi-objective optimization by genetic algorithms: a review. *Proceedings of IEEE International Conference on Evolutionary Computation*. 517-522.

An Industrial and Academic Perspective on Plantwide Control

James J. Downs*. Sigurd Skogestad**

**Eastman Chemical Company, Kingsport, TN 37662, USA*
(e-mail: jjdowns@ eastman.com).
*** Dept. of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway*
(email: skoge@ntnu.no)

Abstract: The purpose of this paper is to stress the importance of selecting the right plantwide control structure and the need for a formalized approach that can encompass the many issues that arise in plantwide control design. Since the concept of process control design based on a holistic view of the process came about, the variety of procedures and approaches to the design problem have illustrated the difficulty of a unified approach. Using examples, a formal design approach is presented to help put in context the need and advantages of using such an approach. The examples deal with disturbance rejection, throughput maximization and economic optimization of plants consisting of parallel units.

Keywords: process control, control structure design, plantwide control, inventory control, throughput

1. INTRODUCTION

Industry uses a variety of approaches to accomplish plantwide control design. The range of tools used spans from engineering judgment to the applications of complex model based algorithms. Over the last 40 years the field of research in this area has attacked this design problem on various levels. Larsson and Skogestad (2000) provide a good review of the various approaches. Design heuristics based on experience, design rules based on case studies, algorithms for objective function minimization, etc. have all contributed to the improvement of how designs can be accomplished (e.g., Downs, (1992), Narraway and Perkins (1993), Luyben (1998), Zheng, Mahajanam and Douglas (1999), Kookos and Perkins (2002), Chen and McAvoy (2003), Vasbinder and Ho (2003), Skogestad (2004), Konda et al. (2005), Ward et al. (2006)). However, the complex nature of the problem and the various depths to which it needs to be solved have resulted in a design procedure that is difficult to piece together from the various approaches that have been put forth. This is not a new issue and almost 20 years ago the "Tennessee Eastman challenge problem" (Downs and Vogel, 1993) was put forward so that various approached could be tested against each other. Nevertheless, today in industry, much of the research in this area has still not gained traction to have the profound influence possible. The purpose of this paper is to stress the need for a formalized yet simple approach that can encompass the many levels that arise in plantwide control design.

In section 2 industrial aspects of plantwide control design are discussed and two examples illustrate how industry may approach the plantwide control problem using a single criterion for design guidance. This points to the need for a more formal procedure which is presented in Section 3. In Section 4 the inclusion of plantwide economic variables is presented and illustrated in Section 5. The paper concludes that the formal approach presented is a step in the direction of helping to organize the design procedure for plantwide control. This paper also illustrates the application of the formal procedure to more complex examples that illustrate plantwide design involves many issues and one-criteria approaches may not be sufficient.

2. STATUS IN INDUSTRY

The traditional approach for designing process control strategies for chemical plants has been to set production rates by setting process feed rates and then to design automatic control systems around each unit operation sequentially through the process. For processes having significant in-process inventory and not too much in the way of recycles, this approach can be used successfully. However, as processes become more complex and at the same time have less in-process inventory, the design of a plant-wide control strategy becomes a more important part of the overall process control design problem. The interrelation of the plant-wide control strategy with the process chemistry and economics requires both control theory and also process knowledge. It has become apparent that the design of plant-wide control strategies involves not only the development and application of process control theory but also, in a more fundamental sense, the development of a methodology one uses to approach the plant-wide control problem.

While we usually think about material balance and energy balance equations applying to a unit operation, they also apply to whole processes and to entire chemical complexes. The time it takes to accumulate and deplete inventories may be longer for large processes or chemical complexes, but the laws of accumulation and depletion of material hold nonetheless. Whereas for a process, we assume the rate of accumulation of each component to be zero, the fact that the control system must ensure that to be the case is often

overlooked. The manipulation of flows, utilities, and the readjustment of process operating conditions to maintain a balance of material and energy entering and leaving an entire process is one of the overriding priorities for the control system (Buckley, 1964). The material balance must be maintained not only from an overall viewpoint but also for each component in the system.

While traditional control theory can be used to approach the control problem as, "Given a process described by a model of the form ... ", the plant-wide control problem requires much more in the development of the problem statement itself. It is not intuitively obvious at the outset what the underlying control problems are - much less how they should be solved. As researchers have begun to explore the plant-wide control area, the application of methods and techniques as applied to case studies has elucidated issues that are difficult to quantify and are in need of further discussion and research.

Despite the ever-increasing incentive, segregation of the process design and control tasks is still common. Two contributing factors to this segregation are: (1) the difficulty of changing from the historical approach of fixing the process design before the control engineer becomes involved, and (2) the difference in the thought pattern of design and control engineers. In addition it can be costly and time consuming to address controllability and operability in a rigorous way at the design stage. The common notion is that process economics are solely determined by the steady-state process design. While the nominal steady-state design point is very important, it loses its distinction if one is unable to maintain plant operation at the design point. Design decisions are often based on steady-state analysis without consideration of controllability, process and product variability, or plant-wide control issues. The basic thought pattern in the design stage usually follows the form, "Given these conditions, create a design to perform this function" (design question), as opposed to, "Given this design, how well will it perform its intended function?" (rating question). As existing plants are pushed to produce greater throughputs, an additional question becomes important, “Given this plant, how can I maximize profit?” (optimization question).

Current industrial practice is usually focused on unit operation control. This viewpoint emanates from the overriding issue of reliable operation. These unit control strategies are simple and understandable by operators and engineers alike and lead to operations that when “sick” can usually be healed without the capabilities of experts. This approach has worked reasonably well for many years. Furthermore, the high costs of building new facilities have led to more retrofits and plants producing products that they were not designed to produce. As plants are campaigned to produce a wider variety of product specifications, control strategies that are simple and perhaps applicable to many different operating points, can result in more reliable operation.

This current design practice is being challenged as process economics drive toward fewer new designs and more operation of existing facilities in new ways. Techniques for plant wide process control design are needed (1) that result in processes that are operated in near optimal fashion while not employing complex control technology and (2) that do not require the care and feeding of control experts. Several approaches that address the attainment of optimal operation of plants while not requiring implementation of complex, perhaps difficult to understand control systems, have emerged. Two of these, self optimizing control design (Skogestad, 2000) and operational strategies based on process chemistry (Ward et al., 2004, 2006) have found particular appeal at Eastman.

The importance of being able to discriminate how process variables need to behave to achieve optimal operation is fundamental when designing plantwide strategies. Often the underlying unit operation strategies can be kept simple and usually SISO while the overall plant wide strategy is optimizing plant operation in a more natural fashion. This approach has wide appeal when plant reliability and control system understandability are required. Each of these approaches builds into the control system a natural “self-optimizing” that is part of normal operation. Contrasted with the centralized approach of using models to determine an optimum and then driving a process to that optimum point, optimization designed in from the bottom up provides the important robustness and reliability component.

From start-up the primary objective for a new plant is to achieve nameplate capacity in a reliable and predicable way. Often times the need for optimization of plant operations comes after the facility has been operational for a few years. By this time top down optimization strategies can be implemented, provided the plant has a good regulatory control system. If the optimization strategy is counterintuitive, then operator understanding can suffer. We can all attest to the uphill battle to achieve routine usage of a control system that, while driving the process to the correct economic conditions, does so in an unusual or difficult to understand fashion.

The importance of having plantwide control strategies that are optimizing in a natural, fundamental way can have long term effects. Operator training and understanding during the early years of plant operation sets thought patterns for years to come. When the need for plant optimization arises, the basic building blocks of how the control system automatically drives plant operation are in place. The process optimizer at this time may only have to make small adjustments to a process that is close to optimum already. The trick, of course, is that these strategies must be basically simple and for the most part SISO. Our experience is that for plants where “self-optimizing” regulatory control strategies have been build in from the beginning, we have been successful with process optimization projects that have been undertaken. On the other hand, for older processes which have control strategies not designed with optimization in mind, we may struggle for years working to gain operator acceptance to a new strategy. Even the simple idea of setting

process throughput at a place other than the process feed can become a difficult endeavour.

Example 2.1 - Changing the production throughput manipulator (TPM) for an esterification plant: Eastman operates many processes that have produced chemicals for over 50 years. Esterification chemistry is well known and has been a workhorse for the company. Units that were built 50 years ago were typically designed with the process throughput set at the feed to the process. Control systems consisted of pneumatic single input / single output controllers that were difficult to change and had a long operating history. As production rates increased over the years due to demand growth and incremental process improvement, the original plantwide strategy would become limiting. The original plant had the standard scheme with the throughput manipulator (TPM) located at the feed as illustrated in Figure 2-1.

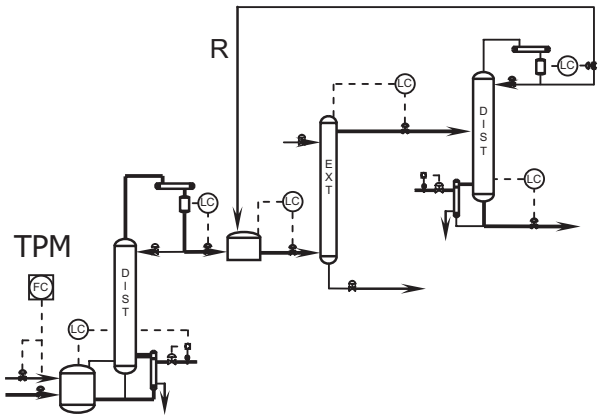


Figure 2-1. Esterification process. Inventory control strategy with column 1 feed rate used as the TPM.

In the late '70's and early '80s' Eastman benefited from implementing a change in the TPM location on numerous plants. Early adoption of this significant change was difficult because of (1) an ingrained mindset toward needing process feeds constant, (2) operator understanding of an "inventory-to-feed" strategy, and (3) the difficulty of reversing the control decision using pneumatic hardware. Today at Eastman, the notion of setting the TPM at a location other than the process feeds is common and is driven by variability propagation and ease of operation requirements. The benefits of choosing the best location for the TPM have also become realized in our capital design process.

For the esterification process the first change was to move the TPM from the process feed rate to the distillate flow rate leaving the first distillation column as shown in Figure 2-2. This strategy worked well for many years because many of the disturbances entering the reactor were directed away from the more sensitive separation portion of the process. The extraction step of the process was intended to wash unreacted alcohol from the ester product. As the extraction step became the process bottleneck, it became evident that its behaviour as a function of organic feed rate was very nonlinear. This nonlinearity stemmed from the fact that increasing organic

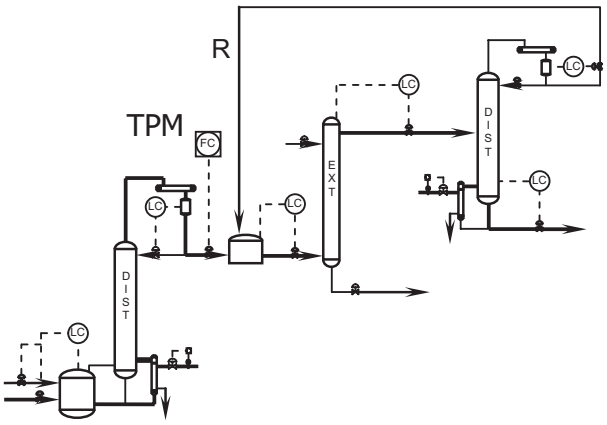


Figure 2-2. Esterification process. Inventory control strategy with column 1 distillate rate used as the TPM.

feed rate resulted in an increasing composition of the alcohol taken from the extractor to the final distillation column. The increase in distillate rate needed to remove the alcohol from the final product would aggravate the situation by increasing the feed rate to the extractor (stream "R" in Figure 2-2). The point at which the process would enter this "windup" varied with the amount of unreacted alcohol reaching this part of the process. This windup in the recycle loop is similar to Luyben's "snowball effect" (Luyben, 1994), but the cause in our case is a limitation in mass transfer rate whereas in Luyben's case it is a limitation in reaction rate. For this process, the windup condition usually took 12-24 hours to get fully engaged. This made it difficult for operators to confidently set the production rate. In addition, what may be a maximum and stable rate today might result in the windup condition tomorrow. The outcome of this uncertainly resulted in operations setting a lower than optimum production rate to guarantee process stability.

A further improvement in locating the TPM occurred when it was relocated to be the feed to the extraction system (Figure 2-3). Obviously, this eliminated variability from propagating to the extractor, but more importantly, it resulted in a self regulating system that avoids the windup should the operator set the TPM too high. In particular, if the TPM is set too high and excess alcohol leaks to the final distillation system, take note of the system response to the extra distillate flow recycled to the extractor. Namely, it results in less flow being drawn from the front end of the process and the extractor, while not at the optimum feed rate, does remain stable. This situation is quite recoverable by operators who note that production rates have fallen, and realize that they have set the extractor feed rate too high. We found that the operators were capable of optimizing the operation once fear of setting the extractor feed too high was removed.

The principle that proved most useful is the idea that the optimum did not lie against or close to a process cliff. The original strategy was very unforgiving once the process was pushed too far. Extractor flooding, loss of liquid/liquid immiscibility, and flooding of the final distillation column

meant several hours of lost production. The ability to experiment with the process without the penalty of passing this "point of no return", gave operators confidence in the control system to recover if they ended up pushing rates too high.

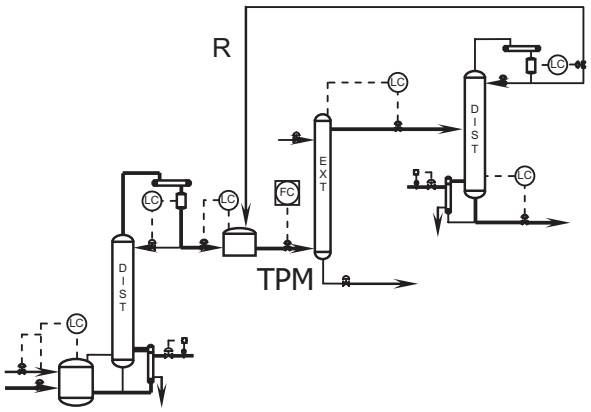


Figure 2-3. Esterification process. Inventory control strategy with extractor feed rate used as the TPM.

Example 2.2 - Control strategy for a liquid-liquid extraction process: During the control design phase one may chose from a variety of criteria to drive the control strategy design and the criterion chosen is usually based upon engineering judgement. The importance of the criterion choice is often not appreciated. The objective of this example is to illustrate how the choice of a design criterion that aims to *propagate disturbances to insensitive locations* results in a particular design. The resulting control strategy can then be compared with those obtained using a more methodical approach.

Consider the extraction process in Figure 2-4 where acid is transferred from the water/acid feed (F) to the extract (E) by use of a solvent (S). The remaining water is the raffinate product (R). The total inventory is self-regulated by overflow of extract, but the interface level (component inventory) does not self regulate. How should this inventory be controlled? Two alternatives are shown in Figure 2-5.

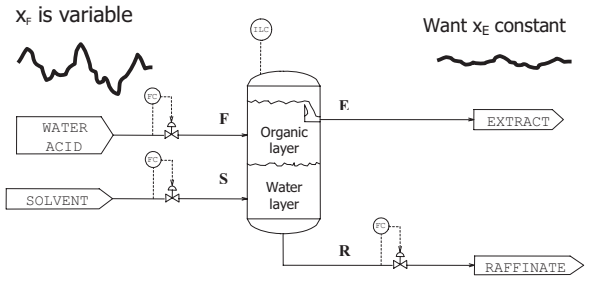


Figure 2-4. Liquid-Liquid Extraction Process

Strategy I. Let aqueous feed F control interface level (with constant outflow R)
Strategy II. Let aqueous outflow R control interface level (with constant feed F).

Both of these structures have been used for extraction control in various services including the example given here. Obviously, both structures work and give the same result if everything is constant (no disturbances). How do the two strategies differ when there are disturbances? To understand the difference we ask the question: "Where does the disturbance go"?

Let x denote the acid fraction, and consider variation (disturbance) in the acid feed fraction x_F by ±1% (30±1%). For strategy I, the resulting variation in the acid composition of the extract product (x_E) is ±0.856% (21.4 ±0.856%) and for strategy II it is ±0.506% (21.4 ±0.506%). For details see the mass balances in Table 1. Thus, strategy II is the preferred strategy of the two if the objective is to have small variations in extract composition, x_E.

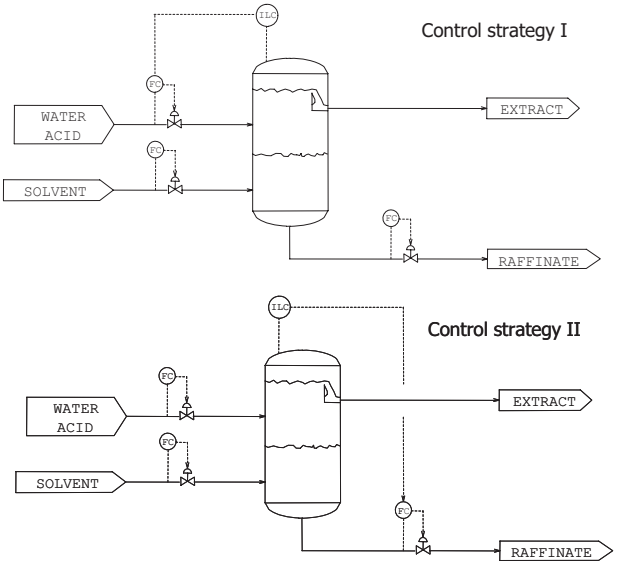


Figure 2-5. Alternative Control Strategies for Liquid-Liquid Extraction

In summary, for strategy I the variation in x_F results in variations in the feed flow, F, and in x_E (with gain 0.856), while for strategy II the variation in x_F results in variations of the outlet flow, R, and to a lesser extent in x_E (with gain 0.506). Strategy II is the preferred strategy of the two if the main objective is to have small variations in x_E. This example suggests that different inventory strategies may result in process variability being transferred to portions of the process that are insensitive to variation or portions in which variability is harmful. The idea of propagating disturbances to insensitive locations gives good insight and can result in good designs. However, for more complex problems and for less experienced engineers a more systematic approach is needed.

	Feed, F				S	Extract, E				Raffinate, R			
Case	Base	I	II	III,IV	All	Base	I	II	III,IV	Base	I	II	III,IV
Water	70	70.568	69	66.774	0	10	10.568	10.33	10	60	60	58.67	56.77
Acid	30 (30%)	31.705 **(31%)**	31 **(31%)**	30 **(31%)**	0	30 (21.4%)	31.705 (22.3%)	31 (21.9%)	30 (21.4%)	0	0	0	0
Solvent	0	0	0	0	100	100	100	100	100	0	0	0	0
Total	100	102.273	**100**	96.77	0	140	142.273	141.33	140	60	**60**	58.67	56.77

Table 1. Mass balances for extraction process: Base case ($x_F = 30\%$) and with disturbance ($x_F = 31\%$) for control strategies I, II, III, and IV. (Assumption: Equilibrium relationship Acid/Water = 3 in extract, E.)

3. A PLANTWIDE CONTROL DESIGN PROCEDURE

No matter what approach we use, the following decisions need to be made when designing a plantwide control strategy:

Decision 1. What to control? Selection of controlled variables (CVs) to achieve
a. Good steady-state performance (economics), and
b. "Stable" operation with little dynamic drift (including selecting CVs related to inventories)

Decision 2. Where to set the production rate? Placement of throughput manipulator (TPM)

Decision 3. How to control the inventories? How to pair the loops? That is, selection of a control configuration that interconnects CVs and MVs.

Often in industrial practice all issues are considered simultaneously without making formal decisions that answer the above three questions. For the extraction process in Example 2.2 the need for good extract composition raised the question of how best to control the aqueous inventory. This naturally leads one to consider the same issues on a broader, plantwide scale. To be effective, a more systematic procedure is helpful.

The plantwide control structure design procedure of Skogestad (2004) consists of the following seven steps:

I. Top-down part

Step 1. Define operational objectives (economics) and constraints.
Step 2. Identify degrees of freedom (MVs) and optimize operation for important disturbances (offline analysis)
Step 3. Select primary (economic) controlled variables
*- **Decision1a***
Step 4. Select location of throughput manipulator
*- **Decision 2***

II. Bottom-up part

Step 5. Structure of regulatory control layer (including inventory control)
*a. Select secondary ("stabilizing") CVs (**Decision 1b**)*
*b. Select "pairings" between CVs and MVs (**Decision 3**)*
Step 6. Structure of supervisory control layer (decentralized, MPCs?)
*- Related to **Decisions 1a and 3***
Step 7. Structure of (and need for) optimization layer (RTO)
*- Related to **Decision 1a***

The top-down part (steps 1-4) is mainly concerned with economics and steady-state considerations are often sufficient. Dynamic considerations are required for steps 4 to 6.

Steps 1 and 2 involve *analysis* of the optimal operation of the plant, and should form the basis for the actual *decisions* in Steps 3 to 7. A detailed analysis in steps 1 and 2 requires that one has available a steady-state model and that one performs optimizations using the model for various disturbances. This is often not done in industrial practice. The model used for design may not be suitable or available, the working relationship between the design and control functions may be weak, or there may not be time to perform this analysis.

Nevertheless, one should at least perform a simplified engineering version of steps 1, 2, and 3 where one thinks through the economics of the present and future operation with aim of using process insight to propose which variables to control, keep constant, from a steady-state economic point of view. In particular, a good engineer can often easily identify the "active constraints" that the control system should maintain. That is, where should one optimally stay at maximum or minimum values of flow, temperature, pressure, composition, etc?

Simplified Step 1-3. Identify degrees of freedom and main disturbances. Based on process insight, select variables to keep constant at steady-state in order to achieve close-to-optimal economic operation (in spite of disturbances).
*- **Decision 1a***

There have been many applications of the above design procedure, e.g. see Araujo et al. (2007), but most of them on academic problems. There exist several other procedures for plantwide control (e.g., Luyben et al., 1998), but they focus mainly on the bottom-up part, and in particular on Step 5. However, making good decisions in step 5 can be difficult without having first gone through the top-down plantwide economic analysis in steps 1 to 4.

Step 4 (location of TPM) was addressed in Example 2.1, and this issue is further discussed in the recent PhD thesis by Aske (2009); see also Aske et al. (2008).

The focus of the rest of this paper is on step 3 (economic CVs). In this respect it is important to notice that the best control structure may vary, and, depending on market conditions, there are two main modes of operation:

Mode I. Maximize efficiency (for a given throughput).

With a given throughput (production rate), the value of the products is usually known, and provided there are degrees of freedom left after satisfying the constraints (specifications), the economic objective is to minimize the use of utilities, maximize raw material yield, and to minimize waste treatment costs. These and other issues that increase specific production costs are the same as maximizing the efficiency. As discussed in section 2 on the industrial status, the control system for a new plant is usually set up to handle this mode of operation well. Changes in production rate are considered a disturbance.

Mode II. Maximize throughput (with production rate as a degree of freedom).

When market demand is good and product prices are high, the profit is maximized by running the plant at maximum throughput. In fact, the first thing that the operation people usually focus on after startup of a new plant is to increase capacity because the opportunities for extra profit in mode II are usually much larger than in mode I (In spite of this there is usually no effort during the design phase to design a control system that can operate at maximum throughput). Operation at maximum throughput usually corresponds to using all degrees of freedom to satisfy active constraints. There will be a bottleneck somewhere in the plant against which operation at maximum throughput will run. Trying to increase the throughput will result in infeasible operation in the bottleneck unit. The maximum flow through the bottleneck unit is then an active constraint, and operation in mode II should be focused on keeping this flow at its maximum (Aske, 2009).

The esterification plant in Example 2-1 is a case of operating in mode II with the extraction section being the bottleneck.

Example 3-1 - Application of the design procedure to Example 2-2: The design criterion for Example 2-2 was that disturbances should be propagated to insensitive locations. At this point we want apply the more systematic plantwide procedure. The process is very simple, so we use the simplified approach for selection of controlled variables (there are no degrees of freedom left for economic optimization once the specifications are satisfied).

Simplified Step 1-3. Identify degrees of freedom (MVs) and main disturbances and based on process insight, select primary controlled variables (Decision 1a).

The extract product flow (E) is on overflow, so there are 3 MVs that can be used for control; the two feed flows (F and S) and the raffinate R. However, at steady state there are only 2 degrees of freedom because the interface level, which has no steady-state effect, needs to be controlled. Further, the throughput is assumed to be given (mode I), which consumes another degree of freedom. We are then left with only 1 steady-state degree of freedom, and thus need to decide on 1 "economic" CV. From process insight it is important to maintain a constant product composition (x_E) so we decide that this should be controlled. There are then no degrees of freedom left for economic optimization.

Decision 1a: The acid product composition x_E should be kept constant. The "economic" CV is therefore $CV_1 = x_E$.

Step 4. Select location of throughput manipulator (TPM) (Decision 2).

The location of the TPM influences the structure (pairing) of the inventory control system in Step 5. The throughput is often located at the main feed, but could generally be anywhere in the process. Since the two proposed control strategies both have a constant solvent feed flow, we assume here that the solvent feed S is the throughput manipulator ***(Decision 2).***

Step 5. Structure of regulatory control layer (including inventory control)

Decision 1b: The total inventory is self-regulated by overflow, but also the interface level between the two liquid phases must be controlled. Thus, CV_2 = interface level.

We must next decide *how* to control the interface level. With solvent feed rate S as the TPM, we have left two candidate MVs: Feed F and outflow R. The main issue for regulatory control is usually dynamics, and from this point of view there does not seem to be any significant difference between the two choices. Another issue for regulatory control is to avoid saturation of the MV, and this tells us that we should prefer the largest flow, which is the feed F. However, one should also think ahead to Step 6, which is the structure of the supervisory layer. Here, the concern is to control acid composition ($CV_1 = x_E$) which depends directly on the feed F but only indirectly on the outflow R. Thus, we would like to "save" F for the supervisory layer.

Decision 3. Use R to control the interface layer (MV_2 = R). This gives inventory control in the direction of flow, which is normal with the throughput set at the feed.

Step 6. Structure of supervisory control layer

Decision 3, continued. The remaining MV_1 = F is used to control acid composition ($CV_1 = x_E$). The final control structure is shown as strategy III in Figure 3-1.

Note that we assumed that the product composition x_E can be measured ($CV_1 = x_E$), but this may not be possible in practice. We then need to find something else to "control" (keep constant). This is what we indirectly did in the previously proposed strategies where we selected

Strategy I: Keep CV_1 = R constant (and use F to control the interface level)

Strategy II: Keep CV_1 = F constant (and use R to control the interface level)

However, both of these strategies give undesired variations in the product composition x_E; we found $\Delta x_E / \Delta x_F = 0.856$ for strategy I and $\Delta x_E / \Delta x_F = 0.506$ for strategy II. It is possible to add a supervisor layer, where one adjusts R (strategy I) or F (strategy II) such that x_E is kept constant. This modification to strategy II is shown as *Strategy III* in Figure 3-1.

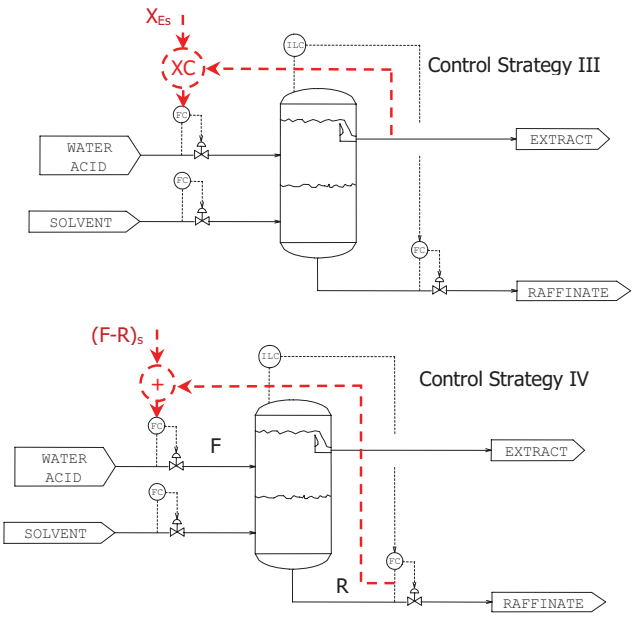


Figure 3-1. Self-Optimizing Strategies for Extraction Process

However, assume there is no online measurement of the extract composition x_E. One option would then be to estimate x_E using a model and available measurements ("soft sensor"), but this is a bit complicated. Is it possible to find a simple strategy (maybe a combination of strategies I and II) that gives $\Delta x_E / \Delta x_F = 0$? Yes, it is!

Strategy IV: Keep the flow difference CV_1 = F – R constant (while at the same time adjusting F and R to control interface level). One possible implementation is shown in Figure 3-1.

Why does strategy IV give constant composition x_E? Controlling the interface level (which indirectly depends on the feed composition x_F) closes the material balance at steady-state. From the total material balance we have E = S + (F-R) so by keeping F-R constant, we have that the flow E remains constant (because the throughput S is constant) and assuming equilibrium it follows that the composition of E must remain constant (again assuming S constant). If the throughput S varies (disturbance) then all flows should be scaled by S to keep x_E constant, so by process insight we derive that an "improved strategy IV" is to keep the variable $CV_1 = (F-R)/S$ constant.

Strategy IV is a special case of a "self-optimizing" measurement combination, as discussed below. In fact, since we have $n_d = 1$ disturbance (x_F) and $n_u = 1$ steady-state degrees of freedom, we have from the nullspace theorem (Alstad and Skogestad, 2007) that self-optimizing control can be obtained by controlling a combination of $n_d + n_u = 2$ independent measurements. The flows (MVs) R and F are here candidate "measurements", so a possible controlled variable is $CV_1 = h_1 F + h_2 R$, where in general the optimal h_1 and h_2 can be found from the nullspace theorem. In this example, we found by process insight that the optimal choice is $h_1 = 1$ and $h_2 = -1$ (strategy IV).

4. SELECTION OF ECONOMIC (PRIMARY) CVs

In the above example, we found that the flow difference F-R is a good primary CV. How do we select primary (economic) CVs in a systematic manner (step 3)?

We make the standard assumption here that a steady-state analysis is sufficient for studying the economics. The question is: How can we turn optimization into a setpoint problem? The issue is to find some "magic" variable, c, to keep constant. The obvious "magic" variable is the gradient of the cost function, $J_u = dJ/du$, which should be zero at the optimum point, independent of disturbances. However, before we look at this idea, let us look in a bit more detail in Steps 1 to 3 in the proposed procedure for selecting economic CVs.

Step 1. Define operational objectives (cost J) and constraints

In many cases a simple economic cost is used:

> Profit = - J = value products – cost feeds – cost utilities (energy)

Other operational issues, such as safety and environmental impact are usually formulated as constraints. For cases with good marked conditions we often have a constrained optimum and the cost function can be simplified to J = - TP (mode II, maximum throughput).

Other cost functions are also possible. For example, consider the extraction process. Here, the optimum is to keep a constant product composition x_E, but this is not possible, even at steady-state, because there is no online measurement. We therefore want to control something else that gives *indirect control* of the primary output (Hori et al., 2005). The cost function is then $J = (x_E - x_{ES})^2$.

Step 2. Identify degrees of freedom and optimize operation for various disturbances.

One approach is to use a steady-state flowsheet simulator, if available, to optimize operation (with respect to the degrees of freedom) for various disturbances. In many cases, simpler models and approaches may be used. Typical "disturbances" include feed composition, feed rate, reaction rate constants, surroundings, values of constraints and prices.

Step 3. Select primary (economic) controlled variables

The issue is to select the primary (economic) controlled variables (CVs). That is, for what should we use the (steady-state) degrees of freedom? What should we control?

1. Control active constraints. The active constraints come out of the analysis in step 3 or may in some cases be identified based on physical insight. The active constraints should be selected as CVs because the optimum is not "flat" with respect to these variables. Thus, there is usually a significant economic penalty if we "back off" from the active constraints, *so tight control of the active constraints is usually desired.*

Specifically, in mode II the feed rate should be adjusted to keep the bottleneck unit operating at its active constraints. Any back-off from the active constraints will reduce the flow through the bottleneck unit and give a loss in feed flow (production) which can never be recovered.

2. Identify "self-optimizing" variables related to the (possibly) remaining unconstrained degrees of freedom. These are "magic" variables which when held constant result in close-to-optimal operation (with a small loss), in spite of the presence of disturbances. The term "magic" is used to signify that the choice may have a significant effect on the economics (loss), and that it is not generally obvious what a good choice is. A good self-optimizing variable should give a "flat" optimum, which means that tight control of these variables is usually not required (as opposed to the active constraints). Note that the different self-optimizing variables must be found for each region of active constraints.

There are two main possibilities for selecting self-optimizing CVs:

1. Select single measurements as CVs (however, it is difficult to find single measurements in a systematic manner, so one must often use the "brute force" approach)
2. Use measurements combinations as CVs (here, methods exist to find optimal combinations).

To identify good candidates for a controlled variable, c, we may use the following four requirements (Skogestad, 2000):

Requirement 1. Its optimal value is insensitive to disturbances (so that the optimal variation Δc_{opt} is small).

Requirement 2. It is easy to measure and control accurately (so that the implementation error n is small).

Requirement 3. Its value is sensitive to changes in the manipulated variable, u; that is, the gain, G, from u to c is large (so that even a large error in controlled variable, c, results in only a small error in u. Equivalently, the optimum should be 'flat' with respect to the variable, c.

Requirement 4. For cases with two or more controlled variables, the selected variables should not be closely correlated.

All four requirements should be satisfied. For example, for a marathon runner, the heart rate may be a good "self-optimizing" controlled variable (to keep at constant setpoint). Let us check this against the four requirements. The optimal heart rate is weakly dependent on the disturbances (requirement 1) and the heart rate is easy to measure (requirement 2). The heart rate is relatively sensitive to changes in power input (requirement 3). Requirement 4 does not apply since this is a problem with only one unconstrained input (the power).

In addition to the above requirements, some systematic approaches to evaluate and find good "self-optimizing" CVs (especially associated with the unconstrained degrees of freedom) are:

Approach 1 - Brute force. Conceptually, the simplest approach for finding candidate CVs is the "brute force" approach where one considers the economic loss imposed by keeping a candidate set of CVs constant when disturbances occur (rather than re-optimizing their values),

$$\text{Loss} = J(\text{CV} = \text{constant}, d) - J_{opt}(d)$$

The term "brute force" is used is because one must do a separate evaluation of each candidate set of CVs. The "brute force" approach is the most general and exact method, but also the most time consuming method because there are essentially an infinite number of possible CVs (at least if measurement combinations are included) that can be suggested, and for each of them we need to do computations to find the cost for each disturbance.

The "brute force" approach was essentially what we initially tried with strategies I and II for the extraction process, where we evaluated the change in product composition ($\Delta x_E / \Delta x_F$) resulting from a disturbance in feed composition.

Approach 2 - Use analytic expressions or insight about the optimum. This is not a general approach, but it may be

very effective for cases where it works. One useful method is to start from the fact that at the optimum the gradient of the cost J with respect to the degrees of freedom should be zero:

At optimum: Gradient = $J_u = dJ/du = 0$

These are also known as the necessary condition of the optimum (NCO) (Srinivasan, et al). It seems obvious that the gradient CV = J_u is the "ideal" self-optimizing variable (Halvorsen and Skogestad, 1997), However, it may be difficult to obtain the expression for J_u or it may depend on non-measured variables.

Approach 3 - Exact local method and optimal measurement combinations. The details are found in Halvorsen et al (2003), Alstad and Skogestad (2007) and Alstad et al. (2009). For the case single measurements as CVs, this is a "local" version of the brute force approach. However, the evaluation is much more efficient. In addition, the "nullspace method" can be used to find truly optimal measurement combinations, as was done in strategy IV for the extraction process.

Approach 4 - Maximum gain rule. The maximum gain rule (Halvorsen et al., 2003) says that one should control "sensitive" variables with a large scaled gain |G|/span(CV). This captures two main concerns:

1. *The optimal value of the CV should be approximately constant* (independent of disturbances), that is, span(CV) should be small.
2. *The CV should be sensitive to changes in the unconstrained degrees of freedom* (to ensure a flat optimum), that is the gain G = ΔCV / ΔMV should be large.

The maximum gain rule can be derived from the exact local method by making some not too serious assumptions. An important advantage of the maximum gain rule is the insight that it gives.

5. OPTIMAL OPERATION OF PARALLEL UNITS

Let us return to an important problem, often encountered in industrial practice. During the life of production of a product, a company often times expands capacity as demand grows. Early plant design may involve process designs based on incomplete data as time to market drives commercialization timelines. Once operation begins, improved operating conditions, equipment designs, and process topology emerge. When capacity expansion takes places the new capacity may come simply by adding equipment to the existing process or by construction of a parallel plant. The new plant is seldom run in a "stand alone" fashion, but instead may share some unit operations with the existing facility. As expansion continues, the complexity of the topology among the plants can lead to plant wide control problems.

In its simplest form, consider a number of plants operating in parallel, each of differing ages, and each with its own efficiency and yield relationships that are dependent on throughput. How should we optimally load each plant to achieve a target production while minimizing the total production costs? We can derive useful result from the necessary optimality condition $J_u = 0$. We derive that, provided the total production rate is given, it is optimal to load the units such that we have **equal marginal costs in all units** (which corresponds to $J_u = 0$).

Proof. To derive this result, consider *n independent parallel units* with a given total load (e.g., given total feed). Let the total cost be $J = \Sigma J_i$ and let the total feed (or some other limited load for the units) be fixed, $F = \Sigma F_i$. The necessary conditions of optimality is that $J_u = \delta J/\delta u = 0$ where u in this case is the vector of feed rates F_i. Since the total feed is fixed, there are n-1 independent degrees of freedom F_i, and we assume these are the F_i's for n-1 first units (and for unit n we have $F_n = F - \sum_{i=1}^{n-1} F_i$) The units are assumed to be independent which means that the cost in unit i, J_i, depends only on the flow into unit i, F_i. However, note that when we make a change in F_i, we also need to change F_n, and we have $dF_n = - dF_i$. The optimality condition $\delta J/\delta F_i = 0$ for variable F_i then becomes

$$\delta J/\delta F_i = \delta(J_1 + J_2 + \dots J_i + \dots J_n)/\delta F_i = \delta(J_i + J_n)/\delta F_i$$
$$= \delta J_i/\delta F_i - \delta J_n/\delta F_n = 0$$

or $\delta J_i/\delta F_i = \delta J_n/\delta F_n$. Since this must hold for all i units, we have proved that one should operate such that the *marginal cost* $\delta J_i/\delta F_i$ *is the same in same units.* **End proof.**

Urbanczyk and Wattenbarger (1994) applied this criterion to the maximization of oil production of wells that produce both oil and gas, but where the total gas handling capacity is fixed (limited). In their application J_i is the oil production and F_i is the gas production in well i, and the idea is to operate the wells such that $\delta J_i/\delta F_i$ is the same for all wells; that is, by increasing the gas production by a given amount δF_i one gets the same benefit in terms of extra oil production δJ_i in all wells.

Good self-optimizing variables are then the difference in marginal cost between the units (which should be zero). Below we discuss two industrial applications of this idea.

Example 5-1 – Operation of parallel refining systems: Eastman received an industry award for its application of advanced control to optimally load three parallel refining systems. Each system consists of four distillation columns used to refine crude reactor product. The application uses process data to establish operating costs for processing material from crude reactor effluent to saleable product. Based on operating costs, process operation limits, and utility availability, the feed rate to each refining train is adjusted to match reactor production with refining system production. *The allocation of load to each system is adjusted to achieve equal marginal refining costs.*

Example 5-2 – Syngas production in parallel furnaces: For many years Eastman produced synthesis gas by reacting methane and steam in reforming furnaces. The process consisted of 15 furnaces operated in parallel, see Figure 5-1. The effluent gas from the furnaces was combined as feed to three carbon dioxide removal systems. The product syngas from the three carbon dioxide removal systems was combined to form a single product gas used in downstream chemical production. The 15 reforming furnaces, constructed over the span of three decades, each had different energy efficiency characteristics as well as different yield performance as technology advanced. In addition, the three carbon dioxide removal systems were of varying efficiency and performance. Newer systems were better instrumented, had valves that performed better, and had on-line analytical measurements. At any time, there were one to three furnaces down for routine maintenance.

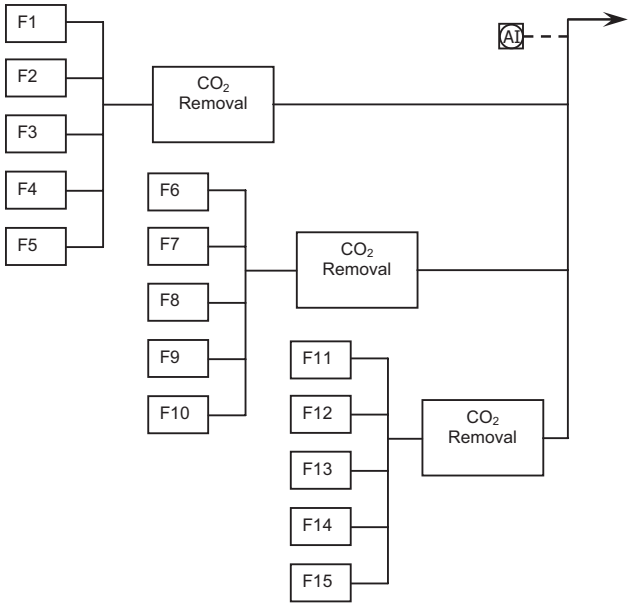


Figure 5-1. Syngas process with fifteen furnaces and three CO_2 removal systems.

The optimum operation of the plantwide system to coordinate pressure and production among the interconnected gas flow network was a significant challenge. The simple objective of matching production of syngas with consumption often ended up varying the production rate on the newest furnace because it could most gracefully handle the needed changes. From an optimization point of view this approach usually resulted in the most efficient units not being operated at their maximum rates.

Normally with units in parallel, an expected "self-optimizing" strategy is to operate with the same outlet conditions (temperatures or compositions) of all parallel units. This would have been a good strategy if the reactors were identical, but, for this example it is more economical to operate each furnace differently based on its particular efficiency and yield profile and then ensure that the combination stream met the total stream specifications. In particular, the newer more efficient furnaces were able to produce a much purer product for the same cost as the older units producing a much less pure product. The purity of the product from each furnace was a relatively weak function of feed rate. The final layer of complexity arises from the efficiency of the carbon dioxide removal system. Each system was connected to a designated set of furnaces so that it was beneficial to operate furnaces linked to the better performing carbon dioxide removal system.

The optimization layer to coordinate the total process production and the allocation of that production to various parallel units was complicated by the presence of crossover lines. These lines added operational flexibility but created an ever increasing complexity of the optimization problem. Local MPC controllers for furnace operation and supervisory control for the carbon dioxide removal systems allowed for near optimal operation at the local level illustrated in Figure 5-2. Overall optimization was approached by production loading strategies and coordination using a supervisory MPC controller. As solutions to this problem were developed, it became clear that technology to guide us on the appropriate degree of decentralization was sparse. Developing a centralized system with all the CV's and MV's in

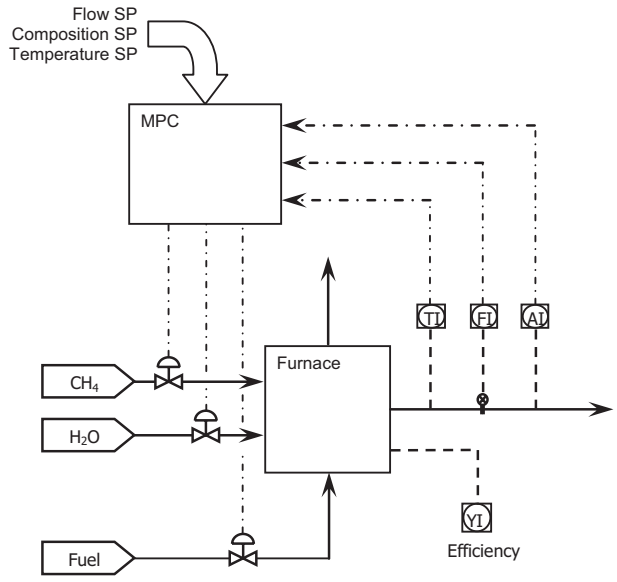


Figure 5-2. Individual syngas furnace control with three manipulated variables and three controlled variables.

the same MPC was problematic due to the routine on-line/off-line operation of the furnaces. Being able to gracefully add and remove systems from the overall control system was critical to success. In addition, measurement reliability often resulted in some furnaces being operated in "local' mode; i.e., not connected to the centralized MPC. The eventual control system needed to be developed and commissioned in reasonable time, needed to be implemented on available hardware, needed to be understood by plant operating staff, and had to be maintainable as process

improvements were made. This led to a decentralized strategy choice as shown in Figure 5-3.

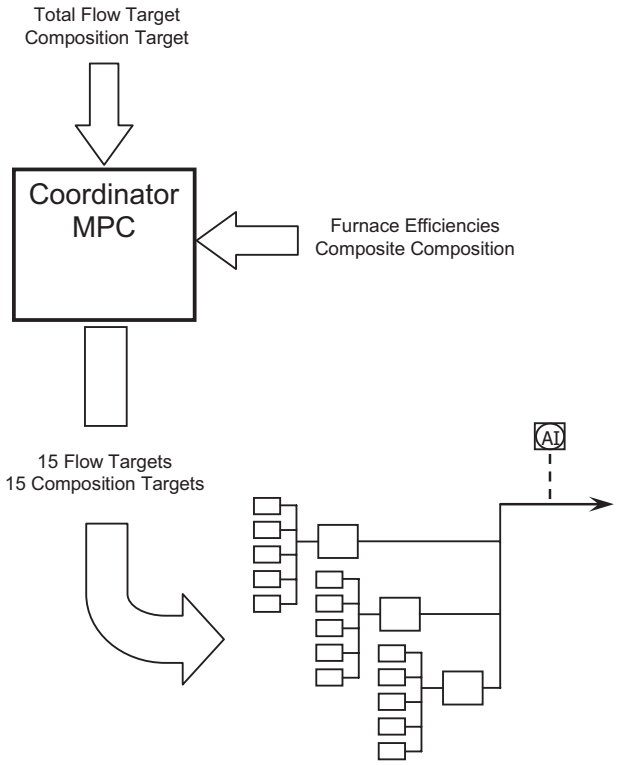


Figure 5-3. Coordination MPC supervising fifteen local furnace MPC controllers.

These examples illustrate the complex nature of an industrial plantwide control problem. The use of a formalized procedure can make known improved strategies that may go undetected when using only one or two design criteria for guidance. The ability to weave in practical issues that complicate implementation is paramount. Using a formalized procedure can help unscramble the vast array of decisions that can overwhelm designers and cause them to continue reliance upon a unit operation focus.

6. DISCUSSION

The design of plantwide control strategies can be seen from two viewpoints. These are (1) the design of control strategies for the regulation of plant material and energy inventories and (2) the design of control strategies for process economic optimization. The design of inventory control strategies determines the manipulated variables that remain for process economic optimization. Concepts are needed that guide the design of the inventory control strategy such that the design of the process economic optimization strategy is made easy. It is clear that if a good job is done during the design of the inventory control strategy, such as setting the TPM near the bottleneck, then the remaining process economic optimization strategy design is made easier. On the other hand, if the inventory control strategy results in key optimization variables being far away from available manipulated variables, then strategies for optimizing process economics will be difficult if not impossible to implement.

The examples illustrate that the inventory control strategy design not only affects the dynamics between manipulated and controlled variables used for optimization, but also can change the gain as well. The emphasis upon placement of the TPM for a process has long been recognized as a key decision in the resulting inventory control strategy. It is becoming more evident that this decision also determines the difficulty of the remaining process economic optimization strategy design. Techniques to determine self-optimizing control variables can be effectively and easily employed if the variables available to optimize the process have good dynamic linkage with their manipulated variable counterparts.

The examples also illustrate that the application of optimization from a top down viewpoint may guide one to select manipulated variables that should remain free for economic optimization while other should be used for inventory control. The formalization of a procedure to organize the design of these two phases of control includes the concepts of: (1) TPM location within the process, (2) control of unit operation process variables against their local constraints, and (3) the development of measurement combinations whose control implies nearness to the economic optimum.

The application of plantwide control design procedures for new plants is certainly an obvious direction of growth. However, the redesign of plantwide control structures for existing plants has been shown to be very beneficial. The known locations of process bottlenecks, known market conditions and product demands, and the operating nuances of a running process all make the plantwide design procedure more understandable and manageable. Using a procedure to determine alternate control structures can lead to new ideas for control that may have been missed for existing processes. As noted, the migration from tried and true control, but inferior, control strategies to new and unfamiliar strategies can be difficult.

7. CONCLUSION

Since the concept of process control design based on a holistic view of the process came about, the variety of procedures and approaches to the design problem have illustrated the difficulty of a “one size fits all approach.” The examples presented illustrate the application of a few industrial design approaches. A more formal design procedure is presented and it is applied to the industrial examples. The importance of addressing process economics in the control design procedure is discussed and the industrial need to run plant at their maximum feed rate (mode II) is emphasized. The use of a plantwide design procedure that

incorporates and organizes the variety of concerns and technical issues in this important area is demonstrated.

REFERENCES

V. Alstad and S. Skogestad, "Null Space Method for Selecting Optimal Measurement Combinations as Controlled Variables", *Ind. Eng. Chem. Res.*, **46** (3), 846-853 (2007).

V. Alstad, S. Skogestad and E.S. Hori, "Optimal measurement combinations as controlled variables", *J. Proc. Control*, **19**, 138-148 (2009).

A.C.B. Araujo, M. Govatsmark and S. Skogestad, "Application of plantwide control to the HDA process. I Steady-state and self-optimizing control"', *Control Engineering Practice*, 15, 1222-1237 (2007).

E.M.B. Aske, S. Strand and S. Skogestad, "Coordinator MPC for maximizing plant throughput", *Computers and Chem. Eng.,* **32**, 195-204 (2008).

E.M.B. Aske, *Design of plantwide control systems with focus on maximizing throughput*, PhD Thesis, Norwegian University of Science and Technology (NTNU) (2009)

P.S. Buckley, *Techniques of Process Control*, Wiley, New York (1964), Chapter 13.

R. Chen and T.J. McAvoy, "Plantwide control system design: Methodology and application to a vinyl acetate process", *Ind. Eng. Chem. Res.*, **42** (20), 4753-4771 (2003).

J.M. Douglas, *Conceptual Design of Chemical Processes*, McGraw-Hill, New York (1988).

J.J. Downs, "Distillation Control in a Plantwide Control Environment", In: *Practical Distillation Control*, W.L. Luyben (ed.), Van Nostrand Reinhold, 413-439 (1992)

J.J. Downs and E.F. Vogel, "A Plant-Wide Industrial Process Control Problem", *Computers and Chem. Eng.,* **17**, 245-255 (1993)

I.J. Halvorsen and S. Skogestad, "Optimizing control of Petlyuk distillation: Understanding the steady-state behavior", *Computers and Chem. Eng.,* **21**, Suppl., S249-S254 (1997).

I.J. Halvorsen, S. Skogestad, J.C. Morud and V. Alstad, "Optimal selection of controlled variables", *Ind. Eng. Chem. Res*., **42** (14), 3273-3284 (2003).

E.S. Hori, S. Skogestad and V. Alstad, "Perfect steady-state indirect control", *Ind. Eng. Chem. Res.*, **44** (4), 863-867 (2005).

NVSNM Konda, G.P. Rangaiah and P.R. Krishnaswamy, "Plantwide control of industrial processes: An integrated framework of simulation and heuristics", *Ind. Eng. Chem. Res.* , **44** (22), 8300-8313 (2005).

I.K. Kookos and Perkins, J.D., "An Algorithmic method for the selection of multivariable process control structures", *J. Proc. Control*, 12, 85-99 (2002).

W.L. Luyben, B.D. Tyreus and M.L. Luyben, *Plantwide Process Control*, McGraw-Hill, New York (1998).

W.L. Luyben, "Snowball Effects in Reactor Separator Processes with Recycle", *Ind. Eng. Chem. Res.* , **33**, 299-305 (1994).

L.T. Narraway, J.D. Perkins, "Selection of control structure based on economics", *Computers and Chem. Eng.,* **18**, S511–S515 (1993).

S. Skogestad "Plantwide control: the search for the self-optimizing control structure", *J. Proc. Control*, **10**, 487-507 (2000).

S. Skogestad, "Control structure design for complete chemical plants", *Computers and Chem. Eng.,* **28** (1-2), 219-234 (2004).

B. Srinivasan, D. Bonvin and E. Visser, "Dynamic optimization of batch processes – III. Role of measurements in handling uncertainty", *Computers and Chem. Eng.,* **27** (1) 27–44 (2003).

E.M. Vasbinder and K.A. Hoo, "Decision-based approach to plantwide control structure synthesis", *Ind. Eng. Chem. Res*., **42**, 4586-4598 (2003).

J.D. Ward, D.A. Mellichamp, and M.F. Doherty, "Importance of Process Chemistry in Selecting the Operating Policy for Plants with Recycle", *Ind. Eng. Chem. Res*., **43**, 3957 (2004).

J.D. Ward, D.A. Mellichamp, and M.F. Doherty, "Insight from Economically Optimal Steady-State Operating Policies for Dynamic Plantwide Control", *Ind. Eng. Chem. Res*., **45**, 1343 (2006).

C.H. Urbanczyk and R.A. Wattenberger, "Optimization of well rates under gas coning conditions", *SPE Advanced Technology Series*, **2**, (2), 61-68 (1994).

A. Zheng, R.V. Mahajanam, J.M. Douglas, "Hierarchical procedure for plantwide control system synthesis", *AIChE Journal*., **45**, (6) 1255-1265 (1999)

Distributed Control

Oral Session

Industrial Implementation of a Coordinator MPC for Maximizing Throughput at a Large-Scale Gas Plant

Elvira Marie B. Aske [*,**] **Stig Strand** [**] **Sigurd Skogestad** [*]

[*] *Department of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway, e-mail: sigurd.skogestad@chemeng.ntnu.no (Sigurd Skogestad)*
[**] *StatoilHydro R&D, Integrated Operations and Process Control, N-7005 Trondheim, Norway*

Abstract: An industrial implementation of a coordinator MPC to maximize throughput at the large-scale Kårstø gas plant is described. The "coordinator MPC" coordinates the flows through the network and not the local MPCs. It uses as degrees of freedom (MVs) the flows not used by the local MPCs (feeds, crossovers), and maximizes the throughput subject to the keeping the remaining capacities in all units zero or positive. A key idea is to use the local MPCs to estimate the remaining capacities in the units (Aske et al., 2008). Although not fully implemented, the coordinator MPC is found to be a promising tool for implementing maximum throughput.

Keywords: Model based control, throughput, implementation, optimization, capacity.

1. INTRODUCTION

This paper describes an actual industrial implementation of the method for maximum throughput proposed earlier by Aske et al. (2008). The application is the Kårstø gas processing plant located in Norway, which receives rich gas and unstabilized condensate through pipelines from more than 30 producing offshore fields. This set high demands, not only to the plant efficiency and its regularity, but also to the plant throughput. Limited gas plant processing capacity means that one or more fields must reduce production or even shut down. Therefore, it is important that the Kårstø plant does not become a "bottleneck" in the Norwegian gas transport system. The Kårstø plant has no recycles or reactors, but it has several independent feeds and parallel flows that make it possible to have multiple bottlenecks at the same time. The bottlenecks may move due to disturbances, thus the throughput maximizing is a dynamic and multivariable problem.

The overall feed rate (or more generally the throughput) affects all units in the plant. For this reason, the throughput is usually not used as a degree of freedom for control of any individual unit, but is instead left as an "unused" degree of freedom (u^c) to be set at the plant-wide level. The throughput at the Kårstø plant is presently set by the operators who manipulate the feed valves to satisfy orders from the gas transport system (operated by another company). The objective of this work is to coordinate the throughput manipulators (u^c) to achieve economic optimal operation.

In general, to optimize the economic operation of a plant, one may use real-time optimization (RTO), normally based on (rigorous) steady-state models. Standard RTO methods require the plant to be close to steady state before performing a reoptimization based on data reconciliation or parameter estimation (Marlin and Hrymak, 1997). However, many plants are rarely at steady state or important economic disturbances occur more frequent than the controlled plant response times. At least in theory, it is then more suitable to use dynamic optimization with a nonlinear model, which may be realized using dynamic RTO (DRTO) or non-linear model predictive controller (MPC) with an economic objective, e.g. Engell (2007); Strand (1991).

In this study, a different approach is used. We assume that optimal economic operation is the same as maximizing plant throughput, subject to achieving feasible operation (satisfying operational constraints in all units) with the available feeds. This corresponds to a constrained operation mode with maximum flow through the bottleneck(s). At maximum throughput, all throughput manipulators (u^c) are used to satisfy active constraints (bottleneck). Thus a nonlinear model of the entire plant is not needed, and instead linear MPC may be used. One option is to combine all the MPCs in the plant into a single application. However, here we choose to decompose the problem by keeping the local MPC applications and introducing a coordinator MPC (Aske et al., 2008) to maximize throughput. The coordinator uses the remaining degrees of freedom (u^c) to maximize the flow through the network subject to satisfying given constraints. The remaining degrees of freedom (u^c) include feed rates, feed splits and crossovers. The constraints are the feasible remaining capacities of the individual units ($R_k > 0$). The feasible remaining capacity R_k is how much more feed unit k can receive while operating within its constraints. For most units, R_k is not a quantity that can be measured, because it depends on the operation of the unit. For

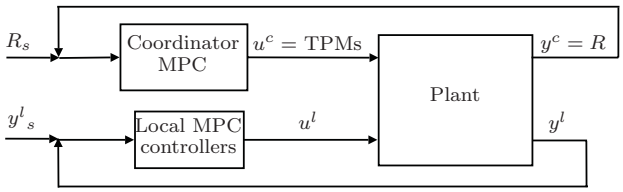


Fig. 1. Plant decomposition by coordinator MPC. The local MPC applications uses u^l to control the local targets y^l_s, whereas the coordinator uses the throughput manipulators (u^c = TPMs) to control the remaining capacity ($y^c = R$) in the units.

example, the capacity may be increased by producing less pure products. A key idea in the approach of Aske et al. (2008) is to use the local MPC to estimate R_k. By estimating R_k for each unit, the plant-wide control problem is decomposed and the application becomes smaller in size and hence easier to understand and maintain. The plant decomposition is illustrated in Fig. 1.

This paper considers about half of the Kårstø gas processing plant. The application presently includes 12 distillation columns, 2 compressor stages, 4 feed valves and 2 crossovers (splits). The main reason for not including the entire plant is that local MPC applications are yet not implemented on all units. All MPC applications at the Kårstø plant use the in-house SEPTIC technology (Strand and Sagli, 2003).

This paper is organized as follows. The local MPC controllers for the individual units are discussed briefly in Section 2. The local MPCs adjust the local degrees of freedom (u^l) such that the operation is locally feasible. However, local feasibility requires that the feed rate to the unit F^l_k is below its maximum capacity, $F^l_{k,max}$, and one of the tasks of the plant-wide coordinator is to make sure that this is satisfied. The maximum capacity for a unit ($F^l_{k,max}$) may change depending on disturbances (e.g. feed composition) and needs to be updated continuously. To estimate $F^l_{k,max}$ by using the already existing models in the local MPCs is discussed in Section 3. Section 4 discusses the coordinator MPC, including control design choices, model development, tuning issues and test runs. Experience from the implementation at the Kårstø site is summarized in Section 5. An extended version of this paper is found in Aske (2009, Ch. 6)

2. LOCAL MPC APPLICATIONS

Presently, all the local MPC applications for the coordinator are on two-product distillation columns. The main control objective for each distillation column is to control the quality of the distillate- (D) and bottoms (B) products. In addition, the column must be kept under surveillance to avoid overloading, where column differential pressure (Δp) is used as an indicator.

The local MPCs are configured with the following controlled variables (CVs), manipulated variables (MVs) and disturbance variables (DVs):

CV (set point + max constraint): Impurity of heavy key component in D.
CV (set point + max constraint): Impurity of light key component in B.
CV (max constraint): Column differential pressure (Δp).
MV: Reflux flow rate set point (L).
MV: Tray temperature set point (T_s).
DV: Column feed flow.

These MVs correspond to the local degrees of freedom (u^l) and the CVs correspond to the local outputs (y^l), see Fig. 1. Some of the columns have additional variables, but in principle, all the columns have the same control configuration.

The local MPC problems are solved at each sample time using a standard two-step approach, where first a steady-state problem is solved with constraint relaxation until the predicted final steady state is feasible, and then the "standard" dynamic MPC problem is solved with the possibly recalculated (reachable) set points and constraints. The high limit differential pressure has the highest priority, followed by impurity limits and then impurity set points. This priority hierarchy may leads to a relaxation of the impurity set points (and in worst case the limits) to avoid exceeding the differential pressure high limit. By using relaxation, the column can handle the given feed rate without flooding the column, but note that the exceeding the limits may result in an unsellable product. In the dynamic optimization part, constraints are handled by adding penalty terms to the objective function.

The local MPC applications are based on experimental step response models. The prediction horizon is 3 to 6 hours and the sample time is 1 minute.

3. ESTIMATE OF REMAINING CAPACITY

In this section, the procedure used by the local MPCs for estimating the remaining capacity in each unit (R_k) is explained.

The remaining capacity for unit k is the difference between the current feed F^l_k and the feasible maximum feed $F^l_{k,max}$

$$R_k = F^l_{k,max} - F^l_k \tag{1}$$

The feed to the local unit F^l_k is assumed to be a DV in the local MPC application. The maximum feed to the unit k is then easily obtained by solving an additional steady-state LP-problem subject to the present initial state, linear model equations and constraints used in the local MPC. $F^l_{k,max}$ is calculated using the end predictions (steady-state model) for the variables. This to include both past MV moves, disturbances and future MV moves for the local MPC. This indirectly assumes that the closed-loop response time for the local MPC is faster than for the coordinator. Note that $F^l_{k,max}$ can change due to updated measurements, disturbances (e.g. feed compositions changes), changes in the constraints and model changes in the local MPCs. The current feed to the unit (F^l_k) is measured, either by a flow transmitter or by a level controller output (valve opening) if a flow transmitter is not available.

The accuracy of the estimated remaining capacity depends on:

- The validity of the models used in the local application. The algorithm uses the end prediction and therefore the steady-state gain is in particular important.
- The appropriate use of gain scheduling for CV-MV pairs with larger nonlinearities. Here "gain scheduling" means that the model gain is updated (scaled) based on the current operation point.
- The CV constraints must reflect the true operational limits and the MV constraints must be reasonable.

Let us explain the first two points in more detail. An incorrect steady-state gain leads to a poor estimate of the remaining capacity and because the coordinator MPC has slow dynamics, it will take a long time before the feedback can correct for the error. A too high remaining capacity estimate lead to a oscillating behavior because of the long delays in the flow network. Another issue is that the operators will not trust the remaining capacity estimates if the estimates are far away compared to their own experience.

The remaining capacity estimate uses the CV constraints and not the CV set points. For a distillation column this implies that the distillate and bottoms quality constraints are used instead of the CV set points because set point deviations are acceptable if the alternative is feed reduction. This leads to an estimated capacity that is often larger than expected by the operators.

For units with several feeds, the LP optimization will maximize the feed with the smallest steady-state gain (smallest predicted effect on capacity), whereas the other feeds will go to zero. However, some feeds cannot be set to zero, because they are outlet from an upstream unit with no possibility for routing it elsewhere. In this case, the LP optimization is set to maximize the feed from the flow line the unit must process and the other feeds are held constant in the optimization.

Compressors are also included in the application, but at present there are no MPC applications implemented on these. To estimate the remaining capacity of the compressors one option could be to consider the percent load (given by the speed). However, it may not always be possible to reach 100% load due to other constraints, for instance the turbine exhaust gas temperature. To consider several constraints, we therefore use MPC applications with no control tasks, but with only CVs and DVs and the models between them to estimate the remaining capacity. A copressor stage consists of several copressors, but local control handles the distribution between parallel compressors (equal distance to the compressor control line), therefore is only one remaining capacity needed at each compressor stage.

At present, the estimates are based on experimental models. However, rigorous models for local units can also be used to predict the remaining capacity. This is attractive for units where experimental modelling is difficult, for example, due to nonlinearities. This illustrates the flexibility with this decomposition where the best available model can be used to predict the remaining capacity.

4. COORDINATOR MPC

4.1 Objective, variables and constraints

The Kårstø plant is shown in Fig. 2 where most of the CVs, MVs and DVs for the coordinator MPC are indicated. The coordinator MPC maximizes sum of the total plant feed which is the sum of the feeds to train 100 (T100), train 200 (T200), train 300 (T300), train 400 (T400) and the dew point control unit (DPCU). The application consists of:

- 6 MVs: 4 feed rates, 1 crossover, 1 feed split.
- 22 CVs: Remaining capacity of 12 distillation columns and 2 compressors stages, 7 other constraints plus the main objective: total plant feed with a high, unreachable set point with lower priority.
- 7 DVs: 3 feed rates, 2 feed compositions, 1 crossover, 1 feed split.

The "other" 7 CV constraints are related to the use of MVs, that is, levels constraints to avoid filling or emptying of buffer tanks and sump volumes, pressure constraints in the pipelines and pressure controller outputs. The CV "total plant feed" is the sum of the plant feeds and is given by

$$\begin{aligned}\text{TOTALFEED} = {} & 20\text{FC1001A} + 20\text{FC2001A} + 27\text{FC3108} \\ & + 27\text{FC3208} + 21\text{FC4125A} + 21\text{FC4225A} + 21\text{FC5219}\end{aligned} \tag{2}$$

where the variables are marked in Fig. 2. In general, the feeds could have different weighting, but at present, their weights are equal. Of the 22 CVs, only the total plant feed is set point controlled; the other CVs are constraints.

The MVs (throughput manipulators) are the feed rates, a crossover between parallel trains (from T100 to T300) and a feed split to T300. Other throughput manipulators that affect the CVs in the sub-application are included as DVs. Later, if the coordinator MPC is extended to the whole plant, most of these DVs will become MVs. The feed compositions (DVs) reflects the gas/liquid split, and determine the split between gas flow to the compressors and liquid flow to the fractionation and are estimated from flow- and temperature measurements.

The objective function in the SEPTIC MPC algorithm is quadratic, while the objective function for the the maximum throughput problem is linear. To obtain a quadratic objective function that fits directly into our quadratic MPC algorithm, we have used the common trick of introducing a quadratic set point deviation term with a high and unreachable set point with a lower priority than the capacity constraints. (Of course, the actual case function used by the coordinator MPC has additional terms and weights). The first step of the coordinator MPC solution will then result in a recalculated (reachable) set point for the total feed.

Each variable (CV, MV and DV) belongs to one or more sub-groups that will be deactivated if one critical variable in the sub-group is deactivated. For instance, if a local MPC application is turned off, the corresponding remaining capacity CV is deactivated, and this critical variable suspends the whole sub-group. By using this condition-based logic, the coordinator MPC can operate

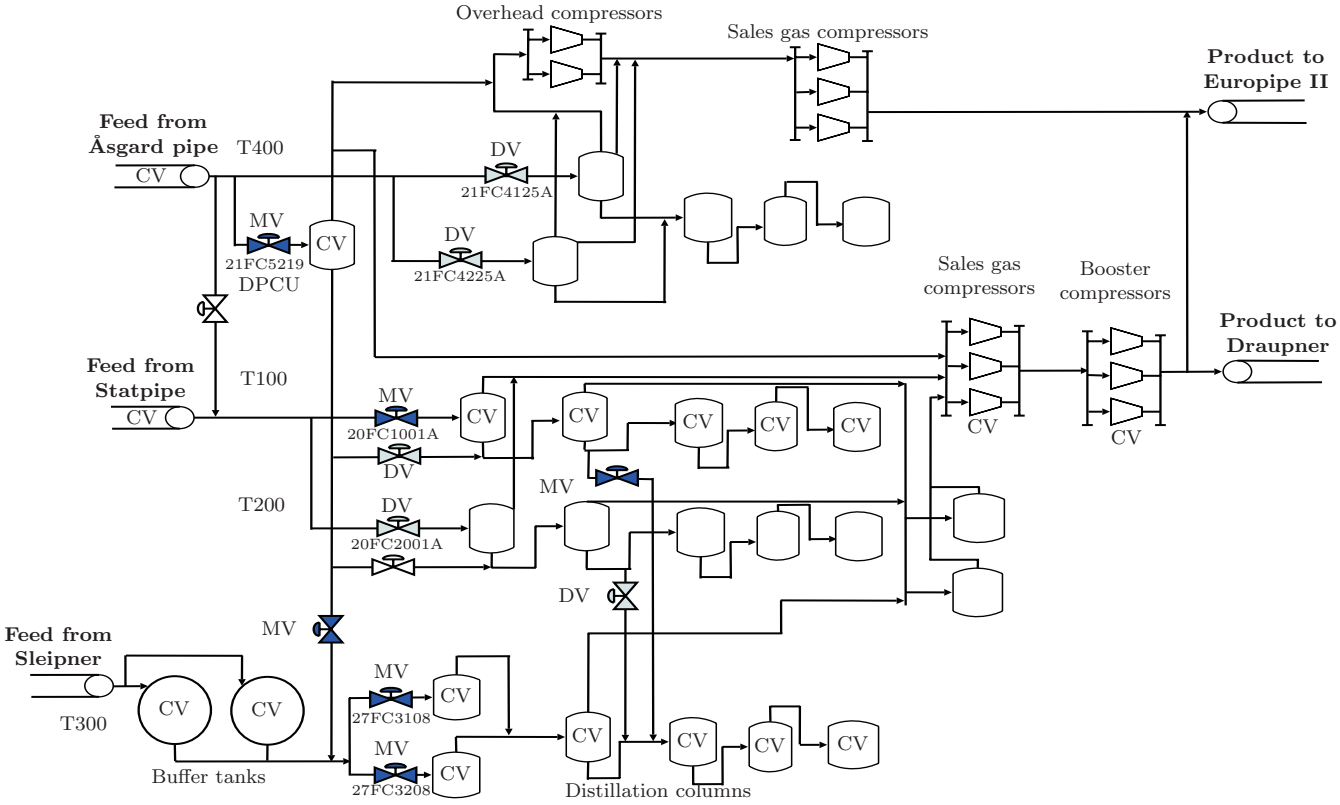


Fig. 2. Overview of the Kårstø plant, including the coordinator MPC variables.

even if parts of the plant are not running or not available for throughput maximization.

The decomposition requires that the coordinator receives three variables from each of the 12 local MPC applications:

- Estimated remaining capacity (value)
- Quality of the remaining capacity value (good/bad)
- Status of the local MPC (on/off)

If the estimated remaining capacity has a bad value, that is, the LP formulation is not feasible, then the status of the remaining capacity CV is set to ERROR and the corresponding MVs, given by the sub-grouping in the coordinator, are then suspended. If a local MPC application is deactivated, then the unit remaining capacity CV is set to OFF in the coordinator and the sub-group in the coordinator is suspended. The coordinator still runs, but the MVs in the sub-group are deactivated. This is done because we require that the local MPC application is active before the coordinator can manipulate on the corresponding unit feed rate.

4.2 Dynamic modelling for the coordinator MPC

The model for the coordinator MPC is a linear dynamic model for the flows through the plant network with the local MPC applications in service. The current implementation of the coordinator uses individual (SISO) step response models, or more precisely a single-input multiple-output representation of a multi-input multi-output system. The advantage with SISO models is that it is easy to adjust the models independently for input-output pairs. However, SISO models imply that the structure of the model is lost and, for instance, disturbances may not propagate as they would in a state-space model. The loss of structure leads to some additional variables around the DPCU.

The models are obtained from step tests and historical plant data. The steady-state gains found from step-tests are verified by calculating the gains using typical feed compositions.

The sampling time for the coordinator MPC is 3 minutes. The prediction and control horizon are set to 6 hours, whereas the longest response models reach steady state at approximately 4.5 hours.

4.3 Tuning the coordinator MPC

The tuning of the coordinator MPC is a trade-off between robustness and MV (e.g. feed) variations on one side and keeping the flows through the bottlenecks close to their maximum on the other side. The coordinator MPC was gradually operating in closed-loop and tuned in several tests in February 2008.

MV tuning From the early tests, it became clear that the trick of using a CV of total plant feed with a high, unreachable set point to maximize throughput, requires ideal values on the MV plant feeds to obtain satisfactory dynamic performance. The ideal values that are added to the MV plant feeds are high and unreachable with a lower priority than the total plant feed set point and have a low penalty on the deviation from the ideal value. The ideal

values are needed to avoid that *all* MVs that constitute the CV total plant feed (see (2)) are reduced dynamically to reach the new recalculated set point for CV total plant feed.

When ideal values (IV) for the MVs are introduced, the rate of change towards the ideal value is specified to obtain ramping rate independent of the penalty on the deviation from ideal value (Strand and Sagli, 2003). The ideal ramping rate is typical set to 500-750 kg/h. Maximum increase and decrease of the MV at each sample is chosen based on typically rate changes operators choose to implement.

CV tuning The most important tuning variables for the CVs are the penalties on constraint violation used in the dynamic step of the MPC algorithm. The constraint violation is "balanced" by using penalties on MV moves to obtain a satisfactory dynamic behavior when CV constraints are violated. Even though a CV constraint is violated, the use of MVs should not be too aggressive to avoid unnecessary throughput variations. Importantly, the CV constraints are not absolute because back off is included to handle disturbances and imperfect control. Specifically, the lower value of the remaining capacities is not set to zero, but rather to a positive back off value, $R_k^l > \text{back off}_k > 0$. The value of the back off is a tuning parameter decided by disturbance handling and model accuracy.

The coordinator MPC has four integrating CVs; two buffer volumes (levels) and two pipelines pressures. For an integrator, the horizon length is a tuning parameter. A shorter horizon length will give a larger slope and allow for larger feed rate changes. The integrating variables have a prediction horizon of 3 hours, which is half the prediction length to the other variables. The prediction horizon is shortened because it is likely that disturbances occur within the 6-hour period that counteracts the change in the integrated variable.

5. EXPERIENCE FROM IMPLEMENTATION

Some experiences from the implementation at the Kårstø site are summarized in this Section.

5.1 Estimate of remaining capacity

For distillation columns that frequently operate close to their capacity limit, the estimated capacity is generally good. For these units we have more experience in the actual operation range, and the models in the local MPC applications are typically obtained in this range. For some columns, the differential pressure is included in the remaining capacity calculation and this improves the estimate.

For control, the initial response for the models is most crucial to obtain good performance. For remaining capacity estimate, the steady-state model gain is most important. A systematic evaluation of the inferential models (estimators of product quality) and models in the local MPC applications is necessary to obtain satisfactory performance of the coordinator MPC. Since some of the local MPC applications were commissioned several years ago, a validation of the models was found necessary.

One observation is that when a large disturbance occurs, the predicted steady-state values may violate their limits and, if this violation is sufficiently large, the LP optimization does not find a feasible solution and the estimate of maximum capacity ($F_{k,max}^l$) fails. The end prediction values are in such cases often not reasonable because the MPC application assumes that the disturbances will maintain constant (possible reduced with a low-pass filter) throughout the prediction horizon, which is rarely the case.

To improve the estimation of remaining capacity, several approaches are used:

- With a known, measured, short-time disturbance: The maximum capacity ($F_{k,max}^l$) is held constant during the period of the disturbance. For example, this is used for the disturbances that occur at each dryer exchange.
- For each unit, a minimum value of the maximum capacity ($F_{k,max}^l$) is included.
- CV constraints included in the local MPCs that should not limit the throughput were replaced with wider constraints. This applies to "non-physical" constraint that may have been added in the MPC for tuning reasons.
- Gain scheduling is included for some differential pressure models.

The main structural weakness in the estimation of remaining capacity is that the LP solver may "give up" to find a solution because there is no possibility for relaxation of constraints. When the LP solver does not find a solution, it returns a "bad quality" value to the coordinator and its variable subgroup is turned off. It would be preferable that the coordinator finds the best possible solution instead of "giving up". This can be realized with a LP solver that includes relaxation of the constraints. This improvement of the LP algorithm is planned to be included in the future.

5.2 Experience with the coordinator MPC

A test run of the coordinator MPC from 07 Feb. 2008 is displayed in Fig. 3. The coordinator is turned on at $t = 18$ min and the coordinator starts to increase the feed to T100 (Fig. 3(a)) until the pipeline pressure in Statpipe reaches its low constraint (Fig. 3(b)). During this start-up period, the crossover flow ramps towards its ideal value (Fig. 3(c)). The remaining capacity in the butane splitter T100 reaches its low constraint (Fig. 3(d)) and the crossover increases again to avoid reduction in the throughput. However, the use of the crossover is "aggressive" and actually generates oscillations in the downstream remaining capacities. The model gain was almost doubled around $t = 250$ minutes and the crossover is now able to control the remaining capacity towards its low constraint. The adjustment of the model gain was based on comparing the model prediction (not shown) and actual value.

The accuracy of the estimate of remaining capacity for demethanizer T100 (Fig. 3(e)) was poor. The model gain from column feed to differential pressure was increased at $t = 320$ minutes, and the new value seems to give a more

correct estimate of the remaining capacity for the column. Again, the adjustment of the model gain was based on comparing model prediction and the actual value. Note that the remaining capacity of the demethanizer T100 became close to zero at about $t = 330$ min and the lower constraint value (back off) was increased at $t = 500$ min to obtain larger operation margins.

Feed composition changes are important disturbances and affect the remaining capacity to the units. The feed composition in the Statpipe (T100) (Fig. 3(f)) is rather stable until $t = 580$ min when the feed becomes significantly heavier and thereafter (at $t = 610$ min) significantly lighter. In this case, the coordinator uses the crossover (Fig. 3(c)) and the T100 feed rate (Fig. 3(a)) to control the remaining capacity for the butane splitter T100 (Fig. 3(d)) at its constraint.

When in closed loop, the coordinator MPC manipulates directly on the plant production. This directly involves the shift manager at Kårstø and close cooperation with the manager at the gas pipeline network (operated by another company) is necessary. The plant is operated by three control panels, so a close dialog between the operator personnel and the shift manager is crucial. The coordinator MPC introduces a "new way of thinking" for both operators and shift managers. The coordinator introduces the back off constraint as a new handle in addition to pressure pipeline constraints, instead of the feed valves.

ACKNOWLEDGEMENTS

The implementation was performed together with Kjetil Meyer, Roar Sørensen, shift managers and operating personnel at the Statpipe and Sleipner panels. All are gratefully acknowledged. The plant operator Gassco and technical services provider StatoilHydro are acknowledged for plant data.

REFERENCES

Aske, E. (2009). *Design of plantwide control systems with focus on maximum throughput*. Ph.D. thesis, NTNU, Trondheim, Norway. Available at the homepage of S. Skogestad.

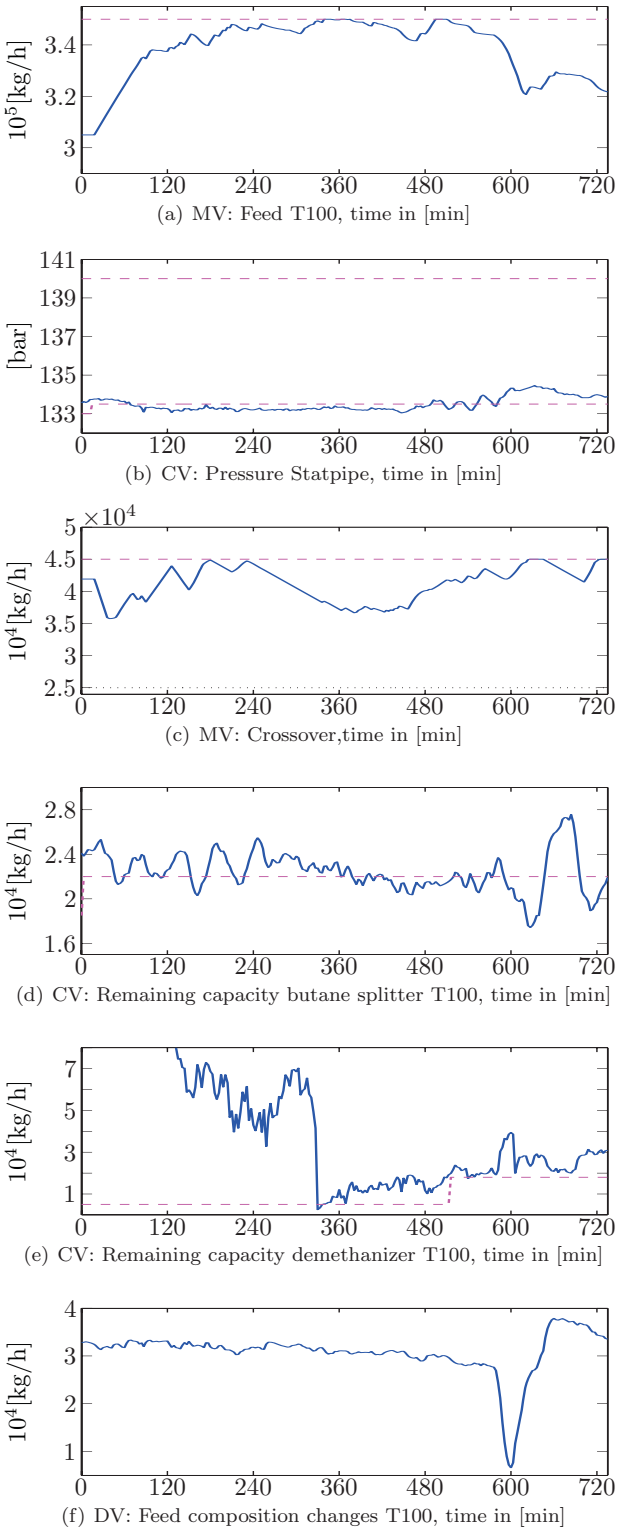
Aske, E., Strand, S., and Skogestad, S. (2008). Coordinator MPC for maximizing plant throughput. *Comput. Chem. Eng.*, 32(1-2), 195–204.

Engell, S. (2007). Feedback control for optimal process operation. *J. Proc. Control*, 17, 203–219.

Marlin, T.E. and Hrymak, A.N. (1997). Real-time operations optimization of continuous processes. In J. Kantor, C. Garcia, and B. Carnahan (eds.), *CPC-5*, 156–164. Lake Tahoe, Nevada.

Strand, S. (1991). *Dynamic Optimization in State-Space Predictive Control Schemes*. Ph.D. thesis, Norwegian Institute of Technology, Trondheim, Norway.

Strand, S. and Sagli, J. (2003). MPC in Statoil - Advantages with in-house technology. *ADCHEM, Hong Kong, 2004*, 97–103.



(a) MV: Feed T100, time in [min]

(b) CV: Pressure Statpipe, time in [min]

(c) MV: Crossover,time in [min]

(d) CV: Remaining capacity butane splitter T100, time in [min]

(e) CV: Remaining capacity demethanizer T100, time in [min]

(f) DV: Feed composition changes T100, time in [min]

Fig. 3. From test run 07 Feb. 2008: $t = 18$ min: turn on, $t = 250$ min: change in model gain for crossover, $t = 320$ min: change in model grain for feed to differential pressure in the demethanizer, $t = 580$ and $t = 610$ min: feed composition change. MV and CV values (solid), high and low limits (dashed) and ideal values (dotted).

Coordination of Distributed Model Predictive Controllers for Constrained Dynamic Processes ⋆

Natalia I. Marcos *, **J. Fraser Forbes** * and **Martin Guay** **

* *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, Canada T6G 2V4*
** *Department of Chemical Engineering, Queen's University, Kingston, Ontario, Canada K7L 3N6*

Abstract: In this paper, a coordinated-distributed model predictive control (MPC) scheme is presented for large-scale discrete-time linear process systems. Coordinated-distributed MPC control aims at enhancing the performance of fully decentralized MPC controllers by achieving the plant-wide optimal operations. The 'price-driven' decomposition-coordination method is used to adjust the operations of the individual processing units in order to satisfy an overall plant performance objective. Newton's method, together with a sensitivity analysis technique, are used to efficiently update the price in the price-driven decomposition-coordination method. The efficiency of the proposed control scheme is evaluated using a model of a fluid catalytic cracking process.

Keywords: Decomposition-coordination methods; Large-scale optimization; Optimal control theory.

1. INTRODUCTION

Since the late seventies, the design of chemical processes has evolved towards integrated operations that have increased plant's efficiency. The improvement in the design of chemical processes included, among other things, energy and mass integration, and the use of recycle streams. As a result, processes became more complex and processing units became more tightly interconnected. Control of such integrated large-scale processes has been typically performed with *decentralized* schemes because of the difficulties in implementation and maintenance of *centralized* control frameworks.

Centralized and decentralized control are two distinct control strategies. In centralized control, no real distinction is made among processing units. The centralized control framework is formulated as a monolithic control problem that incorporates all process variables with no unit-level decomposition. While a centralized strategy can lead to optimal plant-wide performance, it presents some disadvantages (e.g., the large-dimensionality of the control problem and lack of flexibility in terms of operation and maintenance), which make centralized control unsuitable for industrial processes. In decentralized control, each engineering unit is optimized separately by neglecting the interactions with the other units. The decentralized approach is the most commonly used in the industry because of its robustness and its resiliency to systems failures. Nevertheless, decentralized control does not generally lead to the desired plant-wide optimal operations (Lu (2003); Sun and El-Farra (2008)).

⋆ This work is supported by Natural Sciences and Engineering Research Council of Canada (NSERC) and Alberta Ingenuity.

A compromise between centralized and decentralized control is desired in order to improve plant operations. *Distributed* control has emerged as a promising control strategy that can lead to the plant-wide optimal operations, while keeping manageable controllers for each subunit in the plant. In the distributed control framework, it is assumed that each subsystem computes its own optimal solution while considering all or certain degree of interactions with the other subsystems. To attain the desired control performance, information related to each subsystems' optimal solutions is generally exchanged among the subsystems. In this work, we present a coordinated-distributed model predictive control (CDMPC) framework for constrained dynamic processes. In CDMPC control, data is exchanged with each individual MPC controller via a 'coordinator' or 'master'.

1.1 Distributed MPC Control

Distributed MPC control has attracted the attention of many researchers in recent years. Dunbar and Murray (2004) formulated MPC platforms for nonlinear interacting subsystems (multi-vehicle formations) whose state variables are coupled in a single objective function. For linear interconnected systems, Venkat et al. (2005) proposed a communication-based MPC that can converge to a Nash equilibrium. The communication-based MPC was further improved by a cooperation-based MPC that leads to the Pareto optimal feasible solution. Cheng et al. (2008, 2007) proposed a coordinated scheme for MPC steady state target calculation based on Dantzig-Wolfe decomposition and price-driven coordination methods, respectively.

The main contribution of this work is to propose the *price-driven* decomposition-coordination algorithm, as described in Cheng et al. (2007), for the control of constrained process systems whose dynamics are represented by discrete-time models. The CDMPC control scheme presented in this paper achieves the centralized optimal operations and can be implemented when step-response models are available for the process. Since our control formulation uses models obtained from step-test data, it does not need estimation of unavailable process variables (as it might be required when formulating MPC controllers based on state-space models). Furthermore, the proposed CDMPC control scheme allows for bias correction in the predicted outputs through feedback.

An illustration of CDMPC is shown in Fig. 1. The price-driven decomposition-coordination method is used in the formulation of the CDMPC controllers. In the price-driven decomposition-coordination method, the coordinator sets up a price, 'p', for the subsystems' interacting variables (Fig. 1). The price provided by the coordinator is then

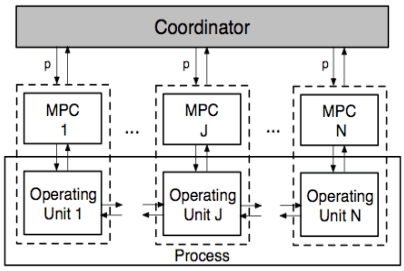


Fig. 1. Illustration of CDMPC Control

adjusted to alter the subunits' calculated control actions towards the overall plant optimum. In this work, the price, p, is updated based on Newton's method. An iterative procedure is established between the coordinator and the subunits until the desired plant-wide optimal solution is achieved.

2. CDMPC CONTROL FOR DYNAMIC PROCESS SYSTEMS

In this section, the CDMPC control scheme is presented. Since we consider the centralized performance as the ideal benchmark, we begin the CDMPC control formulation by decomposing the centralized control problem into N smaller subproblems that are easier to solve. Then, an efficient mechanism is used to achieve the same solution as the one obtained in the centralized control problem.

2.1 Process Model

Consider the overall plant process, modelled by step-response coefficients:

$$y_z(k+l) = \sum_{h=1}^{T-1}\sum_{w=1}^{r} S_{zw,h}\Delta u_w(k+l-h) + \sum_{w=1}^{r} S_{zw,T}\ u_w(k+l-T), \quad \forall z = 1,\ldots,m, \tag{1}$$

where y_z $(\forall z = 1,...,m) \in \Re^m$ denote the process outputs; $u_w \in \Re^r$ and $\Delta u_w \in \Re^r$ $(\forall w = 1,...,r)$ denote the manipulated variables and the change in the manipulated variables, respectively. The coefficients $S_{11,h},...,S_{mr,h}$ represent the step-response coefficients for h^{th} time step. The step-response weight $S_{11,h}$ is the coefficient between Δu_1 and output y_1 for the h^{th} time step. In a similar manner, $S_{mr,h}$ is the coefficient between Δu_r and output y_m for the h^{th} time step.

2.2 Centralized MPC Formulation

For the centralized MPC implementation, it is convenient to arrange process model (1) in a matrix form as following:

$$\hat{Y}(k+1) = \mathcal{S}\Delta\hat{U}(k) + Y^0(k+1) + \hat{D}(k+1), \tag{2}$$

where the output variables, input variables and change in input variables predicted along the prediction horizon H_p and control horizon H_u are defined as:

$$\begin{cases} \hat{Y}(k+1) = [\hat{y}(k+1|k)^\top,...,\hat{y}(k+H_p|k)^\top]^\top, \\ \hat{y}(.) = [\hat{y}_1(.),...,\hat{y}_m(.)]^\top, \\ \Delta\hat{U}(k) = [\Delta\hat{u}(k|k)^\top,...,\Delta\hat{u}(k+H_u-1|k)^\top]^\top, \\ \Delta\hat{u}(.) = [\Delta\hat{u}_1(.),...,\Delta\hat{u}_r(.)]^\top, \hat{u}(.) = [\hat{u}_1(.),...,\hat{u}_r(.)]^\top. \end{cases} \tag{3}$$

The $m \times H_p$ vector of unforced responses $Y^0(k+1)$ is:

$$\begin{cases} Y^0(k+1) = [y^0(k+1)^\top,...,y^0(k+H_p)^\top]^\top, \\ y^0(.) = [y_1^0(.),...,y_m^0(.)]^\top. \end{cases} \tag{4}$$

The vector $\hat{D}(k+1)$ has been incorporated in (2) to correct through feedback the discrepancies between the measured and predicted outputs. The vector $\hat{D}(k+1)$ is defined as:

$$\hat{D}(k+1) = \underbrace{[I_m,...,I_m]^\top}_{H_p\ times}[y(k) - \hat{y}(k|k-1)],$$

where I_m is the $m \times m$ identity matrix. It is assumed that the difference between the measured and predicted outputs at time k remains constant throughout the prediction horizon.
In (2), the matrix of step-response coefficients $\mathcal{S}$ is defined as:

$$\mathcal{S} = \begin{bmatrix} S_1 & 0 & \cdots & 0 \\ S_2 & S_1 & 0 & 0 \\ \vdots & \vdots & \ddots & 0 \\ S_{H_u} & S_{H_u-1} & \cdots & S_1 \\ \vdots & \vdots & \ddots & \vdots \\ S_{H_p} & S_{H_p-1} & \cdots & S_{H_p-H_u+1} \end{bmatrix}, \tag{5}$$

where S_h is the $m \times r$ matrix of step-response coefficients for the h^{th} time step $(\forall h = 1,...,H_p)$:

$$S_h = \begin{bmatrix} S_{11,h} & S_{12,h} & \cdots & S_{1r,h} \\ \vdots & \cdots & \cdots & \vdots \\ S_{m1,h} & \cdots & \cdots & S_{mr,h} \end{bmatrix}. \tag{6}$$

The centralized MPC controller is formulated to minimized the following objective function:

$$\min_{\hat{Y},\Delta\hat{U}} \quad \mathcal{J} = \frac{1}{2}\Big((Y_{sp} - \hat{Y}(k+1))^\top Q(Y_{sp} - \hat{Y}(k+1)) + \Delta\hat{U}(k)^\top R\Delta\hat{U}(k)\Big) \tag{7}$$

subject to:

$$\begin{cases} \text{Process model (2)-(6), and} \\ \hat{u}(k+l|k) = \sum_{h=0}^{l} \Delta\hat{u}(k+h|k) + u(k-1), \\ \Delta\hat{u}(k+h|k) = 0, \qquad H_u \le h \le H_p - 1, \\ y_{min} \le \hat{y}(k+l+1|k) \le y_{max}, \\ u_{min} \le \hat{u}(k+l|k) \le u_{max}, \\ \Delta u_{min} \le \Delta\hat{u}(k+n|k) \le \Delta u_{max}, \\ \forall l = 0,\ldots,H_p-1, \quad \text{and} \quad \forall n = 0,\ldots,H_u-1, \end{cases} \tag{8}$$

where Y_{sp} is the vector of desired set-points, $Q = diag\{Q(l+1)\}$ and $R = diag\{R(n)\}$ are positive definite matrices of appropriate dimensions.

2.3 Decomposition of Centralized MPC Formulation

We propose a decomposition of the overall optimization problem (7)-(8) into N subproblems 'i'. We consider that the plant dynamics and constraints can be decomposed into N subunits, followed by a block decomposition of the tuning matrices Q and R. As a result of the centralized problem decomposition, each subunit i ($\forall i : 1,\ldots,N$) solves its own optimization problem given by:

$$\min_{\hat{Y}_i,\Delta\hat{U}_i,\hat{V}_i} \mathcal{J}_i = \frac{1}{2}\Big((Y_{i\_sp} - \hat{Y}_i(k+1))^\top Q_{ii}(Y_{i\_sp} - \hat{Y}_i(k+1)) + \Delta\hat{U}_i(k)^\top R_{ii}\Delta\hat{U}_i(k)\Big) + p^\top\Theta_i Z_i(k) \tag{9}$$

subject to:

$$\hat{Y}_i(k+1) = \mathcal{S}_{ii}\Delta\hat{U}_i(k) + \hat{V}_i(k) + Y_i^0(k+1) + \hat{D}_i(k+1), \tag{10}$$

$$\hat{V}_i(k) = \mathcal{S}_{ij}\Delta\hat{U}_j(k), \quad \forall j \ne i, \tag{11}$$

and

$$\begin{cases} \hat{u}_i^s(k+l|k) = \sum_{n=0}^{l} \Delta\hat{u}_i^s(k+h|k) + u_i^s(k-1), \\ \Delta\hat{u}_i^s(k+h|k) = 0, \qquad H_u \le h \le H_p - 1, \\ y_{i\_min}^s \le \hat{y}_i^s(k+l+1|k) \le y_{i\_max}^s, \\ u_{i\_min}^s \le \hat{u}_i^s(k+l|k) \le u_{i\_max}^s, \\ \Delta u_{i\_min}^s \le \Delta\hat{u}_i^s(k+n|k) \le \Delta u_{i\_max}^s, \\ \forall l = 0,\ldots,H_p-1, \quad \text{and} \quad \forall n = 0,\ldots,H_u-1. \end{cases} \tag{12}$$

The vector $\hat{y}_i^s$ is a subset of the plant predicted outputs ($\hat{y}_i^s \subset [\hat{y}_1,...,\hat{y}_m]$) and represents the predicted output variables of subsystem i, $\forall i = 1,...,N$. Similarly, the vector $\hat{u}_i^s$ is a subset of the plant predicted inputs ($\hat{u}_i^s \subset [\hat{u}_1,...,\hat{u}_r]$) and represents the predicted input variables of subsystem i, $\forall i = 1,...,N$. According to the proposed decomposition, the predicted change in input variables and predicted outputs can be arranged in vector form as $\Delta\hat{U}_i(k) = [\Delta\hat{u}_i^s(k|k)^\top,...,\Delta\hat{u}_i^s(k+H_u-1|k)^\top]^\top$ and $\hat{Y}_i(k+1) = [\hat{y}_i^s(k+1|k)^\top,...,\hat{y}_i^s(k+Hp|k)^\top]^\top$, respectively.

In (10) and (11), the matrix $\mathcal{S}_{ii}$ corresponds to the step-response coefficient matrix between $\Delta\hat{U}_i(k)$ and predicted output variables $\hat{Y}_i(k+1)$; while the matrix $\mathcal{S}_{ij}$ corresponds to the step-response coefficient matrix between the interacting variables $\Delta\hat{U}_j(k)$ and predicted output variables $\hat{Y}_i(k+1)$. The matrices $\mathcal{S}_{ii}$ and $\mathcal{S}_{ij}$ can be obtained by decomposing matrices (5) and (6) into N subsystems.

Finally, the variables $\hat{V}_i(k)$ represent the interacting or linking variables among the different subunits in the plant. The interacting variables account for the effect that inputs from unit j have on unit i, with $i \ne j$. In the objective function (9), $Z_i(k) = [\hat{Y}_i(k+1)^\top, \Delta\hat{U}_i(k)^\top, \hat{V}_i(k)^\top]^\top$ represents the vector of decision variables for subunit i; the matrix Θ_i is the coefficient matrix for the linking variables, which is constructed according to (10) and (11), and p is a price vector provided by the coordinator.

For simplicity, we re-write problem (9)-(12) as:

$$\min_{Z_i} \quad \mathcal{J}_i = \frac{1}{2}\Big(Z_i(k)^\top \Upsilon_i Z_i(k)\Big) + \Phi_i^\top Z_i(k) + p^\top\Theta_i Z_i(k) \tag{13}$$

subject to:

$$\begin{cases} B_i^{eq} Z_i(k) = b_i^{eq}, \\ B_i^{ineq} Z_i(k) \le b_i^{ineq} \end{cases} \qquad \forall i = 1,...,N. \tag{14}$$

The optimization problem (13)-(14) can be straightforwardly obtained by arranging (9)-(12) in a matrix form for the entire prediction and control horizons. The optimization problem described by (13)-(14) forms part of the price-driven decomposition coordination method. The price-driven decomposition-coordination method was discussed in Jose and Ungar (2000, 1998) to solve algebraic optimization problems such as resource allocation or auction problems. This method was successfully adapted and implemented in Cheng et al. (2007) to solve the MPC steady-state target calculation problem. In this work, we use the price-driven method to solve MPC dynamic calculation problems.

2.4 Coordination of CDMPC Controllers

In the previous section, a decomposition of the overall problem into N smaller subproblems was presented. A key step in the formulation of CDMPC controllers is to design an efficient coordination mechanism that ensures convergence of the distributed optimal solutions to the centralized optimum. In this section, we extend the results obtained in Cheng et al. (2007) for the MPC steady-state target calculation to the MPC dynamic calculation.

As discussed in Jose and Ungar (2000), a large-scale problem:

$$\max_{z_1,\ldots,z_n} \quad \sum_i^n f_i(z_i)$$

$$\text{s.t.} \quad \sum_i^n R_i(z_i) \le \bar{R}, \qquad z_i \in \Omega_i,$$

with $z_i \in \Re^{n_i}$ decision variables, R_i vector of shared resources, and vector $\bar{R}$ representing the availability of shared resources, can be decomposed into N subproblems:

$$\max_{z_i \in \Omega_i} f_i(z_i) - (p + qR_i(z_i))^\top R_i(z_i). \quad (15)$$

In (15), 'p' represents the price vector, and the variable q is a nonnegative scalar that could be assumed to be zero for quadratic programming problems. In this work, we assume $q = 0$.

It was shown in Jose and Ungar (2000) that, when the subproblems present concave objective functions and compact convex feasible sets, they can be successfully coordinated. Moreover, at equilibrium, the following condition is satisfied:

$$\Delta R(p,q) = \sum_i R_i(p,q) - \bar{R} \leq 0,$$

with $p^\top(\Delta R(p,q)) = 0$, and $p \geq 0$.

Coordination of subproblem (13)-(14) for $i : 1, ..., N$ to achieve the plant-wide optimal solution can be performed by using an efficient price-update technique, such as Newton's method. Based on Newton's method, the price vector can be updated as follows (Cheng et al. (2007)):

$$p^{[\kappa+1]} = p^{[\kappa]} - \alpha J^{-1} \Delta R^{[\kappa]}, \quad (16)$$

provided that the matrix J is invertible. In the price update mechanism (16), the superscripts '$[\kappa]$' and '$[\kappa+1]$' denote the iteration steps; α is the step size in Newton's method, $\Delta R^{[\kappa]} = \Delta R(p,q)$, and J can be calculated as:

$$J = \frac{d\Delta R^{[\kappa]}}{dp^{[\kappa]}} = \sum_i \frac{dR_i^{[\kappa]}}{dp^{[\kappa]}}. \quad (17)$$

For the problem formulation described by (13)-(14), the shared resources or linking constraints are defined as $R_i^{[\kappa]} = \Theta_i Z_i^{[\kappa]}$, with $Z_i^{[\kappa]}$ representing the decision variables at each iteration step 'κ'. Therefore, the Jacobian matrix J in (17) becomes:

$$J = \sum_i \frac{dR_i^{[\kappa]}}{dp^{[\kappa]}} = \sum_i \Theta_i \frac{dZ_i^{[\kappa]}}{dp^{[\kappa]}}. \quad (18)$$

The Jacobian matrix (18) requires information of the sensitivity matrix $dZ_i^{[\kappa]}/dp^{[\kappa]}$; that is, in order to efficiently adjust the price vector, the coordinator should be aware of how the price affects the decision variables $Z_i^{[\kappa]}$ at each iteration. A sensitivity analysis was proposed in Wolbert et al. (1994) for an algebraic optimization of a process flowsheet, and it was extended in Cheng et al. (2007) for the MPC steady-state target calculation. This approach can be followed to solve problem (18). By performing a sensitivity analysis, the matrix $dZ_i^{[\kappa]}/dp^{[\kappa]}$ can be calculated. This requires solving the following system of equations:

$$\Gamma_i \begin{bmatrix} \nabla_p Z_i(k) \\ \nabla_p \lambda_i \\ \nabla_{pA} \mu_i \\ \nabla_{pI} \sigma_i \end{bmatrix} = - \begin{bmatrix} \Theta_i^\top \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad (19)$$

where

$$\Gamma_i = \begin{bmatrix} \Upsilon_i & B_i^{eq\top} & {}_A B_i^{ineq\top} & 0 \\ B_i^{eq} & 0 & 0 & 0 \\ {}_A B_i^{ineq} & 0 & 0 & 0 \\ {}_I B_i^{ineq} & 0 & 0 & I \end{bmatrix}, \quad (20)$$

assuming that Γ_i is full-rank. We refer the reader to Cheng et al. (2007) for a detailed derivation of equations (19) and (20).

2.5 Implementation of CDMPC Control Scheme

In the traditional MPC implementation, a control action sequence is determined at each sampling interval by optimizing an objective criterion over a finite-time horizon. Only the first control signal is applied to the process, while the rest of the calculated control inputs are discarded (Camacho and Bordons (1999); Maciejowski (2002)). At the next sampling interval, new process measurements are available and the optimization is repeated to calculate a new control action sequence.

In a CDMPC control platform, the coordinator imposes an extra step to the traditional MPC implementation. Before the control input is applied to the process, the control action calculated by each distributed MPC controller needs to converge to the optimal centralized control action. Convergence of the CDMPC solutions to the centralized performance can be achieved by allowing the coordinator to iteratively adjust the price vector, and therefore the optimal solution of each subsystem, according to the plant-wide objective.

Implementation of the CDMPC controllers is carried out according to the following steps:

(1) **Initialization:** The coordinator sets up an initial price vector $p^{[\kappa]}$ for the interacting variables ($\Theta_i Z_i$, $\forall i = 1, ..., N$) and sends that information to every subsystem.
(2) **Optimization performed by each subsystem:** Based on the price provided by the coordinator, each subsystem solves its own optimization problem (13)-(14) and calculates the resources $R_i^{[\kappa]} = \Theta_i Z_i^{[\kappa]}$; as well as $dZ_i^{[\kappa]}/dp^{[\kappa]}$, according to (19)-(20). This information is communicated back to the coordinator.
(3) **Price update:** The coordinator gathers the information from each subsystem; it calculates $\Delta R^{[\kappa]}$, and J given by (18). Then, the coordinator determines the step size α (with $0 < \alpha \leq 1$) and updates the price vector $p^{[\kappa]}$ as per (16). The new price vector is informed to each subsystem.
(4) **Iteration until convergence:** Steps (2)-(3) are repeated until convergence of the price-driven decomposition coordination algorithm. Convergence of the algorithm is achieved when $||\Delta R^{[\kappa]}|| \leq \epsilon$, where ϵ is a tolerance error.
(5) **Implementation of control action:** Once the decomposition-coordination algorithm converges, the control actions calculated for the first sampling interval are implemented in each subsystem and the optimization problem (steps (1)-(4)) is initiated again for the next receding horizon.

3. SIMULATION EXAMPLE

In this section, a case study is performed to illustrate the effectiveness of the proposed algorithm. We consider a fluid

catalytic cracking (FCC) process given in Grosdidier et al. (1993). A diagram of the FCC system is shown in Fig. 2.

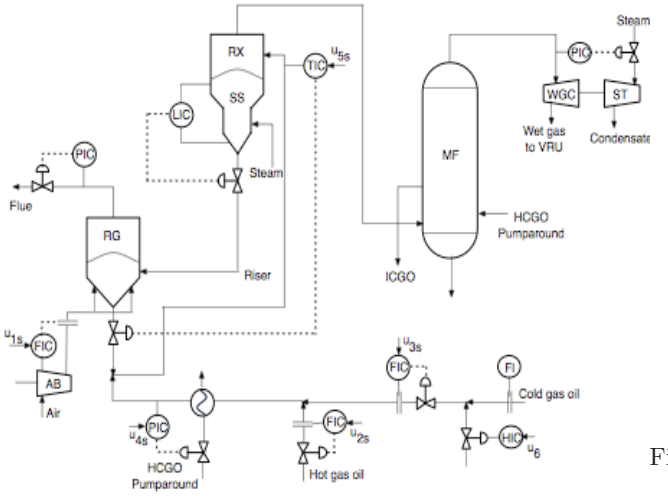


Fig. 2. FCC process (Grosdidier et al. (1993))

In the FCC unit, gas oil is converted into hydrocarbons of shorter chains. A description of the FCC process, together with the limit values for the controlled and manipulated variables are given in Grosdidier et al. (1993). The model of the FCC process, as well as, the models of the regulatory controllers are shown in tables 1 and 2, respectively. The continuous-time transfer function models were obtained through identification analysis of step-test data and include seven outputs and six inputs. The transfer function matrix for the overall process, including the models for the regulatory controllers, can be obtained by multiplying each transfer function model in table 1 by the corresponding input model in table 2, except for transfer functions between $y5-u5$ and $y6-u5$, which do not require that multiplication (Grosdidier et al. (1993)). To implement the CDMPC controllers, step-response models were obtained based on the process dynamics given in tables 1 and 2. The sampling interval used for simulations was 1 $[min]$.

3.1 Simulation Results

We begin by decomposing the centralized problem into two subsystems. The first subsystem includes outputs y_1 to y_3 and inputs u_1 to u_3, while the second subsystem includes outputs y_4 to y_7 and inputs u_4 to u_6. The following parameters were used in the simulation study: weighting matrices $Q(l+1) = diag\{5; 10; 0.001; 5; 5; 5; 0.001\}$, $R(n) = diag\{100; 100; 100; 100; 100; 100\}$, for $l = 0, ..., H_p - 1$ and $n = 0, ..., H_u - 1$. The weighting matrices $Q(.)$ and $R(.)$ are decomposed as $Q_{11}(.) = diag\{5; 10; 0.001\}$ and $R_{11}(.) = diag\{100; 100; 100\}$ for the first subsystem, and $Q_{22}(.) = diag\{5; 5; 5; 0.001\}$ and $R_{22}(.) = diag\{100; 100; 100\}$ for the second subsystem. The prediction horizon H_p and the control horizon H_u considered for the simulation are 50 and 5, respectively.

A set-point change of 0.5 was performed in output y_1 at initial time, while the targets for the remaining outputs were kept at the origin. The results of the simulation are

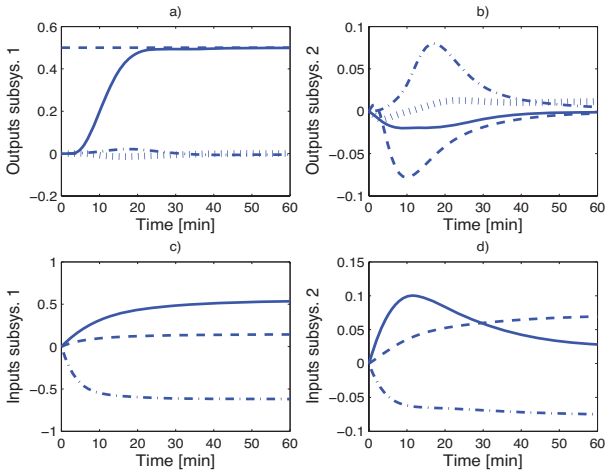


Fig. 3. a) Output variables for subsystem 1: set-point for y_1 (dashed line), y_1 (solid line), y_2 (dotted line), y_3 (dash-dot line); b) Output variables for subsystem 2: y_4(dash-dot line), y_5 (dashed line), y_6 (solid line), y_7 (dotted line); c) Input variables for subsystem 1:u_1 (dashed line), u_2 (dashed-dot line), u_3 (solid line); d) Input variables for subsystem 2: u_4 (dashed-dot line), u_5 (solid line), u_6 (dashed line)

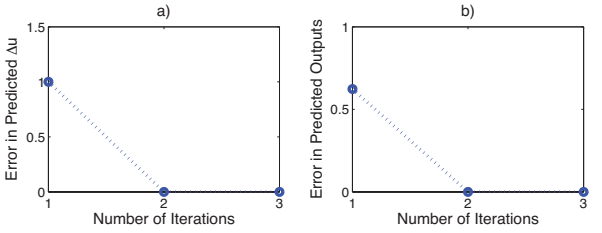


Fig. 4. a) Error in predicted change in input variables $\Delta\hat{U}(k)$; b) Error in predicted output variables $\hat{Y}(k+1)$

presented in Fig. (3)-(4). The closed-loop performance of the CDMPC controllers for subsystems 1 and 2 is shown in Fig. 3, where the trajectories are plotted in deviation variables. It can be seen in Fig. (3a)-(3d) that the CDMPC controllers provide a good performance since output y_1 achieves the new set-point and outputs y_2 to y_7 are stabilized at their new steady-state optimal values. Fig. (4a)-(4b) show the normalized errors of the predicted inputs ($||\Delta\hat{U}_{CDMPC} - \Delta\hat{U}_{cen}||/||\Delta\hat{U}_{cen}||$) and predicted outputs ($||\Delta\hat{Y}_{CDMPC} - \Delta\hat{Y}_{cen}||/||\Delta\hat{Y}_{cen}||$) for the optimization performed at the first sampling time. These prediction errors are calculated as the difference between the CDMPC optimal solutions and the optimal solutions calculated with a centralized MPC controller. It can be observed in Fig. (4a)-(4b) that the solutions achieved with the CDMPC controllers converge to the centralized performance within 2 iterations. The fast convergence observed in this simulation study confirms the results reported in Cheng et al. (2007) when using Newton's method as price-adjustment algorithm for the MPC steady-state target calculation. In the numerical simulations performed for the FCC unit, the same fast solution convergence (2 iterations) was observed within each control execution.

Table 1. FCC process models

	u_1	u_2	u_3	u_4	u_5	u_6
y_1	$\frac{0.097(1.7s+1)e^{-2s}}{19s^2+6.5s+1}$	$\frac{-0.87e^{-2s}}{13s^2+4.9s+1}$	$\frac{-0.092(0.25s+1)e^{-3s}}{3.7s^2+4.7s+1}$	$\frac{0.026e^{-7s}}{12s+1}$	$\frac{-0.074(4.8s+1)}{9.3s^2+3.4s+1}$	$\frac{-(0.48s)e^{-12s}}{(6s+1)(8s+1)}$
y_2	0	$\frac{0.55e^{-4s}}{27s^2+8.7s+1}$	$\frac{0.55e^{-4s}}{10s^2+4.9s+1}$	0	$\frac{0.74(1.7s+1)e^{-2s}}{11s^2+7.3s+1}$	$\frac{0.36e^{-11s}}{33s^2+6.5s+1}$
y_3	0	$\frac{0.14e^{-11s}}{46s^2+8.5s+1}$	$\frac{0.14e^{-6s}}{46s^2+8.5s+1}$	0	$\frac{0.27(16s+1)}{53s^2+23s+1}$	$\frac{0.015(12s+1)e^{-9s}}{66s^2+27s+1}$
y_4	0	$\frac{0.25e^{-11s}}{17s^2+7s+1}$	$\frac{0.25e^{-7s}}{3s+1}$	0	$\frac{0.70}{3s+1}$	$\frac{0.079(6.3s+1)e^{-10s}}{24s^2+12s+1}$
y_5	0	$\frac{0.66e^{-s}}{2.5s+1}$	$\frac{0.66e^{-s}}{2.5s+1}$	$\frac{-0.9e^{-10s}}{6s+1}$	$\frac{1}{2s+1}$	$\frac{-0.54e^{-11s}}{9s+1}$
y_6	0	$\frac{-0.84e^{-s}}{6.1s+1}$	$\frac{-0.90}{1.5s+1}$	$\frac{0.35e^{-10s}}{5s+1}$	$\frac{-(0.64s+1)}{13s^2+7s+1}$	$\frac{0.23(0.5s+1)e^{-14s}}{3.6s^2+11s+1}$
y_7	0	$\frac{0.81}{6s+1}$	$\frac{0.90}{s+1}$	$\frac{-0.35e^{-10s}}{5s+1}$	0.80	$\frac{-0.26e^{-18s}}{7.1s+1}$

Table 2. Models between regulatory controller set-points us_i and process inputs u_i, for $i = 1, ..., 6$

(us_1, u_1)	(us_2, u_2)	(us_3, u_3)	(us_4, u_4)	(us_5, u_5)	(us_6, u_6)
$\frac{1}{(0.75s+1)(4.5s+1)}$	$\frac{1}{(s+1)}$	$\frac{1}{1.7s^2+2.1s+1}$	$\frac{(3.3s+1)e^{-s}}{40s^2+13s+1}$	$\frac{(0.64s+1)}{13s^2+7s+1}$	1

Remark: For the ease of presentation, we decomposed the overall FCC process into two subsystems of similar dimensions. Nevertheless, the CDMPC control scheme can be applied to N number of subsystems of different dimensions. As future work, we will evaluate the efficiency of the CDMPC control scheme on process systems that include more subunits and there is a significant mismatch in the size of the subunits.

4. CONCLUSION

In this paper, we presented a coordinated-distributed model predicted control scheme for constrained dynamic processes. The CDMPC control framework improves the performance of decentralized controllers by achieving the overall plant-wide optimal operations.

An important advantage of CDMPC controllers is the simplicity in the control scheme, which does not require a radical new configuration of the decentralized MPC controllers operating in the plant. The upgrade from the existing decentralized MPC controllers to CDMPC controllers only involves small modifications in the control formulation of each subsystem and the addition of a coordinator.

The price-driven decomposition-coordination algorithm was used to efficiently coordinate the dynamic behavior of the CDMPC controllers. Newton's method was selected to update the price vector during the coordination process. It was shown with a benchmark process system that Newton's method provides a rapid convergence of the unit operations towards the plant-wide optimal performance.

REFERENCES

Camacho, E.F. and Bordons, C. (1999). *Model Predictive Control*. Springer, Berlin.

Cheng, R., Forbes, J.F., and Yip, W.S. (2007). Price-driven coordination method for solving plant-wide mpc problems. *Journal of Process Control*, 17, 429–438.

Cheng, R., Forbes, J.F., and Yip, W.S. (2008). Dantzig-wolfe decomposition and plant-wide mpc coordination. *Computers and Chemical Engineering*, 32, 1507–1522.

Dunbar, W.B. and Murray, R.M. (2004). Receding horizon control of multi-vehicle formations: A distributed implementation. In *43rd IEEE Conference on Decision and Control*, 1995–2002.

Grosdidier, P., Mason, A., Aitolahti, A., Heinonen, P., and Vanhamaki, V. (1993). Fcc unit reactor-regenerator control. *Computers and Chemical Engineering*, 17(2), 165–179.

Jose, R.A. and Ungar, L.H. (1998). Auction-driven coordination for plantwide optimization. In *Proceedings of Foundations of Computer-Aided Process Operation (FOCAPO)*.

Jose, R.A. and Ungar, L.H. (2000). Pricing interprocess streams using slack auctions. *AIChE Journal*, 46(3), 575–587.

Lu, J.Z. (2003). Challenging control problems and emerging technologies in enterprise optimization. *Control Engineering Practice*, 11, 847–858.

Maciejowski, J.M. (2002). *Predictive Control with Constraints*. Prentice Hall, Essex, England.

Sun, Y. and El-Farra, N.H. (2008). Quasi-decentralized model-based networked control of process systems. *Computers and Chemical Engineering*, 32, 2016–2029.

Venkat, A.N., Rawlings, J.B., and Wright, S.J. (2005). Stability and optimality of distributed model predictive control. In *Proceedings of the 44th IEEE Conference on Decision and Control*, 6680–6685.

Wolbert, D., Joulia, X., Koehret, B., and Biegler, L.T. (1994). Flowsheet optimization and optimal sensitivity analysis using analytical derivatives. *Computers and Chemical Engineering*, 18(11/12), 1083–1095.

Integrating Control and Scheduling of Distributed Energy Resources Over Networks ⋆

Yulei Sun, Sathyendra Ghantasala and Nael H. El-Farra *

* *Department of Chemical Engineering & Materials Science, University of California, Davis, CA 95616 USA (e-mail: nhelfarra@ucdavis.edu)*

Abstract: This paper presents an integrated approach for the control and scheduling of Distributed Energy Resources (DERs) that are managed by a central supervisor over a resource-constrained communication network. The objective is to enhance the performance and disturbance-handling capabilities of the DERs while keeping the communication requirements with the supervisor to a minimum in order to reduce the susceptibility of the DERs to communication outages. To this end, the rate of data transfer from the DERs to the supervisor is initially minimized by embedding in the supervisor a set of models that are used to generate the necessary control action when measurements are not transmitted over the network, and then updating the models' states at discrete time instances. Only a subset of the DERs are allowed to transmit their data at any given time to provide updates to their target models according to a certain scheduling strategy. By formulating the networked closed-loop system as a hybrid system, an explicit characterization of the interdependence between the performance of the DERs, the communication rate, the transmission schedule and times, and the plant-models' mismatch is obtained. It is shown that by judicious selection of the transmission schedule and models, it is possible to optimize the performance of the DERs while simultaneously reducing network utilization beyond what is possible with concurrent transmission configurations. The results are demonstrated through an application to a collection of solid oxide fuel cells in a distributed power network.

Keywords: Networked control, model-based control, scheduling algorithms, distributed energy resources, solid oxide fuel cells.

1. INTRODUCTION

Distributed Energy Resources (DERs) are a suite of on-site, grid-connected or stand-alone technology systems that can be integrated into residential, commercial, or institutional buildings and/or industrial facilities. These energy systems include distributed generation, renewable energy sources, and hybrid generation technologies; energy storage; thermally activated technologies that use recoverable heat for cooling, heating, or power. Such distributed resources offer advantages over conventional grid electricity by offering end users a diversified fuel supply; higher power reliability, quality, and efficiency; lower emissions and greater flexibility to respond to changing energy needs. As the number and diversity of DERs on the grid increases, dispatching these resources at the right time and accounting for the flow of energy correctly become complex problems that require reliable monitoring and telemetering equipment, as well as reliable communication and control technologies to enable the integration and inter-operability functions of a broad range of DERs. Some estimates (Lovins et al (2002)) place the market potential for advanced control and communications technologies in DERs at $3.75-$7.5 billion domestically, and at $15-$30 billion worldwide .

While managing DERs over a communication network offers an appealing modern solution to the control of distributed energy generation, it poses a number of challenges that must be addressed before the full economic and environmental potential of DERs can be realized. These challenges stem in part from the inherent limitations on the information transmission and processing capabilities of communication networks, such as bandwidth limitations, network-induced delays, data losses, signal quantization and real-time scheduling constraints, which can interrupt the connection between the central control authority (the supervisor), the generation units and the loads, and consequently degrade the overall control quality if not properly accounted for in the control system design (see, for example, Zhang et al. (2001); Walsh et al. (2002); Hokayem and Abdallah (2004); Xu and Hespanha (2004); Munoz de la Pena and Christofides (2008) and the references therein for discussions and results on control over communication networks). Despite the availability of fast and reliable communication networks, the fact that the distributed power market is primarily driven by the need for super-reliable, high-quality power implies that the impact of even a brief communication disruption (e.g., due to local network congestion or server outage) can be substantial.

⋆ Financial support, in part by NSF, CBET-0747954, and by the UC Energy Institute is gratefully acknowledged.

In sites such as hospitals, police stations, data centers and high-tech plants which cannot afford blackouts, millisecond outages that merely cause lights to flicker will cause costly computer crashes. Such high-stake risks provide a strong incentive for the development of robust control and communication strategies that ensure the desired levels and quality of power supply from the DERs while minimizing their reliance on the communication medium, which in turn minimizes the impact of communication disruptions on power supply.

Over the past decade, several efforts have been made towards the development and implementation of control strategies for DERs (e.g., Wang (2001); Barsali et al. (2002); Ro and Rahman (2003); Marei et al. (2004); Lasseter (2007)). While the focus of these studies has been mainly on demonstrating the feasibility of the developed control algorithms, the explicit characterization and management of communication constraints in the formulation and solution of the DER control problem have not yet been addressed. An effort to address this problem was initiated in Sun et al. (2009) where a model-based networked control approach was developed for a DER that communicates with the central controller over a bandwidth-constrained communication network that is shared by several other DERs. The minimum allowable communication frequency was characterized for the case when all DER sensor suites communicate their measurements over the network concurrently and are given simultaneous access to the network. In addition to controlling the transmission frequencies of individual DERs in the network, another important way of reducing network utilization is to select and dispatch only a subset of the deployed DERs at any given time to communicate with the supervisor. Under this restriction, the stability and performance properties of each DER become dependent not only on the controller design but also on the selection of the scheduling strategy that determines the order and times in which the sensor suites of the DERs transmit their data to the supervisor.

Motivated by these considerations, we focus in this work on the problem of integrating control and scheduling of DERs over resource-constrained communication networks. The objective is to find an optimal strategy for establishing and terminating communication between the DERs and the central controller that minimizes the rate at which each DER must collect and disseminate data to the supervisor without jeopardizing the stability and performance properties of the DERs. The rest of the paper is organized as follows. Following some preliminaries in Section 2, the problem of DER scheduling over the network is formulated and an overview of its solution is presented. Section 3 then presents the networked control structure and describes its implementation under scheduling. The closed-loop system is then formulated and analyzed in Section 4 where a precise characterization of the interdependence between the networked closed-loop performance, the communication rate between the DERs and the supervisor, the scheduling strategy, as well as the accuracy of the models and the choice of the control laws, is provided. This characterization is shown to allow a systematic search for the sensor transmission schedules that enhance the overall performance while simultaneously reducing the unnecessary utilization of the communication medium. The implementation of the networked control and scheduling strategy are demonstrated in Section 5 through an application to a network of solid oxide fuel cell (SOFC) plants managed by a supervisor over a communication network.

2. PRELIMINARIES

2.1 Structure of distributed generation units

We consider an array of n DERs managed by a higher-level supervisor over a shared bandwidth-limited communication network. Each DER is modeled by a continuous–time system with the following state–space description:

$$\begin{aligned} \dot{x}_i(t) &= A_i x_i + B_{i_1} w_i + B_{i_2} u_i \\ z_i(t) &= C_i x_i + D_i u_i, \quad i = 1, \cdots, n \end{aligned} \tag{1}$$

where $x_i \in \mathbb{R}^{n_i}$ denotes the vector of state variables associated with the i–th DER (e.g., exhaust temperatures and rotation speed in turbines and internal combustion engines, operating temperature and pressures in fuel cells), $u_i \in \mathbb{R}^{m_i}$ denotes the vector of manipulated inputs associated with the i–th DER (e.g., inlet fuel flow rate in fuel cells, shaft speed in turbines), $w_i \in \mathbb{R}^{q_i}$ denotes the vector of disturbance inputs, $z_i \in \mathbb{R}^{p_i}$ is the vector of DER performance output signals of interest (e.g., power, voltage and frequency), and A_i, B_{i_1}, B_{i_2}, C_i, and D_i are constant matrices. Each DER has local (on-board) sensors and actuators with some limited built-in intelligence that gives the DER the ability to run autonomously for periods of time when no communication exists with the remote software controller (the supervisor). The local sensors in each DER transmit their data over a shared communication network to the supervisor where the necessary control calculations are carried out and the control commands are sent back to each DER over the communication network. Based on load changes, changes in utility grid power prices and the state and capacity of each DER, the supervisor regulates and coordinates local power generation in the DERs.

2.2 Problem formulation and methodological framework

One of the main problems to be addressed when managing a large number of DERs over a communication network is the large amount of bandwidth required by the different subsystems sharing the communication medium. A trade-off typically exists between the control performance and the extent of network utilization. On the one hand, optimal control of each DER to deliver the required power quality in the presence of process variations and disturbances is best achieved when information (e.g., measurements, control commands) are exchanged continuously between each DER and the supervisor. Minimal network utilization necessary to save on communication costs, on the other hand, favors only limited communication. Proper characterization and management of this tradeoff is an essential first step to the design of resource-aware networked control systems that ensure the desired performance while respecting inherent constraints on the resources of the communication medium. To address this problem, we will focus in this work on minimizing the sensor-controller communication costs under the assumption that the actuators and supervisor are collocated (i.e., the network exists between the sensors and the controller; generalizations to account for actuator-controller communication constraints are possible and the subject of other research work). To this end, we will consider the following approach:

- Initially design for each DER an appropriate feedback control law that regulates its output (in the absence of communication constraints) at the desired set-point decided by the supervisor.
- Reduce the collection and transfer of information from each DER to the supervisor as much as possible to limit the bandwidth required from the network without sacrificing the desired stability and performance properties by using models of the DERs in the supervisor to calculate the control action when measurements are not available.
- Limit the number of DERs that, at any time, transmit their data to update the corresponding target models.
- Find a scheduling strategy for establishing and terminating communications between the DERs and the supervisor that optimizes a certain performance metric for the closed-loop system while simultaneously keeping the communication rate to a minimum.

3. NETWORKED CONTROLLER DESIGN AND SCHEDULING

3.1 Model-based networked control of DERs

In order to reduce network usage, we embed a dynamic model of each DER in the supervisor to provide it with an estimate of the evolution of the states of the DER when measurements are not available. The use of a model at the controller/actuator side to recreate the dynamics of each DER allows the on-board sensors to transmit their data at discrete time instances and not continuously (since the model can provide an approximation of the DER dynamics) thus allowing conservation of network resources. The computational load associated with this step (e.g., model forecasting and control calculations) is justified and supported by the increasing capabilities of modern computing systems used by the central control authority. Feedback from the DER is then performed by updating the state of the model state using the actual state that is provided by its sensors at discrete time instances. The model-based controller is implemented as follows:

$$\begin{aligned} u_i(t) &= K_i\widehat{x}_i(t),\ t \neq t_k^i \\ \dot{\widehat{x}}_i(t) &= \widehat{A}_i\widehat{x}_i(t) + \widehat{B}_{i_2}u_i(t),\ t \in [t_k^i, t_{k+1}^i) \\ \widehat{x}_i(t_k^i) &= x_i(t_k^i),\ \ k = 0, 1, 2, \cdots \end{aligned} \tag{2}$$

where $\widehat{x}_i$ is an estimate of x_i, $\widehat{A}_i$ and $\widehat{B}_{i_2}$ are estimates of A_i and B_{i_2}, respectively, which do not necessarily match the actual dynamics of the i–th DER, (i.e., in general $\widehat{A}_i \neq A_i$, $\widehat{B}_{i_2} \neq B_{i_2}$). The notation t_k^i is used to indicate the k-th transmission time for the sensor suite of the i-th DER in the collection. The model state is used by the controller as long as no measurements are transmitted, but is updated (or re-set) using the true measurement whenever it becomes available from the network.

3.2 Scheduling DER transmissions over the network

A key parameter in the analysis of the control and update laws in Eq.2 is the update period for each DER, $h^i := t_{k+1}^i - t_k^i$, which determines the frequency at which the sensor suite of the i-th DER collects and sends measurements to the supervisor through the network to update the corresponding model state. To simplify the analysis, we consider in what follows the case when the update period is constant and the same for all DERs, so that $t_{k+1}^i - t_k^i := h,\ i = 1, 2, \cdots, n$. The update period is also an important measure of the extent of network utilization, with a larger h indicating a larger reduction in network utilization. Because of the bandwidth limitations on the communication network and in order to further reduce network utilization, we perform sensor scheduling whereby only one DER is allowed to transmit its measurements to the supervisor at any one time, while the other DERs remain dormant for some time before the next DER is allowed to transmit its data (the results can also be generalized to configurations where multiple DERs transmit at the same time). The transmission schedule is defined by: (a) the sequence (or order) of transmitting suites of DERs: $\{s_i,\ i = 1, 2, \cdots, n\},\ s_i \in \mathcal{N} := \{1, 2, \cdots, n\}$,where s_i is a discrete variable that denotes the i-th transmitting entity in the sequence, and (b) the time at which each DER in the sequence collects and transmits measurements. To characterize the transmission times, we introduce the variable: $\Delta t_i := t_k^{s_{i+1}} - t_k^{s_i},\ i = 1, 2, \cdots, n-1$, which is the time interval between the transmissions of two consecutive DERs in the sequence.

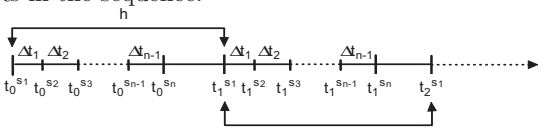


Fig. 1. A schematic of the time-line for the transmissions of DERs in an h-periodic schedule.

Fig.1 is a schematic representation of how DER scheduling is performed. Note that the schedule is h-periodic in the sense that the same sequence of transmitting DERs is executed repeatedly every h seconds (equivalently, each DER transmits its data every h seconds). Note also from the definitions of both h and Δt_i that the transmission times always satisfy the constraint $\sum_{i=1}^{n-1} \Delta t_i < h$. Since the update periods for all DERs are the same, the intervals between the transmission times of two specific DERs are constant, and within any single execution of the schedule (which lasts less than h seconds), each DER can only transmit its measurements through the network and update its model in the supervisor once. This can be represented mathematically by the condition: $s_i \neq s_j$ when $i \neq j$. By manipulating the time intervals Δt_i (i.e., the transmission times) and the order in which the DERs transmit, one can systematically search for the optimal transmission schedule that leads to the largest update period (or the smallest communication rate between the sensor suite of each DER and the supervisor).

4. PERFORMANCE CHARACTERIZATION OF THE SCHEDULED CLOSED-LOOP SYSTEM

4.1 A hybrid system formulation

Defining the model estimation errors by $e_i = x_i - \widehat{x}_i$, where e_i represents the difference between the state of the i-th DER and the state of its model embedded in the supervisor, and introducing the augmented vectors: $\mathbf{e} := [e_1^T\ e_2^T\ \cdots\ e_n^T]^T,\ \mathbf{x} := [x_1^T\ x_2^T\ \cdots\ x_n^T]^T$, it can be shown that the overall networked closed-loop system of Eqs.1-2 can be formulated as a combined discrete–continuous (hybrid) system of the form:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \Lambda_{11}\mathbf{x}(t) + \Lambda_{12}\mathbf{e}(t) + \bar{B}_N w(t) \\ \dot{\mathbf{e}}(t) &= \Lambda_{21}\mathbf{x}(t) + \Lambda_{22}\mathbf{e}(t) + \bar{B}_N w(t),\ \ t \neq t_k^i \\ e_i(t_k^i) &= 0,\ \ i = 1, 2, \cdots, n,\ \ k = 0, 1, 2, \cdots, \end{aligned} \tag{3}$$

where $\bar{B}_N = [B_{1_1}^T\ B_{2_1}^T\ \cdots, B_{n_1}^T]^T$, and the DER states evolve continuously in time while the estimation errors are

reset to zero at each transmission instance. Note, however, that unlike the case of simultaneous DER transmissions (where no scheduling takes place) which was investigated in Sun et al. (2009), not all models within the supervisor are updated (and hence not all estimation errors are reset to zero) at each transmission time. Instead, only the model of the transmitting DER is updated using the measurements provided by its sensor suite.

Referring to Eq.3, Λ_{11}, Λ_{12}, Λ_{21}, and Λ_{22} are all $m \times m$ constant, block-diagonal matrices, where $m = \sum_{i=1}^{n} n_i$ and n_i is the dimension of the i-th state vector. These matrices are linear combinations of A_i, B_{i_2}, $\widehat{A}_i$, $\widehat{B}_{i_2}$, K_i, which are the matrices used to describe the dynamics, the models, and the control laws of the different DERs. The explicit forms of these matrices are given by: $\Lambda_{11} = \text{diag}\{A_i + B_{i_2}K_i\}$, $\Lambda_{12} = \text{diag}\{-B_{i_2}K_i\}$, $\Lambda_{21} = \text{diag}\{\widetilde{A}_i + \widetilde{B}_{i_2}K_i\}$, $\Lambda_{22} = \text{diag}\{\widehat{A}_i + \widetilde{B}_{i_2}K_i\}$, where $\widetilde{A}_i = A_i - \widehat{A}_i$, and $\widetilde{B}_{i_2} = B_{i_2} - \widehat{B}_{i_2}$. Defining the augmented state vector $\xi(t) := [\mathbf{x}^T(t)\ \mathbf{e}^T(t)]^T$, the dynamics of the overall closed-loop system can be written as:

$$\begin{aligned} \dot{\xi}(t) &= \Lambda\xi(t) + B_N w(t),\ t \neq t_k^i \\ e_i(t_k^i) &= 0,\ \ i = 1, 2, \cdots, n,\ \ k = 0, 1, 2, \cdots \\ \mathbf{z}(t) &= C_N\xi(t) \end{aligned} \tag{4}$$

where $\Lambda = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix}$, $B_N = [\bar{B}_N^T \ \ \bar{B}_N^T]^T$, $C_N = [\text{diag}\{C_i + D_iK_i\} \ \ \text{diag}\{-D_iK_i\}]$, and $\mathbf{z} := [z_1^T\ z_2^T\ \cdots\ z_n^T]^T$ is the overall performance output of the DER collection.

4.2 Performance characterization using extended H_2-norm

Our objective in this section is to assess the performance of the networked scheduled closed-loop system subject to disturbances and explicitly characterize its dependence on the update period and the DER transmission schedule to determine an optimal schedule and update period that ensure minimal influence of the disturbances on the performance output of the closed-loop system. As a performance metric, we choose the extended H_2-norm introduced originally in Montestruque and Antsaklis (2006). This performance measure, which is an H_2-like norm that is suitable for analyzing periodic networked control systems, captures the 2-norm of the performance output response when the closed-loop system is initialized at the steady-state and an impulse disturbance is introduced in the input at $t = t_0$ (see Montestruque and Antsaklis (2006) for other types of performance measures that can be used). The following theorem explicitly characterizes the performance output response in terms of the control, communication and scheduling design parameters. The proof can be obtained by solving the system of Eq.4 within each subinterval of the time-line in Fig.1, and is omitted for brevity.

Theorem 1. Consider the system of Eq.4 with a transmission schedule $\{s_1, s_2, \cdots, s_n\}$ and the initial condition $\xi(t_0^{s_1}) = [\mathbf{0}\ \ \mathbf{e}^T(t_0^{s_1})]^T = \xi_0$, with $e_{s_1}(t_0^{s_1}) = 0$, subject to an impulse disturbance $w = \delta(t - t_0^{s_1})$. Then:

(a) For $t \in [t_k^{s_i},\ t_k^{s_{i+1}})$, $i = 1, 2, \cdots, n-1$, $k = 0, 1, 2, \cdots$, the performance output response is given by:

$$\mathbf{z}(t) = C_N e^{\Lambda(t - t_k^{s_i})}\Gamma_i(\Delta t_i, I_s^{s_i})M^k B_N \tag{5}$$

(b) For $t \in [t_k^{s_n},\ t_{k+1}^{s_1})$, $k = 0, 1, 2, \cdots$, the performance output response is given by:

$$\mathbf{z}(t) = C_N e^{\Lambda(t - t_k^{s_n})}\Gamma_n M^k B_N \tag{6}$$

where

$$M(h) = I_s^{s_1} e^{\Lambda(h - \sum_{i=1}^{n-1}\Delta t_i)}\Gamma_n \tag{7}$$

$$\Gamma_i = \left\{ \begin{array}{ll} \displaystyle\prod_{i-1-\mu=0}^{i-2} I_s^{s_{\mu+1}} e^{\Lambda\Delta t_\mu}, & \text{for } i \geq 2 \\ I, & \text{for } i = 1 \end{array} \right\} \tag{8}$$

$$I_s^{s_i} = \begin{bmatrix} I & O & \cdots & O \\ O & H_1 & \cdots & O \\ \vdots & \vdots & & \vdots \\ O & O & \cdots & H_n \end{bmatrix},\ H_i = \left\{ \begin{array}{ll} I, & i \neq s_i \\ O, & i = s_i \end{array} \right. \tag{9}$$

for $i = 1, 2, \cdots, n$, $t_{k+1}^{s_i} - t_k^{s_i} = h$ and $\Delta t_i = t_k^{s_{i+1}} - t_k^{s_i}$, for $i = 1, 2, \cdots, n-1$.

Remark 1: The expression in Eq.5 captures the response of the performance output during the time periods between the transmissions of two consecutive DERs in a given execution of the schedule, while the expression in Eq.6 provides the response for the time period between the transmission of the last DER in a given execution and the transmission of the first DER in the next execution. As expected the responses are parameterized by the transmission sequence (which determines the structure of the matrices $I_s^{s_i}$) as well as the transmission times (which are determined by Δt_i). Note from the term M^k (which captures the growth of the response due to the repeated execution of the transmission schedule) that a necessary and sufficient condition for the responses to be stable is to have all the eigenvalues of the matrix M strictly inside the unit circle (e.g., see Sun and El-Farra (2008) for further details on the characterization of closed-loop stability).

Based on the result of Theorem 1, the extended H_2-norm for the scheduled networked closed-loop system, $\| G \|_{H_2}$, can be calculated using the following defining relation:

$$\| G \|_{H_2} = \text{trace}(B_N^T X B_N)^{1/2} \tag{10}$$

where X is the solution to the discrete Lyapunov equation:

$$M^T(h, I_s^{s_i}, \Delta t_i)XM(h, I_s^{s_i}, \Delta t_i) - X + \sum_{i=1}^{n} W_i = 0, \tag{11}$$

W_i is a matrix computed as:

$$W_i = \int_0^{\Delta t_i} \Gamma_i^T e^{\Lambda^T t} C_N^T C_N e^{\Lambda t}\Gamma_i dt,\ \ i = 1, 2, \cdots, n \tag{12}$$

and $\Delta t_n := h - \sum_{i=1}^{n-1}\Delta t_i$.

Remark 2: The relations of Eqs.10-12 provide a generalization of the extended H_2-norm calculation to networked control systems with scheduled sensor transmissions. In the limit as $\Delta t_i \to 0$, for $i = 1, \cdots, n-1$, (i.e., simultaneous transmissions), these relations reduce to the ones developed originally in Montestruque and Antsaklis (2006) for non-scheduled networked control systems.

Remark 3: By examining Eqs.10-12, it can be seen that $\| G \|_{H_2}$ depends on the interplay between the plant-model mismatch for each DER, the controller gains, the update period, the time intervals between transmissions, as well as the transmission sequence, which altogether provide handles that can be tuned to optimize the performance of the networked closed-loop system subject to disturbances. For example, the extended H_2-norm can be used to

compare different schedules (by varying the transmission sequence and times) to determine which schedules achieve the best performance with the least communication rate between the DERs and the supervisor. Alternatively, if the schedule is fixed by the network access constraints, the performance index can be used to compare the performance levels achieved by using different models and different controllers. The performance criterion can therefore be used to formulate various kinds of optimization problems.

5. SIMULATION STUDY: A NETWORK OF SOLID OXIDE FUEL CELLS

As an illustrative example, we consider a network of three solid oxide fuel cell (SOFC) plants that communicate with the supervisor over a shared communication network . The plants have different dynamic characteristics due to the differences in sizes and capacities of the individual fuel cell stacks. The supervisor is responsible for maintaining the power output of each SOFC plant at a desired set-point by manipulating the inlet fuel flow rate in the presence of disturbances in the inlet air flow rate. Measurements from the sensor suite of each SOFC plant can be received by the supervisor only through the communication network, while the actuator suite of each plant is assumed to have un-interrupted access to the supervisor (ideal actuator-controller links). Under standard modeling assumptions, a dynamic model of the following form can be derived for each SOFC stack from material and energy balances (Mursheda et al. (2007)):

$$\begin{aligned}
\dot{p}_{H_2} &= \frac{T_s}{\tau^*_{H_2} T^* K_{H_2}}(q^{in}_{H_2} - K_{H_2} p_{H_2} - 2K_r I) \\
\dot{p}_{O_2} &= \frac{T_s}{\tau^*_{O_2} T^* K_{O_2}}(q^{in}_{O_2} - K_{O_2} p_{O_2} - K_r I) \\
\dot{p}_{H_2O} &= \frac{T_s}{\tau^*_{H_2O} T^* K_{H_2O}}(q^{in}_{H_2O} - K_{H_2O} p_{H_2O} + 2K_r I) \\
\dot{T}_s &= \frac{1}{m_s C_{ps}} \sum q^{in}_i \int_{T_{ref}}^{T_{in}} C_{p,i}(T) dT \\
&\quad - \sum q^{out}_i \int_{T_{ref}}^{T_{in}} C_{p,i}(T) dT - \dot{n}^r_{H_2} \triangle \hat{H}^o_r - V_s I
\end{aligned} \tag{13}$$

where, p_i is the partial pressure of component i (i: H_2, O_2, H_2O), T_s is the stack temperature, q^{in}_i is the inlet molar flow rate of component i, m_s and C_{ps} are the mass and average specific heat of fuel cell materials excluding gases, $C_{p,i}$ is the specific heat of gas component i, $\triangle \hat{H}^o_r$ is the specific heat of reaction, I is the load current, $\tau^*_i := V/K_i R T^*$ is a time constant for i-th component, K_i is the valve molar constant for component i, and $K_r = N_0/4F$, N_0 is the number of cells in the stack, F is Faraday's constant, V_s is the overall stack voltage:

$$V_s = N_0 \left[\triangle E_0 + \frac{RT_s}{2F} \ln \frac{p_{H_2} p^{(0.5)}_{O_2}}{p_{H_2O}} \right] - r_0 \exp\left[\alpha\left(\frac{1}{T_s} - \frac{1}{T_0}\right)\right] I \tag{14}$$

where r_0 is the internal resistance at T_0, α is the resistance slope (only ohmic losses are included, while activation and concentration losses are neglected), and $\triangle E_0$ is the standard cell potential. Linearizing the SOFC plants around the desired set-points yields a system of the form of Eq.1 with $n = 3$, where x_i, u_i, w_i and z_i are the dimensionless state, manipulated input (inlet fuel flow rate), disturbance (inlet air flow rate) and power output for the i-th plant, respectively. To regulate the power output of each fuel cell, a feedback controller of the form $u_i = K_i x_i$, is designed and implemented. The explicit forms of the plants and controller matrices are omitted due to space limitations.

5.1 Performance under scheduled sensor transmissions

In this section, we investigate the impact of varying the DER transmission schedule, the intervals between transmissions, and the plant-model mismatch on the total power output of the SOFC network which is chosen as the performance output. As mentioned in Section 3, we focus on scheduling configurations where at each transmission time, only the sensor suite of one SOFC plant is allowed to transmit its measurement updates to the supervisor. To quantify the mismatch between each plant and its model that is embedded in the supervisor, we consider as an example parametric uncertainty in Cp_{H_2} and define $\delta_1 = (Cp^m_{H_2} - Cp_{H_2})/Cp_{H_2}$, where $Cp^m_{H_2}$ is a nominal value used in the model, as a measure of model accuracy (any other set of uncertain parameters can also be considered and analyzed in a similar fashion). We initialize the closed-loop SOFC plants at the desired set-points and introduce a unit impulse disturbance in the inlet flow rate of air to each plant. The power outputs of the individual fuel cells are chosen as the performance outputs. Fig.2(a) shows the dependence of the extended H_2-norm of the entire SOFC network on the update period, h, under the six possible sensor transmission schedules listed in Table 1 when imperfect models are embedded in the supervisor (each model with parametric uncertainty $\delta_1 = 5$) and the transmission times are fixed such that $\Delta t_1 = \Delta t_2 = h - \Delta t_1 - \Delta t_2$. It can be seen that among all possible schedules, schedule 4 provides the best performance since for any update period it yields the smallest $\| G \|_{H_2}$. Note also that this schedule yields an improved performance over the non-scheduled (i.e., concurrent) transmission configuration shown by the solid profile. Not only is the minimum extended H_2-norm smaller for the scheduled configuration, but the optimal update period is also larger, which implies that the rate at which each plant needs to collect and transmit measurements to the supervisor under the scheduled configuration is smaller, thus leading to bigger savings in the overall utilization of the communication network resources. The reason for the performance improvement can be understood in light of the fact that forcing the different SOFC plants to transmit their data and update their target models in the supervisor at different times (rather than simultaneously) creates opportunities for providing a more targeted correction to the estimation errors of the different models, where the models with the largest plant-model mismatch can receive more timely updates than would be feasible under simultaneous transmissions. This in turn helps reduce the rate at which each SOFC plant in the communication network must collect and transmit data.

Table 1. SOFC plant transmission schedules

Schedule	$s_1, s_2, s_3, s_1, s_2, s_3, \cdots$
1	$1, 2, 3, 1, 2, 3, \cdots$
2	$1, 3, 2, 1, 3, 2, \cdots$
3	$2, 1, 3, 2, 1, 3, \cdots$
4	$2, 3, 1, 2, 3, 1, \cdots$
5	$3, 1, 2, 3, 1, 2, \cdots$
6	$3, 2, 1, 3, 2, 1, \cdots$

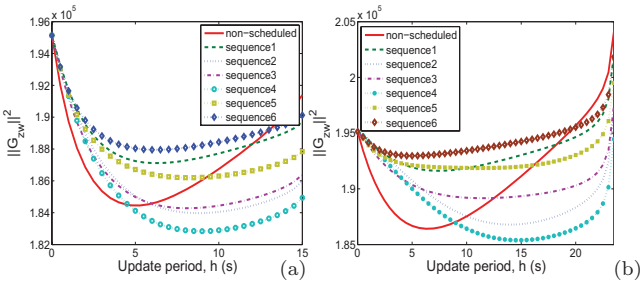


Fig. 2. Dependence of the norm of the power output vector for the SOFC network on the update period for different sensor transmission sequences under (a) a model-based scheme, and (b) a zero-order hold scheme.

Fig.2(b) shows how the extended H_2-norm of the overall SOFC network under each schedule varies as h is varied when a zero-order hold scheme is used. In this case, the supervisor holds the last measurement received from the individual SOFC plant until the next time a measurement is transmitted and received from the network (this corresponds to using models with $\widehat{A}_i = O$ and $\widehat{B}_{i_2} = O$). It can be seen that the optimal update period obtained under scheduling is also larger in this case than the one obtained under simultaneous transmissions.

5.2 Dependence of overall performance on model quality

In this part, we investigate the effect of model uncertainty on the overall SOFC network performance. Fig.3 depicts the dependence of $\| G \|_{H_2}$ of the entire SOFC network on both δ_1 and the update period when the sensors' transmission follows schedule 4. As expected, for a given overall performance level, the range of feasible update period shrinks as the plant-model mismatch increases. The predictions of Fig.3(a) are further confirmed by the closed–loop power output profile in Fig.3(b) which shows that under the same update period of $h = 9$ s and sensor transmission sequence 4, the networked closed-loop system performs better with a relatively accurate model ($\delta_1 = -2$) than the one with an inaccurate model ($\delta_1 = -10$).

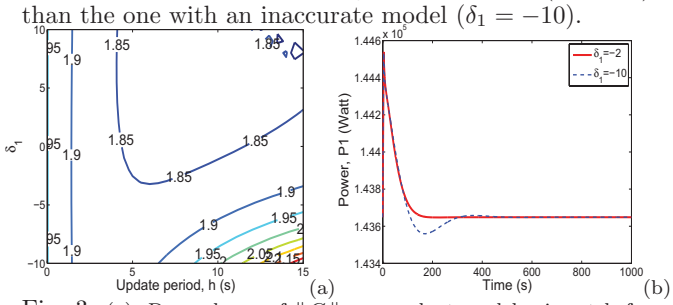


Fig. 3. (a): Dependence of $\| G \|_{H_2}$ on plant-model mismatch for various h. (b): First SOFC plant power output profile under the networked control system with different models.

5.3 Performance dependence on the transmission times

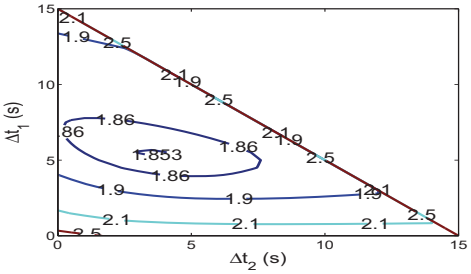


Fig. 4. Dependence of $\| G \|_{H_2}$ on Δt_1 and Δt_2 under schedule 4.

Fig.4 is a contour plot showing the dependence of $\| G \|_{H_2}$ on Δt_1 and Δt_2 for a fixed update period ($h = 15$ s) when the SOFC plants transmit according to schedule 4 and a zero-order hold model is considered. In comparison with the performance achieved in the case when $\Delta t_1 = \Delta t_2 = h - \Delta t_1 - \Delta t_2$ ($\| G \|_{H_2} = 1.854 \times 10^5$; see Fig.2(b)), it can be seen that an improved performance is attained ($\| G \|_{H_2} = 1.853 \times 10^5$) by varying the transmission times such that $\Delta t_1 = 5.5$ s and $\Delta t_2 = 3.5$ s.

REFERENCES

Barsali, S., Ceraolo, M., Pelacchi, P., and Poli, D. (2002). Control techniques of dispersed generators to improve the continuity of electricity supply. In *Proceedings of IEEE Power Eng. Soc. Win. Mtg.*, 789–794.

Hokayem, P. and Abdallah, C. (2004). Inherent issues in networked control systems: A survey. In *Proceedings of American Control Conference*, 329–336. Boston, MA.

Lasseter, R.H. (2007). Microgrids and distributed generation. *Journal of Energy Engineering*, 133, 144–149.

Lovins et al, A.B. (2002). *Small Is Profitable: The Hidden Economic Benefits of Making Electrical Sources the Right Size*. Rocky Mountain Institute.

Marei, M.I., El-Saadany, E.F., and Salama, M.M.A. (2004). A novel control algorithm for the DG interface to mitigate power quality problems. *IEEE Transactions on Power Delivery*, 19, 1384–1392.

Montestruque, L.A. and Antsaklis, P.J. (2006). Performance evaluation for model-based networked control systems. In *Netw. Emb. Sens. and Cntrl., LNCIS*, volume 331, 231–249. P.J. Antsaklis, P. Tabuada (Eds.), Springer-Verlag, Berlin.

Munoz de la Pena, D. and Christofides, P.D. (2008). Lyapunov-based model predictive control of nonlinear systems subject to data losses. *IEEE Trans. Automat. Contr.*, 53, 2067–2089.

Mursheda, A., Huang, B., and Nandakumara, K. (2007). Control relevant modeling of planer solid oxide fuel cell system. *Journal of Power Sources*, 163, 830–845.

Ro, K. and Rahman, S. (2003). Control of grid-connected fuel cell plants for enhancement of power system stability. *Renewable Energy*, 28, 397–407.

Sun, Y. and El-Farra, N.H. (2008). Quasi-decentralized model-based networked control of process systems. *Comp. & Chem. Eng.*, 32, 2016–2029.

Sun, Y., Ghantasala, S., and El-Farra, N.H. (2009). Networked control of distributed energy resources: Application to solid oxide fuel cells. In *Proceedings of American Control Conference,* to appear. St. Louis, MO.

Walsh, G., Ye, H., and Bushnell, L. (2002). Stability analysis of networked control systems. *IEEE Trans. Contr. Syst. Tech.*, 10, 438–446.

Wang, H.F. (2001). Multi-agent co-ordination for the secondary voltage control in power-system contingencies. *IEE Proceedings on Generation, Transmission and Distribution*, 148, 61–66.

Xu, Y. and Hespanha, J.P. (2004). Optimal communication logics for networked control systems. In *Proceedings of the 43rd IEEE Conference on Decision and Control*, 3527–3532. Atlantis, Paradise Island, Bahamas.

Zhang, W., Branicky, M.S., and Phillips, S.M. (2001). Stability of networked control systems. *IEEE Contr. Syst. Mag.*, 21, 84–99.

Distributed Model Predictive Control of Nonlinear Process Systems Subject to Asynchronous Measurements

Jinfeng Liu [*] **David Muñoz de la Peña** [**]
Panagiotis D. Christofides [*,***,1]

[*] *Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA, 90095-1592, USA.*
[**] *Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Camino de los Descubrimientos S/N, 41092, Sevilla, Spain.*
[***] *Department of Electrical Engineering, University of California, Los Angeles, CA, 90095-1592, USA.*

Abstract: In this work, we address distributed model predictive control of nonlinear process systems subject to asynchronous measurements. Assuming that there exists an upper bound on the interval between two successive measurements of the process state, two separate Lyapunov-based model predictive controllers that coordinate their actions and take asynchronous measurements explicitly into account are designed. The proposed distributed control design only requires one directional communication between the two distributed controllers and provides the potential of maintaining stability and performance in the face of new or failing actuators. The results are illustrated through a chemical process example.

Keywords: Distributed model predictive control; Nonlinear systems; Networked control systems; Process control; Asynchronous measurements.

1. INTRODUCTION

We are currently witnessing an augmentation of the existing, dedicated local control networks, with additional networked (wired and/or wireless) actuator/sensor devices which have become cheap and easy-to-install the last few years. Such an augmentation in sensor information and networked-based availability of data has the potential (Ydstie (2002); Neumann (2007); Christofides et al. (2007)) to be transformative in the sense of dramatically improving the ability of the control systems to optimize process performance and prevent or deal with abnormal situations more quickly and effectively. However, augmenting dedicated, local control systems (LCS) with control systems that may utilize real-time sensor and actuator networks gives rise to the need to design/redesign and coordinate separate control systems that operate on a process. Model predictive control (MPC) is a natural control framework to deal with the design of coordinated, distributed control systems because of its ability to handle input and state constraints, and also because it can account for the actions of other actuators in computing the control action of a given set of control actuators in real-time. Motivated by the lack of available methods for the design of networked control systems (NCS) for chemical processes, in a previous work (Liu et al. (2008)), we introduced a decentralized control architecture for systems with continuous and asynchronous measurements. In this architecture, the local, pre-existing control system uses continuous sensing and actuation and an explicit control law. On the other hand, the NCS uses networked (wired or wireless) sensors and actuators and has access to heterogeneous, asynchronous measurements that are not available to the LCS. The NCS is designed via Lyapunov-based model predictive control (LMPC). Following up on this work, in another recent work (Liu et al. (in press)), we proposed a distributed model predictive control method for the design of networked control systems where both the pre-existing local control system and the networked control system are designed via Lyapunov-based model predictive control.

With respect to available results on distributed MPC design, several distributed MPC methods have been proposed in the literature that deal with the coordination of separate MPC controllers that communicate in order to obtain optimal input trajectories in a distributed manner (Rawlings and Stewart (2007); Dunbar (2007); Richards and How (2007); Keviczky et al. (2006); Magni and Scattolini (2006); Raimondo et al. (2007)). All of the above results on distributed MPC design are based on the assumption of continuous sampling and perfect communication between the sensor and the controller. However, one may encounter asynchronous measurement samplings because of measurement difficulties in process control applications.

In this work, we address distributed model predictive control of nonlinear process systems subject to asynchronous measurements. Assuming that there exists an upper bound on the interval between two successive measurements of the process state, two separate Lyapunov-based model predictive controllers that coordinate their actions and

[1] Corresponding author: Panagiotis D. Christofides. Tel.:+1 310 825 2046; fax: +1 310 206 4107; e-mail: pdc@seas.ucla.edu.

take asynchronous measurements explicitly into account are designed. Sufficient conditions are derived for the proposed distributed control design to guarantee that the state of the closed-loop system is ultimately bounded in a region that contains the origin. In addition, the proposed distributed control design only requires one directional communication between the two distributed controllers and provides the potential of maintaining stability and performance in the face of new or failing actuators. The results are illustrated through a chemical process example.

2. PRELIMINARIES

2.1 Control problem formulation

We consider nonlinear process systems described by the following state-space model

$$\dot{x}(t) = f(x(t), u_1(t), u_2(t), w(t)) \tag{1}$$

where $x(t) \in R^{n_x}$ denotes the vector of process state variables, $u_1(t) \in R^{n_{u_1}}$ and $u_2(t) \in R^{n_{u_2}}$ are two separate sets of manipulated inputs and $w(t) \in R^{n_w}$ denotes the vector of disturbance variables. The two manipulated inputs are restricted to be in two nonempty convex sets $U_1 \subseteq R^{n_{u_1}}$ and $U_2 \subseteq R^{n_{u_2}}$ and the disturbance vector is bounded, i.e., $w(t) \in W$ where

$$W := \{w \in R^{n_w} \text{ s.t. } |w| \leq \theta, \theta > 0\}\,[2]\,.$$

We assume that f is a locally Lipschitz vector function and $f(0,0,0,0) = 0$. This means that the origin is an equilibrium point for the nominal system (system (1) with $w(t) = 0$ for all t) with $u_1 = 0$ and $u_2 = 0$. System (1) is controlled with the two sets of manipulated inputs u_1 and u_2, which could be multiple inputs of a system or a single input divided artificially into two terms (e.g., $\dot{x}(t) = \hat{f}(x(t), u(t), w(t))$ with $u(t) = u_1(t) + u_2(t)$).

2.2 Lyapunov-based controller

We assume that there exists a Lyapunov-based controller $u_1(t) = h(x(t))$ which satisfies the input constraint on u_1 for all x inside a given stability region and renders the origin of the nominal closed-loop system asymptotically stable with $u_2(t) = 0$. Using converse Lyapunov theorems (Massera (1956); Lin et al. (1996)), this assumption implies that there exist functions $\alpha_i(\cdot)$, $i = 1, 2, 3, 4$ of class $\mathcal{K}$ [3] and a continuously differentiable Lyapunov function V for the nominal closed-loop system that satisfy the following inequalities

$$\begin{gathered} \alpha_1(|x|) \leq V(x) \leq \alpha_2(|x|) \\ \frac{\partial V(x)}{\partial x} f(x, h(x), 0, 0) \leq -\alpha_3(|x|) \\ |\frac{\partial V(x)}{\partial x}| \leq \alpha_4(|x|) \\ h(x) \in U_1 \end{gathered} \tag{2}$$

for all $x \in D \subseteq R^{n_x}$ where D is an open neighborhood of the origin. We denote the region Ω_ρ [4] $\subseteq D$ as the stability region of the closed-loop system under the control $u_1 = h(x)$ and $u_2 = 0$.

[2] $|\cdot|$ denotes Euclidean norm of a vector.
[3] A continuous function $\alpha : [0, a) \to [0, \infty)$ is said to belong to class $\mathcal{K}$ if it is strictly increasing and $\alpha(0) = 0$.
[4] We use Ω_r to denote the set $\Omega_r := \{x \in R^{n_x} | V(x) \leq r\}$.

By continuity and the local Lipschitz property assumed for the vector field $f(x, u_1, u_2, w)$ and the fact that the manipulated inputs u_1 and u_2 are bounded in convex sets, there exists a positive constant M such that

$$|f(x, u_1, u_2, w)| \leq M \tag{3}$$

for all $x \in \Omega_\rho$, $u_1 \in U_1$, $u_2 \in U_2$ and $w \in W$. In addition, by the continuous differentiable property of the Lyapunov function V and the Lipschitz property assumed for the vector field $f(x, u_1, u_2, w)$, there exist positive constants L_x, R_x, R_w such that

$$|\frac{\partial V}{\partial x} f(x, u_1, u_2, 0) - \frac{\partial V}{\partial x} f(x', u_1, u_2, 0)| \leq L_x |x - x'| \tag{4}$$

and

$$|f(x, u_1, u_2, w) - f(x', u_1, u_2, 0)| \leq R_x |x - x'| + R_w |w| \tag{5}$$

for all $x, x' \in \Omega_\rho$, $u_1 \in U_1$, $u_2 \in U_2$ and $w \in W$.

These constants will be used in section 4 in the proof of the main results of the present work.

2.3 Modeling of asynchronous measurements

Most control systems assume that measurements from sensors are obtained in a continuous periodic pattern. However, in many chemical processes, this assumption does not hold due to a host of measurement difficulties. In this case, the system is subject to asynchronous measurements. In the present work, we assume the state of system (1), $x(t)$, is sampled and available asynchronously at time instants t_k where $\{t_{k\geq 0}\}$ is a random increasing sequence of times. The distribution of $\{t_{k\geq 0}\}$ characterizes the time needed to obtain a new measurement in the case of asynchronous measurements. In general, there exists the possibility of arbitrarily large (but finite) periods of time in which a new measurement is not available. In such a case, it is not possible to provide guaranteed stability properties, because there exists a non-zero probability that the system operates in open loop for a period of time large enough for the state to leave the stability region. In order to study the stability properties in a deterministic framework, in the present work, we assume that there exists an upper bound T_m on the interval between two successive measurements, i.e., $\max_k\{t_{k+1} - t_k\} \leq T_m$. This assumption is reasonable from a process control perspective.

3. DISTRIBUTED LMPC

3.1 Distributed LMPC formulations

In our previous work (Liu et al. (in press)), we introduced a distributed model predictive control method where both the pre-existing LCS and the NCS are designed via Lyapunov-based model predictive control as shown in Fig. 1. The LMPCs computing the input trajectories of the LCS (i.e., u_1) and the NCS (i.e., u_2) are referred to as LMPC 1 and LMPC 2, respectively. Under the assumption of continuous and flawless measurements, in Liu et al. (in press), it was proved that this control scheme guarantees practical stability of the closed-loop system and has the potential to maintain the closed-loop stability and performance in the face of new or failing actuators (for example, the failure of the actuator of the NCS (zero input) does not affect the closed-loop stability) and to reduce

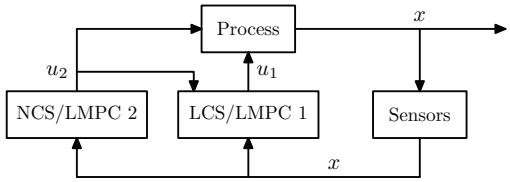


Fig. 1. Distributed LMPC design for networked control systems with continuous measurements (i.e., $x(t)$ is available to the controllers at $t_k = t_{k-1} + \Delta$ where Δ is a fixed sampling time for all k).

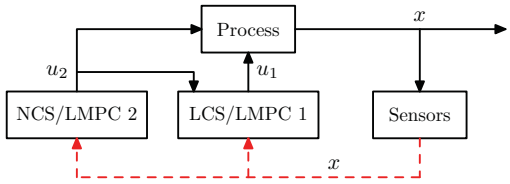


Fig. 2. Distributed LMPC design for networked control systems subject to asynchronous measurements.

computational burden in the evaluation of the optimal manipulated inputs compared with a centralized LMPC. However, when asynchronous measurements are present as shown in Fig. 2, these results do not hold. In this work, the distributed model predictive control method is extended to take into account asynchronous measurements explicitly, both in the constraints imposed on the LMPCs and in the implementation strategy.

In the presence of asynchronous measurements, the controllers need to operate in open-loop between successive new state measurements. We propose to take advantage of the model predictive control scheme to update the input based on a prediction obtained using the model. This is achieved by having the control actuators to store and implement the last computed optimal input trajectory. The proposed implementation strategy in the presence of asynchronous measurements is as follows:

(1) When a measurement $x(t_k)$ is available at t_k, LMPC 2 computes the optimal input trajectory of u_2;
(2) LMPC 2 sends the entire optimal input trajectory to its actuators and also sends the entire optimal input trajectory to LMPC 1.
(3) Once LMPC 1 receives the entire optimal input trajectory for u_2, it evaluates the future input trajectory of u_1;
(4) LMPC 1 sends the entire optimal input trajectory to its actuators.
(5) When a new measurement is received ($k = k+1$), go to step 1.

Note that in the proposed distributed scheme, only LMPC 2 is required to send its optimal input trajectory to LMPC 1 each time when a new measurement is available. This minimizes the communications required between the two controllers. Note also that the communication between LMPC 1 and LMPC 2 is in general done using a reliable link, and hence, it is not subject to data losses or delays.

We first design the optimization problem that charcterizes LMPC 2. This optimization problem depends on the latest state measurement $x(t_k)$, however, LMPC 2 does not have any information about the value that u_1 will take. In order to take a decision, LMPC 2 must assume a trajectory for u_1 along the prediction horizon. To this end, the Lyapunov-based controller $u_1 = h(x)$ is used. LMPC 2 is based on the following optimization problem:

$$\min_{u_{d2}\in S(\Delta)} \int_0^{N\Delta} L(\tilde{x}(\tau), u_{d1}(\tau), u_{d2}(\tau))d\tau \tag{6a}$$

$$\dot{\tilde{x}}(\tau) = f(\tilde{x}(\tau), h(\tilde{x}(j\Delta)), u_{d2}(\tau), 0), \quad \forall \tau \in [j\Delta, (j+1)\Delta) \tag{6b}$$

$$\dot{\hat{x}}(\tau) = f(\hat{x}(\tau), h(\hat{x}(j\Delta)), 0, 0), \forall\, \tau \in [j\Delta, (j+1)\Delta) \tag{6c}$$

$$u_{d2}(\tau) \in U_2, \forall \tau \in [0, N\Delta) \tag{6d}$$

$$\tilde{x}(0) = \hat{x}(0) = x(t_k) \tag{6e}$$

$$V(\tilde{x}(\tau)) \leq V(\hat{x}(\tau)), \forall \tau \in [0, N_R\Delta) \tag{6f}$$

where $S(\Delta)$ is the family of piece-wise constant functions with sampling time Δ, N is the prediction horizon,

$$L(x, u_1, u_2) = x^T Q_c x + u_1^T R_{c1} u_1 + u_2^T R_{c2} u_2$$

is the performance index, Q_c, R_{c1} and R_{c2} are positive definite weight matrices that define the cost, $\tilde{x}$ is the predicted trajectory of the nominal system with u_2 being the input trajectory computed by the LMPC of Eq. 6 (i.e., LMPC 2) and u_1 being the Lyapunov-based controller h applied in a sample-and-hold fashion with $j = 0, ..., N-1$, $\hat{x}$ is the predicted trajectory of the nominal system with u_1 being h applied in a sample-and-hold fashion and $u_2 = 0$, $x(t_k)$ is the state measurement obtained at t_k and N_R is the smallest integer that satisfies the inequality $T_m \leq N_R\Delta$. To take full advantage of the nominal model in the computation of the control action, we take $N \geq N_R$.

The optimal solution to this optimization problem is denoted by $u^*_{d2}(\tau|t_k)$. Once this optimal input trajectory of u_2 is available, it is sent to LMPC 1 as well as the control actuators controlled by LMPC 1.

In order to inherit the stability properties of the Lyapunov based controller, u_2 must satisfy the constraint (6f) which guarantees that the predicted decrease of the Lyapunov function from t_k to $t_k + N_R\Delta$, if $u_1 = h(x)$ and $u_2 = u^*_{d2}$ are applied, is at least equal to the one obtained if the Lyapunov-based controller h is applied in a sample-and-hold fashion. Note that we have considered input constraints, see Eq. 6d.

The optimization problem of LMPC 1 depends on the latest state measurement $x(t_k)$ and the decision taken by LMPC 2 (i.e., u^*_{d2}). This allows LMPC 1 to compute an input u_1 such that the closed-loop performance is optimized, while guaranteeing that the stability properties of the Lyapunov-based controller are preserved. Specifically, LMPC 1 is based on the following optimization problem:

$$\min_{u_{d1}\in S(\Delta)} \int_0^{N\Delta} L(\check{x}(\tau), u_{d1}(\tau), u_{d2}(\tau))d\tau \tag{7a}$$

$$\dot{\check{x}}(\tau) = f(\check{x}(\tau), u_{d1}(\tau), u_{d2}(\tau), 0), \forall \tau \in [0, N\Delta) \tag{7b}$$

$$\dot{\tilde{x}}(\tau) = f(\tilde{x}(\tau), h(\tilde{x}(j\Delta)), u_{d2}(\tau), 0), \quad \forall \tau \in [j\Delta, (j+1)\Delta) \tag{7c}$$

$$u_{d2}(\tau) = u^*_{d2}(\tau|t_k), \forall \tau \in [0, N\Delta) \tag{7d}$$

$$u_{d1}(\tau) \in U_1, \forall \tau \in [0, N\Delta) \tag{7e}$$

$$\check{x}(0) = \tilde{x}(0) = x(t_k) \tag{7f}$$

$$V(\check{x}(\tau)) \leq V(\tilde{x}(\tau)), \ \forall \tau \in [0, N_R\Delta) \tag{7g}$$

where $\check{x}$ is the predicted trajectory of the nominal system if $u_2 = u^*_{d2}$ and $u_1 = u_{d1}$ are applied, and $\tilde{x}$ is the predicted trajectory of the nominal system if $u_2 = u^*_{d2}$ and the Lyapunov-based controller h are applied in a sample-and-hold fashion.

The optimal solution to this optimization problem is denoted by $u^*_{d1}(\tau|t_k)$. The contractive constraint (7g) guarantees that the predicted decrease of the Lyapunov function from t_k to $t_k + N_R\Delta$, if $u_1 = u^*_{d1}$ and $u_2 = u^*_{d2}$ are applied, is at least equal to the one obtained when $u_1 = h(x)$ and $u_2 = u^*_{d2}$ are applied.

Note that the trajectory $\tilde{x}(\tau)$ predicted by constraint (7c) is the same optimal trajectory predicted by LMPC 2. This trajectory and the two contractive constraints (6f) and (7g) allow proving the closed-loop stability properties of the proposed controller.

The manipulated inputs of the proposed control scheme are defined as follows:

$$\begin{array}{l} u_1(t) = u^*_{d1}(t - t_k|t_k),\ \forall t \in [t_k, t_{k+1}) \\ u_2(t) = u^*_{d2}(t - t_k|t_k),\ \forall t \in [t_k, t_{k+1}). \end{array} \quad (8)$$

Note that, as explained before, the controllers apply the last evaluated optimal input trajectory between two successive state measurements.

4. STABILITY PROPERTIES

In this section, we present the stability properties of the proposed distributed control scheme. We prove that the contractive constraints (6f) and (7g) guarantee that the proposed distributed control scheme inherits the stability properties of the Lyapunov-based controller (implemented in sample and hold and using the model to estimate the state of the system when a new measurement is not available). This property is presented in Theorem 1 below. To state this theorem, we need the following propositions.

Proposition 1. (c.f. Muñoz de la Peña and Christofides (2008)). *Consider the nominal sampled trajectory $\hat{x}$ of system (1) in closed-loop with the Lyapunov-based controller h applied in a sample-and-hold fashion and $u_2(t) = 0$. Let $\Delta, \epsilon_s > 0$ and $\rho > \rho_s > 0$ satisfy*

$$-\alpha_3(\alpha_2^{-1}(\rho_s)) + \alpha_4(\alpha_1^{-1}(\rho))L_x M\Delta \leq -\epsilon_s/\Delta. \quad (9)$$

Then, if $\rho_{\min} < \rho$ where

$$\rho_{min} = \max\{V(\hat{x}(t + \Delta)) : V(\hat{x}(t)) \leq \rho_s\} \quad (10)$$

and $\hat{x}(0) \in \Omega_\rho$, the following inequality holds

$$V(\hat{x}(k\Delta)) \leq \max\{V(\hat{x}(0)) - k\epsilon_s, \rho_{min}\}. \quad (11)$$

Proposition 1 ensures that if system (1) with $w(t) = 0$ for all t under the control law $u_1 = h(x)$ implemented in a sample-and-hold fashion and $u_2 = 0$ starts in Ω_ρ, then it is ultimately bounded in $\Omega_{\rho_{\min}}$. The following proposition provides an upper bound on the deviation of the state trajectory obtained using the nominal model, from the real-state trajectory when the same control input trajectories are applied.

Proposition 2. (c.f. Liu et al. (2008)). *Consider the following state trajectories*

$$\begin{array}{l} \dot{x}_a(t) = f(x_a(t), u_1(t), u_2(t), w(t)) \\ \dot{x}_b(t) = f(x_b(t), u_1(t), u_2(t), 0) \end{array} \quad (12)$$

with initial states $x_a(t_0) = x_b(t_0) \in \Omega_\rho$. There exists a class $\mathcal{K}$ function $f_W(\cdot)$ such that

$$|x_a(t) - x_b(t)| \leq f_W(t - t_0), \quad (13)$$

for all $x_a(t), x_b(t) \in \Omega_\rho$ and all $w(t) \in W$ with

$$f_W(\tau) = \frac{R_w\theta}{R_x}(e^{R_x\tau} - 1).$$

The following proposition bounds the difference between the magnitudes of the Lyapunov function of two different states in Ω_ρ.

Proposition 3. (c.f. Liu et al. (2008)). *Consider the Lyapunov function $V(\cdot)$ of system (1). There exists a quadratic function $f_V(\cdot)$ such that*

$$V(x) \leq V(\hat{x}) + f_V(|x - \hat{x}|) \quad (14)$$

for all $x, \hat{x} \in \Omega_\rho$ with

$$f_V(s) = \alpha_4(\alpha_1^{-1}(\rho))s + Ms^2.$$

In Theorem 1 below, we provide sufficient conditions under which the proposed distributed LMPC design (8) guarantees the closed-loop stability of system (1) in the presence of asynchronous measurements.

Theorem 1. Consider system (1) in closed-loop with the distributed LMPC design (8) based on a controller $h(x)$ that satisfies (2). Let $\Delta, \epsilon_s > 0$, $\rho > \rho_{min} > 0$, $\rho > \rho_s > 0$ and $N \geq N_R \geq 1$ satisfy (9),(10) and the following inequality

$$-N_R\epsilon_s + f_V(f_W(N_R\Delta)) < 0. \quad (15)$$

If $x(t_0) \in \Omega_\rho$, then $x(t)$ is ultimately bounded in $\Omega_{\rho_c} \subseteq \Omega_\rho$ where

$$\rho_c = \rho_{min} + f_V(f_W(N_R\Delta)).$$

Proof: In order to prove that the closed-loop system is ultimately bounded in a region that contains the origin, we will prove that $V(x(t_k))$ is a decreasing sequence of values with a lower bound.

The proof consists of two parts. In the first part, we will prove that the stability results stated in Theorem 1 hold for the case where $t_{k+1} - t_k = T_m$ for all k and $T_m = N_R\Delta$. The proof of the stability results for the general case, that is $t_{k+1} - t_k \leq T_m$ for all k and $T_m \leq N_R\Delta$, will be shown in the second part.

Part 1: In this part, we prove that the stability results stated in Theorem 1 hold in the case that $t_{k+1} - t_k = T_m$ for all k and $T_m = N_R\Delta$. This case corresponds to the worst possible situation in the sense that LMPC 1 and LMPC 2 need to operate in open-loop for the maximum possible amount of time.

In order to simplify the notation, we will denote $\tilde{x}(t)$ the nominal closed-loop trajectory of system (1) with $u_1 = h$ implemented in a sample-and-hold fashion and $u_2 = u^*_{d2}$ from $x(t_k)$, $\hat{x}(t)$ the nominal closed-loop trajectory of system (1) under the Lyapunov-based controller $u_1 = h$ implemented in a sample-and-hold fashion and $u_2 = 0$ from $x(t_k)$, and denote $\check{x}(t)$ the nominal closed-loop trajectory of system (1) with $u_1 = u^*_{1d}$ and $u_2 = u^*_{2d}$ from $x(t_k)$.

By Proposition 1 and the fact that $t_{k+1} = t_k + N_R\Delta$, the following inequality can be obtained:

$$V(\hat{x}(t_{k+1})) \leq \max\{V(\hat{x}(t_k)) - N_R\epsilon_s, \rho_{\min}\}. \quad (16)$$

From the contractive constraints (6f) and (7g) in LMPC 2 and LMPC 1, the following inequality can be written:

$$V(\check{x}(t)) \leq V(\tilde{x}(t)) \leq V(\hat{x}(t)), \forall t \in [t_k, t_k + N_R\Delta). \quad (17)$$

From inequalities (16) (17) and taking into account that $\hat{x}(t_k) = \tilde{x}(t_k) = \check{x}(t_k) = x(t_k)$, the following inequality is obtained:

$$V(\check{x}(t_{k+1})) \leq \max\{V(x(t_k)) - N_R\epsilon_s, \rho_{\min}\}. \quad (18)$$

When $x(t) \in \Omega_\rho$ for all times (this point will be proved below), we can apply Proposition 3 to obtain the following inequalities:

$$\begin{aligned} V(x(t_{k+1})) \leq\ & V(\check{x}(t_{k+1})) \\ & + f_V(|\check{x}(t_{k+1}) - x(t_{k+1})|). \end{aligned} \quad (19)$$

Applying Proposition 2 we obtain the following upper bound on the deviation of $\check{x}(t)$ from $x(t)$:

$$|x(t_{k+1}) - \check{x}(t_{k+1})| \leq f_W(N_R\Delta) \quad (20)$$

From inequalities (19) and (20), the following upper bound on $V(x(t_{k+1}))$ can be written:

$$V(x(t_{k+1})) \leq V(\check{x}(t_{k+1})) + f_V(f_W(N_R\Delta)). \quad (21)$$

Using inequality (18), we can re-write inequality (21) as follows:

$$\begin{aligned} V(x(t_{k+1})) \leq\ & \max\{V(x(t_k)) - N_R\epsilon_s, \rho_{\min}\} \\ & + f_V(f_W(N_R\Delta)). \end{aligned} \quad (22)$$

If condition (15) is satisfied, from inequality (22), we know there exists $\epsilon_w > 0$ such that the following inequality holds:

$$V(x(t_{k+1})) \leq \max\{V(x(t_k)) - \epsilon_w, \rho_c\} \quad (23)$$

which implies that if $x(t_k) \in \Omega_\rho/\Omega_{\rho_c}$, then $V(x(t_{k+1})) < V(x(t_k))$, and if $x(t_k) \in \Omega_{\rho_c}$, then $V(x(t_{k+1})) < \rho_c$. Using inequality (23) recursively, it is proved that if $x(t_0) \in \Omega_\rho$, the closed-loop trajectories of system (1) under the proposed distributed LMPC design (8) satisfy

$$\limsup_{t\to\infty} V(x(t)) \leq \rho_c.$$

This proves that the closed-loop system is ultimately bounded in Ω_{ρ_c} for the case where $t_{k+1} - t_k = T_m$ for all k and $T_m = N_R\Delta$.

Part 2: In this part, we extend the results proved in Part 1 to the general case, that is, $t_{k+1} - t_k \leq T_m$ for all k and $T_m \leq N_R\Delta$ which implies that $t_{k+1} - t_k \leq N_R\Delta$. The proof is divided into two cases. The first case is that $t_{k+1} - t_k \leq \Delta$. In this case, the stability results hold as shown in Liu et al. (in press). The second case is that $\Delta < t_{k+1} - t_k \leq N_R\Delta$. Because f_V and f_W are convex and strictly increasing functions of their arguments (see Propositions 2 and 3 for the expressions of f_V and f_W) and following similar steps in Part 1, we can show that inequality (22) still holds. This proves that the stability results stated in Theorem 1 hold.

5. APPLICATION TO A CHEMICAL PROCESS

The process considered in this example is a three vessel, reactor-separator process consisting of two continuously stirred tank reactors (CSTRs) and a flash tank separator.A feed stream to the first CSTR F_{10} contains the reactant A which is converted into the desired product B. The desired product B can then further react into an undesired side-product C. The effluent of the first CSTR along with additional fresh feed F_{20} makes up the inlet to the second CSTR. The reactions $A \to B$ and $B \to C$ (referred to as 1 and 2, respectively) take place in the two CSTRs in series before the effluent from CSTR 2 is fed to a flash tank. The overhead vapor from the flash tank is condensed and recycled to the first CSTR and the bottom product stream is removed. A small portion of the overhead is purged before being recycled to the first CSTR. All the three vessels are assumed to have static holdup. The dynamic equations describing the behavior of the system, obtained through material and energy balances under standard modeling assumptions, can be found in Liu et al. (in press).

Each of the tanks has an external heat input. The manipulated inputs to the system are the heat inputs, Q_1, Q_2 and Q_3, and the feed stream flow rate to vessel 2, F_{20}.

The process was numerically simulated using a standard Euler integration method. Process noise was added to simulate disturbances/model uncertainty and it was generated as autocorrelated noise of the form $w_k = \phi w_{k-1} + \xi_k$ where $k = 0, 1, \ldots$ is the discrete time step of 0.001 hr, ξ_k is generated by a normally distributed random variable with standard deviation σ_p, and ϕ is the autocorrelation factor and w_k is bounded by θ_p, that is $|w_k| \leq \theta_p$.

We assume that the measurements of the temperatures T_1, T_2, T_3 and the measurements of mass fractions x_{A1}, x_{B1}, x_{A2}, x_{B2}, x_{A3}, x_{B3} are available asynchronously at time instants $\{t_{k\geq 0}\}$ with an upper bound $T_m = 3\Delta$ on the maximum interval between two successive measurements, where Δ is the controller sampling time and chosen to be $\Delta = 0.02\ hr = 1.2\ min$.

For each set of steady-state inputs Q_{1s}, Q_{2s}, Q_{3s} and F_{20s} corresponding to a different operating condition, the process has one steady-state x_s. The control objective is to steer the process to the steady state

$$x_s^T = [0.61, 0.39, 425.9, 0.61, 0.39, 422.6, 0.35, 0.63, 427.3].$$

The process belongs to the following class of nonlinear systems

$$\dot{x}(t) = f(x(t)) + g_1(x(t))u_1(t) + g_2(x(t))u_2(t) + w(t)$$

where $x^T = [x_1\ x_2\ x_3\ x_4\ x_5\ x_6\ x_7\ x_8\ x_9] = [x_{A1} - x_{A1s}\ x_{B1} - x_{B1s}\ T_1 - T_{1s}\ x_{A2} - x_{A2s}\ x_{B2} - x_{B2s}\ T_2 - T_{2s}\ x_{A3} - x_{A3s}\ x_{B3} - x_{B3s}\ T_3 - T_{3s}]$ is the state, $u_1^T = [u_{11}\ u_{12}\ u_{13}] = [Q_1 - Q_{1s}\ Q_2 - Q_{2s}\ Q_3 - Q_{3s}]$ and $u_2 = F_{20} - F_{20s}$ are the manipulated inputs which are subject to the constraints $|u_{1i}| \leq 10^6\ KJ/hr$ $(i = 1, 2, 3)$ and $|u_2| \leq 3\ m^3/hr$, and $w = w_k$ is a time varying noise.

To illustrate the theoretical results, we first design the Lyapunov-based controller $u_1 = h(x)$ which can stabilize the closed-loop system and the explicit expression of the controller can be found in Liu et al. (in press). We consider a Lyapunov function $V(x) = x^T P x$ with P being the following weight matrix

$$P = diag\,^5\,(5.2 \times 10^{12}\ [4\ 4\ 10^{-4}\ 4\ 4\ 10^{-4}\ 4\ 4\ 10^{-4}]).$$

The values of the weights in P have been chosen in a way such that the Lyapunov-based controller $h(x)$ satisfies the input constraints, stabilizes the closed-loop system and provides good closed-loop performance.

Based on the Lyapunov-based controller $h(x)$, we design LMPC 1 and LMPC 2. The prediction horizons of both LMPC 1 and LMPC 2 are chosen to be $N = 6$ and N_R is

[5] $diag(v)$ denotes a matrix with its diagonal elements being the elements of vector v and all the other elements being zeros.

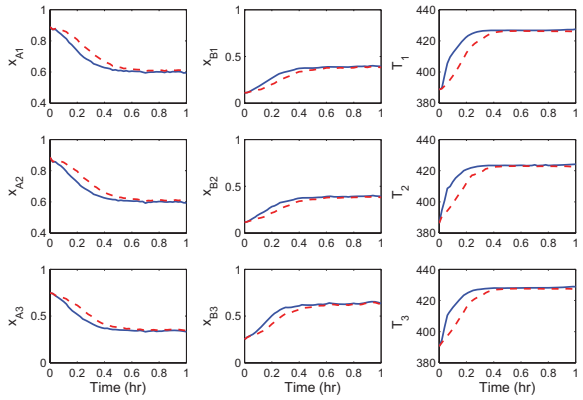


Fig. 3. State trajectories of the process under the proposed distributed LMPC design (8) (solid lines) and the original distributed LMPC design in Liu et al. (in press) (dashed lines) under continuous measurements.

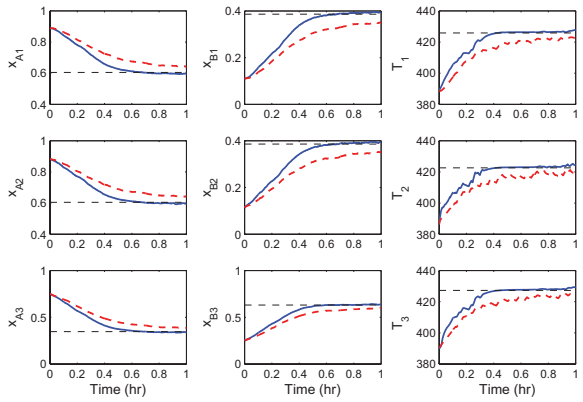


Fig. 4. State trajectories of the process under the proposed distributed LMPC design (8) (solid lines) and the original distributed LMPC design in Liu et al. (in press) (dashed lines) in the presence of asynchronous measurements.

chosen to be 4 so that $N_R\Delta \geq T_m$. The weight matrices for the LMPC designs are chosen as: $Q_c = diag(10^3 Q_v)$ with $Q_v = [2\ 2\ 0.0025\ 2\ 2\ 0.0025\ 2\ 2\ 0.0025]$, $R_{c1} = diag([5 \cdot 10^{-12}\ 5 \cdot 10^{-12}\ 5 \cdot 10^{-12}])$ and $R_{c2} = 100$.

We first carried out simulations to compare the proposed distributed LMPC design (8) with the original distributed LMPC design in Liu et al. (in press) in the case where no asynchronous measurements are present (i.e., state measurements $x(t_k)$ are available continuously with the interval between two successive measurements being Δ). The state trajectories under the two control designs are shown in Fig. 3. From Fig. 3, we can see that both the proposed and the original distributed LMPC designs stabilize the closed-loop system at the desired steady state.

We also carried out another set of simulations to compare both control laws in the presence of asynchronous measurements. To model the time sequence $\{t_{k\geq 0}\}$, we use an upper bounded random Poisson process. The Poisson process is defined by the number of events per unit time W. The interval between two successive concentration sampling times (events of the Poisson process) is given by $\Delta_a = \min\{-ln\chi/W, T_m\}$, where χ is a random variable with uniform probability distribution between 0 and 1. This generation ensures that $\max_k\{t_{k+1} - t_k\} \leq T_m$. In this example, W is chosen to be $W = 20$. The state trajectories of the system in closed-loop with both controllers are shown in Fig. 4. From Fig. 4, we can see that the proposed distributed LMPC design, which takes into account asynchronous measurements explicitly, can stabilize the closed-loop state at the desired steady state; however, the original distributed LMPC design failed to drive the closed-loop state to the desired steady state.

REFERENCES

Christofides, P.D., Davis, J.F., El-Farra, N.H., Clark, D., Harris, K.R.D., and Gipson, J.N. (2007). Smart plant operations: Vision, progress and challenges. *AIChE Journal,* Perspective article, 53, 2734–2741.

Dunbar, W.B. (2007). Distributed receding horizon control of dynamically coupled nonlinear systems. *IEEE Transactions on Automatic Control*, 52, 1249–1263.

Keviczky, T., Borrelli, F., and Balas, G.J. (2006). Decentralized receding horizon control for large scale dynamically decoupled systems. *Automatica*, 42, 2105–2115.

Lin, Y., Sontag, E.D., and Wang, Y. (1996). A smooth converse lyapunov theorem for robust stability. *SIAM Journal on Control and Optimization*, 34, 124–160.

Liu, J., Muñoz de la Peña, D., Ohran, B., Christofides, P.D., and Davis, J.F. (2008). A two-tier architecture for networked process control. *Chemical Engineering Science*, 63, 5394–5409.

Liu, J., Muñoz de la Peña, D., and Christofides, P.D. (2009). Distributed model predictive control of nonlinear process systems. *AIChE Journal*, 55, 1171-1184.

Magni, L. and Scattolini, R. (2006). Stabilizing decentralized model predictive control of nonlinear systems. *Automatica*, 42, 1231–1236.

Massera, J.L. (1956). Contributions to stability theory. *Annals of Mathematics*, 64, 182–206.

Muñoz de la Peña, D. and Christofides, P.D. (2008). Lyapunov-based model predictive control of nonlinear systems subject to data losses. *IEEE Transactions on Automatic Control*, 53, 2076–2089.

Neumann, P. (2007). Communication in industrial automation: What is going on? *Control Engineering Practice*, 15, 1332–1347.

Raimondo, D.M., Magni, L., and Scattolini, R. (2007). Decentralized MPC of nonlinear system: An input-to-state stability approach. *International Journal of Robust & Nonlinear Control*, 17, 1651–1667.

Rawlings, J.B. and Stewart, B.T. (2007). Coordinating multiple optimization-based controllers: New opportunities and chanllenges. In *Procedings of 8th IFAC Symposium on Dynamics and Control of Process*, volume 1, 19–28. Cancun, Mexico.

Richards, A. and How, J.P. (2007). Robust distributed model predictive control. *International Journal of Control*, 80, 1517–1531.

Ydstie, E.B. (2002). New vistas for process control: Integrating physics and communication networks. *AIChE Journal*, 48, 422–426.

Predictive Control of Nonlinear Chemical Processes under Asynchronous Measurements and Controls

P. Varutti * **and R. Findeisen** *

* *P.Varutti and R.Findeisen are with the Institute of Automation Engineering, Otto-von-Guericke University, 39016 Magdeburg, Germany ({paolo.varutti,rolf.findeisen}@ovgu.de).*

Abstract: In many process control problems measurement and control instances might not be available in a periodically-equally-distributed way. Moreover, due to the sensor processing time, actuators/sensors calibration, or computation, inevitable delays can often arise. Also information losses caused, for example, by temporary components failure, or the presence of unreliable communication media, might represent a non-trivial problem. This leads to asynchronous availability of measurement and control inputs, i.e. the controller, sensor, and actuator work in an event-driven, rather than a continuous way. In order to avoid instability and performance loss all these issues must be considered during the control design. In this paper, it is shown that predictive control methods based on continuous time models can be used to stabilize event-based nonlinear systems under variable delays, and limited information losses. It is demonstrated that by using the suggested approach asymptotic convergence is ensured.

Keywords: nonlinear model predictive control, continuous time systems, event-based control, delays, information losses, process control

1. INTRODUCTION

In many cases, continuous time systems are controlled by means of periodically-equally-distributed sampling times, commonly assumed to be known a priori. However, practical control problems are often intrisictly asynchronous, i.e. the dynamics of the system depends on some –maybe exogenous– event. Examples are multi fold: sensors such as chromatographers or laboratory measurements of compositions could need long time due to calibration or limited processing capabilities. The energy of an actuator might be limited, e.g. it must first be "charged" before applying the input. Measurements/actuation might demand human interaction, often unpredictable or inefficient. Moreover, it happens frequently that systems are subject to input/output delays caused, for example, by computational time, communication, and/or sensor/actuator slow dynamics. It might also occur that part of the exchanged information is lost, e.g. due to components failure, or the use of unreliable communication media. If these issues are not taken into account, performance loss or instability of the closed loop can arise. Event-based and asynchronous control is a very active field, Brockett and Liberzon (1998); Heemels et al. (2008). However, most of the work focuses only on linear unconstrained systems, without considering explicitly either time delays or information losses.

In this paper, a solution for the formerly introduced problems, especially suitable to process control applications, is presented. In particular, the suggested solution relies on Predictive Control (PC), which fits well the nature of event-based/asynchronous systems, since the sampling times do not have to neither be equally distant nor to be known a priori (see Fontes (2001); Findeisen (2006)). The compensation capabilities of PC with respect to measurement and computational delays have already been assessed in inter alia Chen et al. (2000); Findeisen and Allgöwer (2004)), where asymptotic stability with respect to such delays has been established. In this work, we focus on the complete asynchronous case, including delays and losses on the actuation and measurement side, which is significantly more challenging. By using smart-sensors and smart-actuators, i.e. components capable of full-duplex communication with the controller, it is possible to achieve closed loop asymptotic convergence.

In the next section PC and the problem under consideration are formally presented. In Section 3, an asynchronous PC solution to compensate delays and information losses is introduced. Results on asymptotic convergence are provided. Simulation results on a Continuous Stirred Tank Reactor (CSTR) are reported in Section 4.

2. PROBLEM STATEMENT

We consider the problem of controlling the nonlinear time-continuous process of the form

$$\dot{x} = f(x,u), \quad x(0) = x_0, \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m. \tag{1}$$

It is assumed that the whole state x is available only at discrete instants t_i. The objective is to stabilize the system around the origin, i.e. $\|x\| \to 0$ for $t \to \infty$, under the state and input constraints $x \in X \subset \mathbb{R}^n$, $u \in U \subset \mathbb{R}^m$. The state constraints X, e.g. max temperature, and the input constraints U, e.g. max valve opening, are assumed to be closed sets. It is also assumed that $f(0,0) = 0$, and f is sufficiently differentiable. The controller should provide for every state measurement $x(t_i)$ a piece of input trajectory

$$u(t) = u(t; x(t_i)), \text{ for } t \in (t_i, t_{i+1}], \quad (2)$$

i.e. the calculated input trajectory is applied open loop in between consecutive recalculation times. This kind of control is commonly called sampled-data open loop feedback. Note that the recalculation times t_i do not need to be known a priori, e.g. in the event-based case when a measurement is triggered once some conditions, such as deviation from the product specifications, are met. This fits well such an asynchronous frame, and it can be used to provide better performance, thanks to the possibility of adjusting on-the-fly the recalculation frequency.

2.1 Predictive Control

In this section, we summarize the basic idea of PC (for more details see Mayne et al. (2000); Findeisen (2006)). The idea is to use a model of the process to be controlled, in order to repeatedly solve an optimization problem, based on the state prediction provided by the model itself. Then, only the first piece of trajectory is implemented and the problem is re-solved with the new measurement. The following definition will be useful for the remainder of the paper.

Definition 1. (Partition). Every series $\pi = (t_i)$, $i \in \mathbb{N}$, $t_i \in \mathbb{R}^+$, such that $t_0 = 0$, $t_i < t_{i+1}$ and $t_i \to \infty$ is called partition.

For every $t_i \in \pi$, $x(t_i)$ is measured, and

$$\min_{\bar{u}(\cdot)} \int_{t_i}^{t_i+T_p} F(\bar{x}(\tau), \bar{u}(\tau))d\tau + E(\bar{x}(t_i + T_p)), \quad (3a)$$

$$\text{s.t. } \dot{\bar{x}}(t) = f(\bar{x}(t), \bar{u}(t)), \quad \bar{x}(t_i) = x(t_i), \quad (3b)$$

$$\bar{u}(t) \in U, \quad t \in (t_i, t_{i+1}], \quad (3c)$$

$$\bar{x}(t) \in X, \quad (3d)$$

$$\bar{x}(t_i + T_p) \in \mathcal{E}, \quad (3e)$$

is solved, where $\bar{\cdot}$ denotes the controller internal variables. The solution is an optimal control signal $u^*(t; x(t_i))$, for $t \in [t_i, t_i + T_p]$, where T_p represents the finite prediction horizon. For sake of simplicity, prediction and control horizon are supposed to be equal, i.e. $T_c = T_p$. The control input is then implemented for the time-span $(t_i, t_i + \delta]$, i.e.

$$u(t) = u^*(t; x(t_i)), \text{ for } t \in (t_i, t_i + \delta], \quad (4)$$

where δ represents the interval between two consecutive recalculation times, i.e. $\delta = (t_{i+1} - t_i), \forall t_i, t_{i+1} \in \pi$. Stability can then be achieved by properly choosing the cost functional $F(x, u)$, the terminal cost $E(x)$, the terminal region $\mathcal{E} \subset X$, and the prediction horizon T_p, see Mayne et al. (2000); Fontes (2001); Findeisen (2006). As formerly mentioned, it is commonly assumed that the recalculation intervals $\delta = (t_{i+1} - t_i)$ are constant and known a priori. Here, however, these assumptions are relaxed, allowing for the recalculation instants to be time-varying and unknown a priori. The only requirement on δ is given by Assumption 2.

Assumption 2. Given the prediction horizon T_p, $\beta \in \mathbb{R}^+$,

$$\beta < \delta = (t_{i+1} - t_i) < T_p, \forall t_i, t_{i+1} \in \pi. \quad (5)$$

In the remainder of the paper, $\overline{\delta}$ will be used to refer to the maximum recalculation interval $\delta = (t_{i+1} - t_i)$. Additionally, the following theorem will be useful for the final results.

Theorem 3. (Asynchronous Predictive Control).
Consider the closed-loop system given by (1), (3)-(4). If

i) Assumption 2 is satisfied.
ii) $\forall x_0 \in \mathcal{E} \subseteq X, \exists \bar{u}(\tau) \in U, \tau \in [0, T_p]$ where

$$x(\tau) \in \mathcal{E}, \quad (6a)$$

$$\text{for } \dot{x}(\tau) = f(x(\tau), \bar{u}(\tau)), \quad x(0) = x_0, \quad (6b)$$

$$\text{and } \frac{\partial E}{\partial x} f(x(\tau), \bar{u}(\tau)) + F(x(\tau), \bar{u}(\tau)) \leq 0. \quad (6c)$$

iii) The optimal control problem is solvable for a time t_0.

Then, $\lim_{t \to \infty} \|x(t)\| = 0$.

Proof. The proof comes directly from Findeisen (2006) and the results about PC stability, see inter alia Fontes (2001); Mayne et al. (2000). It must be ensured that δ is smaller than T_p.

Remark 4. The solution of the optimal control problem, as well as the closed loop stability, are based only on the discrete time measurements $x(t_i)$ at $t_i \in \pi$, where π does not have to be known a priori. This makes PC a very appealing solution for event-based problems.

Remark 5. Note that Theorem 3 states only asymptotic converge, and not asymptotic stability in the Lyapunov sense. The former is a weaker property, meaning that even withouth disturbances, the system can temporary drift away from the equilibrium point before converging to the equilibrium. Proving asymptotic stability would in the first step require to rigorously define stability for discrete event systems, since the partition π is not known a priori. This is way beyond the focus of this paper.

2.2 Delays and Information Losses

In a closed loop controlled system, it is quite common to face delays and/or information losses, e.g. due to components failures. Essentially, there can be three delay sources (see Figure 1):

i) *Measurement delays*, which can be due to measurement elaboration, observer reconstruction, slow sensor dynamics, but also the time required for a signal to reach the controller.
ii) *Computational delays*, which represent the time required by the controller to calculate the control input.
iii) *Actuation delays*, which can be due to slow actuator dynamics, but also signal transportation.

The following assumption on the delays is made:

Assumption 6. $\tau_s(t), \tau_c(t), \tau_a(t)$ are nondeterministic with arbitrary probability distribution, but ultimately limited, i.e.

$$\tau_s(t) \in [0, \overline{\tau_s}], \tau_c \in [0, \overline{\tau_c}], \tau_a(t) \in [0, \overline{\tau_a}]. \quad (7)$$

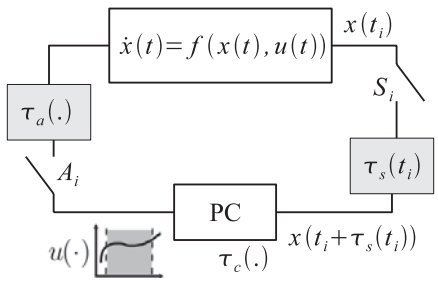


Fig. 1. Sketch of an event-based system subject to delays and information losses.

Both the sensor-to-controller and controller-to-actuator channel can suffer from information losses, which, for example, can be modeled as Bernoullian variables $A_i \sim \mathcal{B}(1-p_a)$, and $S_i \sim \mathcal{B}(1-p_s)$, such that

$$A_i = \begin{cases} 1, & \text{if a control input is received} \\ 0, & \text{otherwise} \end{cases},$$

$$S_i = \begin{cases} 1, & \text{if a measurement is received} \\ 0, & \text{otherwise} \end{cases}.$$

p_a, and p_s represent the loss probabilities for the actuation and the measurement link respectively. Sensor, controller, and actuator are event-driven, such that measurement and control information are dispatched only when necessary. It is also assumed that the following statements are fulfilled.

Assumption 7. Either a common global time, or a set of synchronized clocks is available, such that a common unique time t is established among the components.

Assumption 8. All exchanged information is time-stamped.

3. ASYNCHRONOUS PC

Model-based approaches, such as PC, represent an intuitive and natural way to handle input/output delayed systems. In this section, we show how PC can be used in an event-based way to control asynchronous systems by preserving stability, in the sense of asymptotic convergence, and simultaneously reducing both exchanged information and computational requirements. Note that delays are very common on a daily basis, therefore the presented approach represents a good solution for a wide class of problems, e.g control under actuator/sensor slow dynamics, heavy computation, control over networks, and/or limited resources.

3.1 Compensating Delays

Measurement Delays Assume for the moment that no information is lost. When a measurement $[x(t_i)|t_i]$ is dispatched at a time $t_i \in \pi$, where $x(t_i) \in \mathcal{X}$ is the state value, while t_i is its time-stamp, if there is a measurement delay $\tau_s(t)$, then the information will be available to the controller only at $(t_i + \tau_s(t_i))$, i.e. the controller has to use some piece of information which is outdated and does not correspond to the actual state of the system under control. Therefore, it is necessary to compensate this delay in order to solve the correct control problem. Since a model of the system is available at the controller side, and no mismatch is present, under Assumptions 7-8, it is possible to determine the delay simply by comparing the time-stamp with the global time t, i.e. $\tau_s(t_i) = (t - t_i)$. By means of forward prediction through the local model, possible since it is known what input is applied to the plant (no actuation delay), one can obtain the state prediction

$$\overline{x}(t_i + \tau_s(t_i)) = x(t_i + \tau_s(t_i)). \tag{8}$$

Theorem 9. (Measurement Delay Compensation).
Given the closed loop system (1),(3)-(4), if

i) Theorem 3 is satisfied in the nominal case, i.e. without measurement delays.
ii) $T_p > \overline{\tau_s} + \overline{\delta}$.

Then, $\lim_{t\to\infty} \|x(t)\| = 0$.

Proof. The proof follows from Theorem 3, when the state prediction (8) is used to compensate the delay $\tau_s(t_i)$. More details can be found in Findeisen (2006).

Compensation of Actuation Delays Less trivial is the compensation of computational and actuation delays. In this case, in fact, if the delays are nondeterministic, the actual applied input is not known for sure to the controller. Thus, it is not possible to obtain (8) correctly since $u^*(\cdot)$ is not uniquely determined. As formerly stated, this kind of delays are common in real application, e.g. due to slow actuator dynamics or reduced computational capabilities, and if not explicitly considered can worsen considerably the closed loop performance, or bring to instability. To solve this problem, in some way the applied input must be made deterministic. This can be achieved by using future input trajectories and buffer them in the actuator till the moment they can be used, see Alldredge and M. S. Branicky (2008); Findeisen and Varutti (2009); Varutti and Findeisen (2008). In fact, since Assumption 6 must hold, one can consider the worst case for $\tau_a(t)$ and $\tau_c(t)$, namely $\overline{\tau_a}$ and $\overline{\tau_c}$, in which the state prediction

$$\overline{x}(t_i + \tau_s(t_i) + \overline{\tau_a} + \overline{\tau_c}) = x(t_i) + \int_{t_i}^{t_i+\tau_s(t_i)+\overline{\tau_a}+\overline{\tau_c}} f(x(\tau), u(\tau))d\tau \tag{9}$$

is obtained by using the measurement $x(t_i)$. (9) is then used to solve the optimal control problem and the corresponding solution is despatched to the actuator with a new time-stamp, buffered and used once its time-stamp matches with the global time t, i.e.

$$u^*(\tau; \overline{x}(t_i + \tau_s(t_i) + \overline{\tau_a} + \overline{\tau_c})), \tag{10}$$

for $\tau \in (t_i + \tau_s(t_i) + \overline{\tau_a} + \overline{\tau_c}), t_{i+1} + \tau_s(t_{i+1}) + \overline{\tau_a} + \overline{\tau_c})]$ is sent as $[u^*(\cdot)|(t_i + \tau_s(t_i) + \overline{\tau_a} + \overline{\tau_c})]$. The overall algorithm is reported in Algorithm 1, Appendix A. It can be proved that under Assumption 7, and

$$T_p > \overline{\delta} + \overline{\tau_s} + \overline{\tau_a} + \overline{\tau_c}, \tag{11}$$

Algorithm 1 stabilizes the delayed system.

Theorem 10. (Worst Case Compensation).
Given the nonlinear continuous time system (1) and the the predictive controller obtained from (3)-(4), and (9), by applying Algorithm 1, under (11) the closed loop system is stable, in sense of asymptotic convergence, if the nominal controller, i.e. the controller subject to no delays obtained from Theorem 3, stabilizes the system.

Proof. The proof follows directly from Theorem 3, and 9, first, by proving recursive feasibility, and then convergence –see Findeisen (2006); Findeisen and Varutti (2009); Varutti and Findeisen (2008) for more details–.

Remark 11. Assumption 7 is required to have a common time-frame among the components. This can represent a problem for fast dynamical systems, since the state-of-the-art synchronization algorithms cannot guarantee high precision.

3.2 Information Loss Compensation

It has been assumed till now that the communication is not affected by any information loss. In reality, however, information losses might occur due, for example, to unreliable communication media, or some temporary components failure. As in the delay case, the major problem is represented by losses/failures in the actuation channel. In fact, if an information loss $S_i = 0$ occurs for $t_i \in \pi$, the controller can still use the last available state

$$x(t_k), \text{ for } t_k \in \pi, \text{ s.t. } S_k = 1, \tag{12}$$

and the nominal model for (1) to calculate the prediction

$$\overline{x}(t_k + \sum_{j=k}^{i} \tau_s(t_j) + \overline{\tau_a} + \overline{\tau_c}), \text{ for } t_j \in \pi, \text{ s.t. } S_j = 0, \quad (13)$$

by using the compensation approach presented in Section 3.1, if the applied input is uniquely determined.

Remark 12. Notice that $t_k + \sum_{j=k}^{i} \tau_s(t_j)$ can be substituted by the global time t, moment in which the controller receives a new measurement.

On the contrary, if a dropout $A_i = 0$ occurs, then the control input is not uniquely known at the controller side, and thus (13) cannot be accurately calculated. In Varutti and Findeisen (2008); Findeisen and Varutti (2009) a solution for the problem was found by using *prediction consistent feedbacks*, i.e. feedbacks that under information losses are able to keep the difference between state prediction and actual state negligible and hence guarantee convergence under a limited amount of dropouts.

In this paper, a different approach is considered. In particular, the following assumptions are made:

Assumption 13. An acknowledge mechanism is available on the actuator side.

Assumption 14. The acknowledgments have high priority, and they cannot be dropped.

Assumption 15. The acknowledgments are delivered instantaneously.

This is equal to saying that for every input received by the actuator an acknowledgment with the time-stamp of the latest successfully delivered information is sent back to the controller.

We show later how Assumption 15 can be relaxed. On the contrary, Assumption 14 is a fundamental condition since it is well known from theoretical results that no handshake protocol can solve a coordination problem under acknowledgment losses –see "The two Generals Paradox", Tanenbaum (2008)–. In the case of chemical processes, however, timely submission of acknowledgments is often not a problem, since there are frequently slow measurement and actuation devices. The used communication networks are often sufficiently fast and provide the possibility of having high priority acknowledgments. Although restrictive, these conditions are necessary to allow the controller to correctly reconstruct the applied input sequence. Algorithm 2 in Appendix A illustrates the procedure to compensate simultaneously delays and information losses. Differently from Algorithm 1, the entire control input trajectory $u^*(\tau; x(t_i))$, for $\tau \in (t + \overline{\tau_a} + \overline{\tau_c}, t_i + T_p]$ is sent to the actuator with time-stamp $(t + \overline{\tau_a} + \overline{\tau_c})$. In this way, when some information is dropped either in the down- or in the up-link, the actuator can still utilize the old input trajectory to control the system. Note that in an event-based setup measurement losses are transparent to the controller. This means that no new control input is generated and dispatched to the actuator. However, since the whole control trajectory is sent, if the number of consecutive losses S_i is less than T_p, the actuator can still apply the latest received input. On the other hand, the controller can establish immediately that some information has gone lost thanks to the timer that is implicitly set by time-stamping the control inputs. In fact, if the current time exceeds the former time-stamp, from Assumption 15, it is known for sure that no new control input has arrived. An alternative would be to index the control trajectories and use an error mechanism instead, i.e. an error is sent every time some control information is lost. However, in the asynchronous case utilizing acknowledgments is more efficient, since the actuator does not have to wait till the next successfully received trajectory to realize that some information went lost. Finally, note that a non-resend policy has been chosen, i.e. if a control trajectory is lost, the controller does not transmit it again but it simply records the event in order to update its local copy of the currently applied control input. This seems to be a more logical choice since the sequence would probably arrive when its applicability time is already expired.

Theorem 16. (Convergence Under Information Losses). Given the closed loop system (1), (3)-(4), (13) if

i) Assumptions 13-15 are satisfied.
ii) Algorithm 2 is used.
iii) The prediction horizon T_p is such that

$$T_p > \overline{\delta} + n \cdot \overline{\tau_s} + \overline{\tau_c} + m \cdot \overline{\tau_a}, \quad (14)$$

where n,$m \in \mathbb{N}/\{0\}$ represent, respectively, the number of consecutive losses in the measurement and in the actuation.

Then, $\lim_{t\to\infty} \|x(t)\| = 0$.

Proof. The proof follows from Theorem 3 and 9. The use of acknowledgments allows the controller to reconstruct in real time the correct applied control sequence. Thus, feasibility and convergence can be proved.

As formerly stated, Assumption 15 can be relaxed by allowing the acknowledgments to be subject to nondeterministic delays $\tau_{ack}(t)$. In this case, however, the following assumption is required.

Assumption 17. $\tau_{ack}(t)$ is nondeterministic with arbitrary probability distribution, but limited, i.e $\tau_{ack}(t) \in [0, \overline{\tau_{ack}}]$.

By using a worst case compensation approach similar to Algorithm 1, and 2, one can ensure asymptotic convergence of the closed loop system by considering, instead, the state prediction

$$\overline{x}(t_k + \sum_{j=k}^{i} \tau_s(t_j) + \overline{\tau_a} + \overline{\tau_c} + \overline{\tau_{ack}}), \quad (15)$$

for $t_j \in \pi$, such that $S_j = 0$.

Corollary 18. (Convergence Under Information Losses). The closed loop system (1), (3)-(4), is stable, in the sense of asymptotic convergence, if

i) The conditions for Theorem 3, 9, 10, and Assumption 17 are satisfied.
ii) A modified variant of Algorithm 2, such that the prediction (15) is used and $[u^*(\tau; x(t_i))|ts]$, for $ts = t+\overline{\tau_{ack}}+\overline{\tau_s}+\overline{\tau_c}+\overline{\tau_a}$, $\tau \in [t+\overline{\tau_{ack}}+\overline{\tau_s}+\overline{\tau_c}+\overline{\tau_a}, t_i+T_p]$ is utilized.
iii) $T_p > n \cdot \overline{\tau_s} + \overline{\tau_c} + m \cdot \overline{\tau_a} + \overline{\tau_{ack}}$, where n,$m \in \mathbb{N}/\{0\}$ represent the number of consecutive losses in the measurement/actuation.

4. SIMULATION RESULTS FOR A CSTR

The formerly presented method has been applied to a CSTR, where an irreversible exothermic reaction, $A \rightarrow B$, takes place in a constant volume, cooled by a single coolant stream at temperature T_c –see Figure 2–. The overall

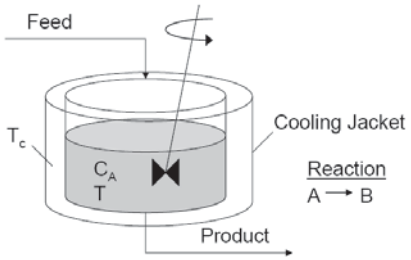


Fig. 2. Scheme of CSTR under study.

system is modeled as:

$$\dot{C}_A(t) = \frac{F}{V}(C_{Af} - C_A(t)) - k_0 C_A(t) e^{-\frac{E}{RT_r(t)}},$$
$$\dot{T}_r(t) = \frac{F}{V}(T_{rf} - T_r(t)) - \frac{k_0 \Delta H}{\rho c_p} C_A(t) e^{-\frac{E}{RT_r(t)}} + \frac{UA}{\rho c_p V}(T_c(t) - T_r(t)).$$

The meaning and the values of all the parameters are explained in Henson and Seborg (1997). Under the nominal condition $T_c^{nom} = 103.4$ K, the system has the three equilibrium points depicted in Figure 3. The objective is to stabilize the unstable saddle point $(0.52, 398.97)$ by manipulating the control input $u(t) = T_c(t)$ under the input constraints $T_c(t) \in [275, 370]$ K. If we represent

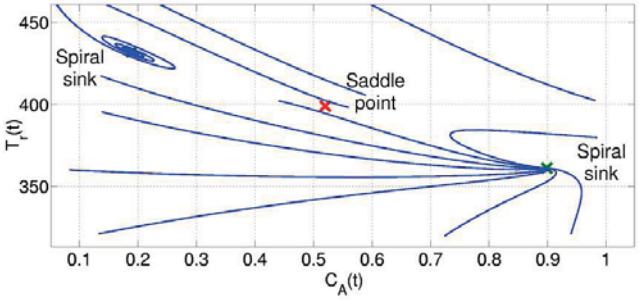


Fig. 3. Phase plot of the system for the nominal case $T_{c,nom} = 302$ K.

the state as the vector $x(t) = [C_A(t) \quad T_r(t)]^T$, the cost functional to be minimized is given by

$$J(u, x) = \int_{t_i}^{t_i + T_p} (x^T Q x + u^T Q u) d\tau,$$

where $Q = I$, $R = 1$, and $T_p = 1.5$ min. For sake of simplicity, the control trajectory is held constant between consecutive recalculation times. Terminal penalty and terminal region constraints have been chosen such that closed loop stability in the nominal case is achieved.

In Figure 4, the results for the asynchronous nominal case (no delays and no failures) are presented. Compared to a classical sampling approach with constant recalculation interval $\delta = 0.15$ min, the asynchronous controller presented in Theorem 3 obtains extremely similar results by saving up to 30% of computational effort and exchanged information. The partition π is implicitly determined by the absolute error on the product concentration, i.e. the sensor regularly checks the concentration $C_A(t)$, but it sends a measurement to the controller only when

$$\|C_A(t) - 0.52\| > \epsilon, \text{ with } \epsilon = 2 \cdot 10^{-3},$$

for longer than 0.15 minutes. In a second simulation

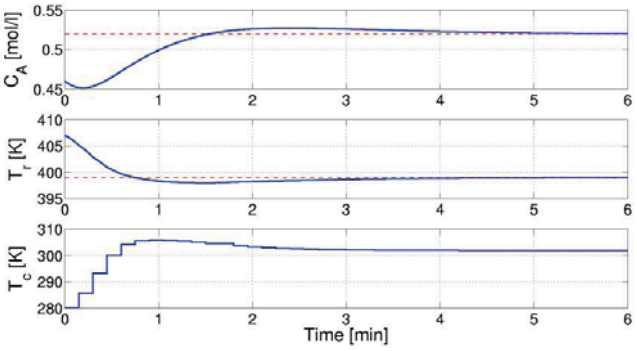


Fig. 4. Event-based controller under nominal conditions.

both measurement and actuation delays are considered. It is assumed that τ_s lies between 0 and 15 seconds, while τ_a between 0 and 5 seconds. Both are modeled as uniform variables. For sake of simplicity, no computational delay is considered. However, notice that this delay could be implicitly considered as part of the input delay τ_a. Furthermore, it is assumed that the probability loss at the actuator side p_a is equal to 5%, while no information from the sensor is lost. It is also supposed that Assumption 15 is verified. The results for the compensated and the uncompensated case are presented in Figure 5. As one can see, the proposed method is able to stabilize effectively the unstable saddle point with performance comparable to the nominal case, reported in Figure 4. Note that the controllers presented in Findeisen and Allgöwer (2004); Chen et al. (2000) are not able to handle delays on the actuation side at all.

5. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, it was shown how PC can be used to control event-based/asynchronous systems, such as chemical processes, in which the recalculation times do not need to be known a priori. Moreover, it was depicted how delays and information losses/failures on the actuators/sensors, common in control systems, need to be taken into account to avoid instability. Whereas delays can be compensated easily with forward prediction, one can exploit bidirectional communication with the actuators in order to establish an acknowledgment mechanism to counteract information losses/failures. Two algorithms able to guarantee asymptotic convergence for nonlinear continuous time systems were presented. Through the simulation of a CSTR, it was shown firstly that asynchronous PC can actually reduce both computational requirements and exchanged information, but also compensate effectively delays and information losses while keeping closed loop stability and good performance. Future work should concentrate on how to include directly in the optimization problem the

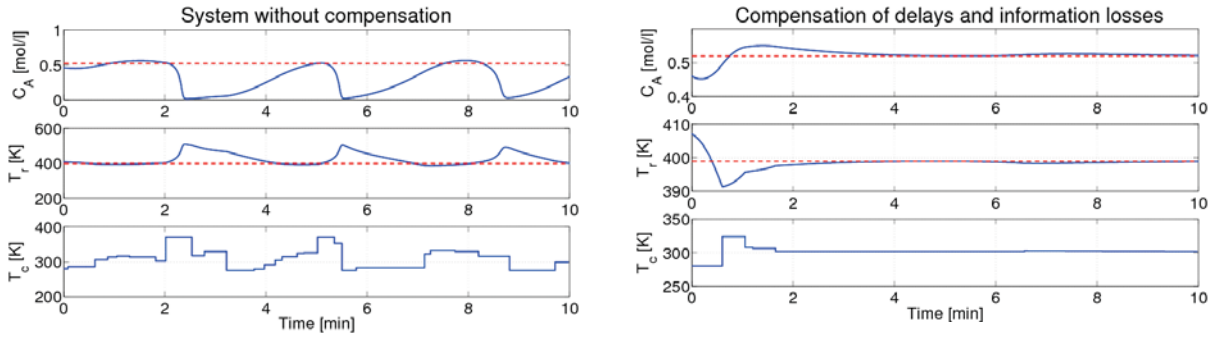


Fig. 5. Results for the closed loop system subject to delays and information losses with and without compensation.

exchanged information. Moreover, the method should be extended to include also robustness.

REFERENCES

Alldredge, G. and M. S. Branicky, V.L. (2008). Play-back buffers in networked control systems. evaluation and design. *Amer. Cont. Conf.*, 3106–3113.

Brockett, R. and Liberzon, D. (1998). Quantized feedback systems perturbed by white noise. In *Proc. 37th IEEE Conference on Decision and Control*, volume 2, 1327–1328.

Chen, W., Ballance, D., and O'Reilly, J. (2000). Model predictive control of nonlinear systems: Computational delay and stability. In *IEE Proceedings, Part D 147(4)*, 387–394.

Findeisen, R. (2006). *Nonlinear Model Predictive Control: a Sampled-Data Feedback Perspective*, volume 8. VDI Verlag.

Findeisen, R. and Allgöwer, F. (2004). Computational delay in nonlinear model predictive control. In *Proc. Int. Symp. Adv. Contr. of Chem. Proc.*, 427–432. Hong Kong, PRC.

Findeisen, R. and Varutti, P. (2009). Stabilizing nonlinear predictive control over nondeterministic communication networks. In L. Magni, D. Raimondo, and F. Allgwer (eds.), *Nonlinear Model Predictive Control: Towards New Challenging Applications*, Lecture Notes in Control and Information Sciences. Springer. In printing.

Fontes, F. (2001). A general framework to design stabilizing nonlinear model predictive controllers. *Sys. Contr. Lett.*, 42(2), 127–143.

Heemels, W.P., Sandee, J.H., and Van den Bosch, P. (2008). Analysis of event-driven controllers for linear systems. *Inter. J. of Cont.*, 81, 571–590.

Henson, M. and Seborg, D. (1997). *Feedback Linearizing Control*. Prentice Hall.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., and Scokaert, P.O.M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789–814.

Tanenbaum, A.S. (2008). *Computer Networks*. Prentice Hall PTR, 4th edition.

Varutti, P. and Findeisen, R. (2008). Compensating network delays and information loss by predictive control methods. (submitted to ECC09).

Appendix A

Algorithm 1 (Worst Case Compensation)

$\forall t_i \in \pi$; t = current time;

Sensor:

(1) Measure $x(t_i)$.
(2) Send $[x(t_i)|ts]$, with $ts = t_i$, to the controller.
(3) Go to 1.

Controller:

`buffer` $= [x(t_i)|ts]_{old}$;
`control_input` $= \{[u^*(\cdot)|ts_0]\}$;
(1) If $[x(t_i)|ts]_{new}$ arrives
  a) If $ts_{new} \leq ts_{old}$, then discard.
  b) Else `buffer` $= [x(t_i)|ts]_{new}$.
(2) $\tau_s = (t - t_i)$ for $t_i = ts$.
(3) Calculate (9), from $u^*(\cdot; x(t_i)) \in$ `control_input`.
(4) Solve the o.c.p. for (9) $\longrightarrow$
  $u^*(\tau; x(t_i))$, for $\tau \in (t_i + \tau_s(t_i) + \overline{\tau_a} + \overline{\tau_c}, t_{i+1} + \tau_s(t_{i+1}) + \overline{\tau_a} + \overline{\tau_c}]$.
(5) Send $[u^*(\tau; x(t_i))|ts]$, with $ts = (t + \overline{\tau_a} + \overline{\tau_c})$.
(6) Insert $[u^*(\tau; x(t_i))|ts]$ in `control_input`.
(7) Go to 1.

Actuator:

`buffer` $= \{[u^*(\cdot)|ts_0], \ldots, [u^*(\cdot)|ts_n]\}$, for $ts_0 < t < ts_1 \ldots < ts_n$;
`applied_input` $= [u^*(\cdot)|ts_0]$;
(1) If $[u^*(\cdot)|ts]_{new}$ arrives
  a) Insert $[u^*(\cdot)|ts]_{new}$ in `buffer`.
  b) Sort `buffer` by increasing ts.
(2) `temp` = first element of `buffer`.
(3) If $ts_{\texttt{temp}} = t$
  a) `applied_input` = `temp`.
  b) Remove first element from `buffer`.
(4) Go to 1.

Algorithm 2 (Information Loss Compensation)

$\forall t_i \in \pi$; t = current time;

Sensor:

(1) Measure $x(t_i)$.
(2) Send $[x(t_i)|ts]$, with $ts = t_i$.
(3) Go to 1.

Controller:

`buffer` $= [x(t_i)|ts]_{old}$;
`control_input` $= \{[u^*(\cdot)|ts_0]\}$;
`delivered` = `true`;
(1) If $[\cdot|ts]$ arrives:
(2) Case($[x(t_i)|ts]$):
  If $ts_{new} \leq ts_{old}$, then discard.
  Else `buffer` $= [x(t_i)|ts]_{new}$.
  $\tau_s = (t - t_i)$ for $t_i = ts$.
  Calculate (13), from $u^*(\cdot; x(t_i)) \in$ `control_input`.
  Solve the o.c.p. for (13) $\longrightarrow$
  $u^*(\tau; x(t_i))$, for $\tau \in (t + \overline{\tau_a} + \overline{\tau_c}, t_i + T_p]$.
  `delivered` = `false`.
  Send $[u^*(\tau; x(t_i))|ts]$, with $ts = (t + \overline{\tau_a} + \overline{\tau_c})$.
  Wait until `delivered`=`true` OR t $\geq$ ts.
  If `delivered` = `true`, then insert $[u^*(\tau; x(t_i))|ts]$ in `control_input`.
  Else if t $\geq$ ts, then `delivered` = `false`, use old input in `control_input`.
  Go to 1.
(3) Case($[ack|ts]$):
  Set `delivered` = `true`.
  Go to 1.

Actuator:

`buffer` $= \{[u^*(\cdot)|ts_0], \ldots, [u^*(\cdot)|ts_n]\}$,
for $ts_0 < t < ts_1 \ldots < ts_n$;
`applied_input` $= [u^*(\cdot)|ts_0]$;
(1) If $[u^*(\cdot)|ts]_{new}$ arrives
  a) Send $[ack|ts]$ to the controller.
  b) Insert $[u^*(\cdot)|ts]_{new}$ in `buffer`.
  c) Sort `buffer` by increasing ts.
(2) `temp` = first element of `buffer`.
(3) If $ts_{\texttt{temp}} = t$
  a) `applied_input` = `temp`.
  b) Remove first element from `buffer`.
(4) Go to 1.

Biological Systems

Oral Session

Analysis, Control, and Operational Optimization of a *Zymomonas mobilis* Reactor with Equilibrium Multiplicity

Jorge O. Trierweiler and Fabio C. Diehl

Group of Intensification, Modeling, Simulation, Control and Optimization of Processes – GIMSCOP
Department of Chemical Engineering, Federal University of Rio Grande do Sul – UFRGS
Rua Luiz Englert, s/n CEP: 90040-040 - Porto Alegre - RS - BRAZIL
E-MAIL: {jorge, fcdiehl}@enq.ufrgs.br

Abstract: For a successful application of any industrial *Z. mobilis* facility, it is necessary to have an efficient and simple control strategy. This paper analyzes the control and optimization problem of a continuous ZM bioreactor modeled by Jöbses et al. (1986). This system has steady state multiplicity in part of the operating range. The idea is to maintain the process close to the manifold border where is achievable the highest ethanol production. Based on a systematically analysis of the operational controllability using the nonlinear RPN indices it is identified that the process can be controlled using a linear controller. Finally the paper proposes a variable transformation that makes easy to maintain the bioreactor close to the optimum.

Keywords: nonlinear degree measurement, RPN methodology, bifurcation, bioreactor control, process optimization

1 INTRODUCTION

Zymomonas mobilis has attracted considerable interest over the past decades as a result of its unique metabolism and ability to rapidly and efficiently produce ethanol from simple sugars. However, despite its apparent advantages of higher yields and faster specific rates when compared to yeasts, no commercial scale fermentations currently exist which use *Z. mobilis* for the manufacture of fuel ethanol. In addition to ethanol depending on the substrate other fermentation products can occur, such as lactic acid, acetic acid, formic acid, acetone, and sorbitol. See (Rogers *et al.*, 2007), for a detailed review.

In the literature, *Zymomonas mobilis* has been proposed as a more promising microorganism than conventional yeast *Saccharomyces cerevisiae* for industrial production of ethanol (Rogers *et al.*, 2007). A major drawback of this microorganism is that it exhibits sustained oscillations (i.e., Hopf bifurcation) for low dilution rates (i.e., $D_f \leq 0.1\ h^{-1}$) when grown in continuous mode. This leads to decreased ethanol productivity and less efficient use of available substrate (Zhang and Henson, 2001). Various models have been proposed to describe the oscillatory dynamics of continuous *Zymomonas mobilis* cultures. Two of them are the Daugulis *et al.* (1997) and Jöbses *et al.* (1986) models. Even though the model predictions can be considered similar at low dilution rates, where the models have been fitted to the experimental data, they are quite different for higher dilution rates (Trierweiler and Diehl, 2009).

The Jöbses's model was fitted to experimental data with low dilution rate (i.e., $D_f \leq 0.1\ h^{-1}$) and middle inlet substrate concentration (i.e., $C_{S0} \cong 150\ kg/m^3$). Later, it was extrapolated outside of this operating region by Elnashaie *et al.* (2006), who have found a much more profitable operating region at higher dilution rates ($D_f \cong 2.0\ h^{-1}$) and inlet concentrations ($C_{S0} \cong 200\ kg/m^3$). Notwithstanding the Jöbses's model has not been validated at this region, our contribution will assume that this extrapolation is acceptable and we will propose a control strategy to maintain the system working at this more profitable operating region.

This paper is structured as follows: In section 2 the Jöbses's models is presented, section 3 it is analyzed the operational controllability which is used as basis for the proposed control strategy developed in section 4 and later validated by simulation. Final conclusions and remarks are then summarized in section 5.

2 MODEL DESCRIPTION AND OPERATING POINT DEFINITIONS

Since the Jöbses's model can predict a branch with higher ethanol production, which has been experimentally confirmed (at least for low dilution rates) by Elnashaie *et al.* (2006), we decide to analyze the control problem of a continuous bioreactor with the Jöbses *et al.* (1986) kinetic model, which is shortly described in the next subsection.

2.1 Model Description

The Jöbses's model consists of the following 4 differential equations:

$$\frac{dC_S}{dt} = -\frac{\mu_{max}.C_S.C_E}{Y_{SX}.(K_S+C_S)} - m_S.C_X + D_f.(C_{S0} - C_S)$$
$$\frac{dC_X}{dt} = \frac{\mu_{max}.C_S.C_E}{(K_S+C_S)} + D_f.(C_{X0} - C_X) \quad (1)$$
$$\frac{dC_E}{dt} = \frac{K_E.(C_P-c_1)(C_P-c_2).C_S.C_E}{(K_S+C_S)} + D_f.(C_{E0} - C_E)$$
$$\frac{dC_P}{dt} = \frac{\mu_{max}.C_S.C_E}{Y_{PX}.(K_S+C_S)} + m_P.C_X + D_f.(C_{P0} - C_P)$$

where C_S is the substrate (glucose) concentration, C_X is the biomass (*Z. mobilis*), C_P is the product (ethanol) concentration, and C_E is an auxiliary variable used to lag the effect of the ethanol concentration in the kinetic model. The insertion of the $K_E.(C_P - c_1)(C_P - c_2)$ parcel together with C_E makes possible the model to depict the oscillatory behavior of the Hopf bifurcation. In the model, dilution rate D_f is the inversion of the residence time and is defined as the relation between the inlet flow rate and the bioreactor volume. The model parameters are summarized in Table 1.

Table 1: Model parameters

Parameters	Values	Parameters	Values
$K_E\ \left[\frac{m^6}{kg^2.h}\right]$	0.00383	$m_S\ \left[\frac{kg}{kg.h}\right]$	2.160
$c_1\ \left[\frac{kg}{m^3}\right]$	59.2085	$m_P\ \left[\frac{kg}{kg.h}\right]$	1.100
$c_2\ \left[\frac{kg}{m^3}\right]$	70.5565	$(Y_{SX}, Y_{PX})\ \left[\frac{kg}{kg}\right]$	(0.02445, 0.05263)
$K_S\ \left[\frac{kg}{m^3}\right]$	0.500	$\mu_{max}\ \left[\frac{1}{h}\right]$	1.0

Fig. 1 shows the steady-state solutions for ethanol concentration (C_P) in function of dilution rate (D_f) and inlet substrate concentration (C_{S0}).

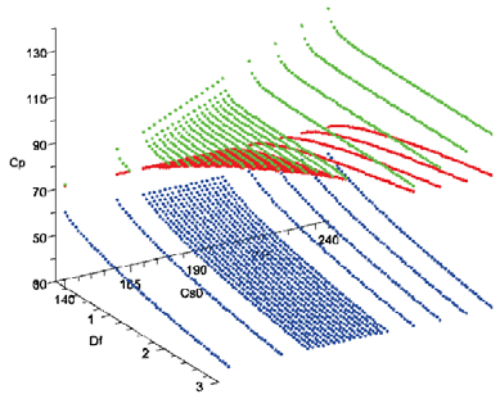


Fig. 1: Steady-state ethanol concentration (C_P) as a function of Dilution rate (Df) and inlet substrate concentration (C_{S0}).

2.2 Operating points

The optimal condition for the bioreactor is achieved with a high ethanol production. For the Jöbses's model, it can be shown that the main decision operating criterion for the best operating point is the ethanol production (EP), which is given by the multiplication of the dilution rate (D_f) and the ethanol concentration (C_P), i.e.,

$$EP = D_f.C_P. \quad (2)$$

Fig. 2 is produced for the iso-inlet substrate concentrations C_{S0} of 180, 200, and 220 kg/m^3 that were already shown in Fig. 1. Note that the dashed and dashdot lines are the same for all three operating conditions, therefore only one line for each branch is shown in Fig. 2.

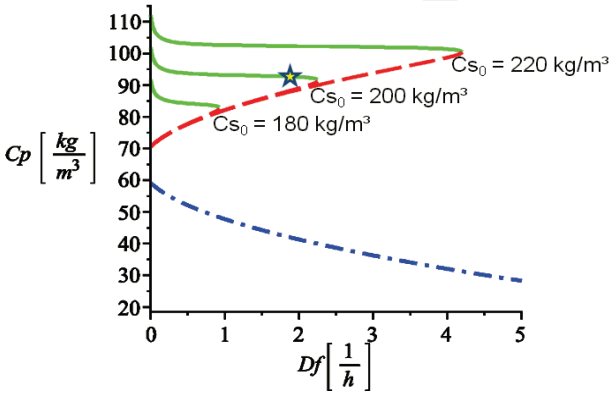


Fig. 2: Steady-state ethanol concentration (C_P) for three inlet substrate concentrations (C_{S0}).

Considering the curves generated for $C_{S0} = 200\ kg/m^3$ in Fig. 2, we can see that in the range of $0 < D_f < 2.25\ h^{-1}$ the system has three possible steady-state solutions. Two of them are stables (solid and dashdot lines) and one unstable (i.e., the middle branch illustrated by dashed line). The best operating point is located in the above curve close to the saddle point formatted by the intersection of the stable and unstable branches. The star in the above branch depicts a typical optimal operating point. From a control point of view the main difficult is to maintain the system working in this point avoiding a migration to the bellow operating point – the dash-dot line in Fig. 2. This migration can occur if the dilution rate is above the corresponding to the bifurcation saddle point or by a reduction to the inlet substrate concentration. Fig. 3 shows the case where the dilution rate is increased from 2 h^{-1} to 2.5 h^{-1} and then at 10 h again reduced to 2 h^{-1}. To bring the system to the more profitable operating point it is necessary to apply a pulse in the inlet ethanol concentration (C_{P0}) as shown at 15 h.

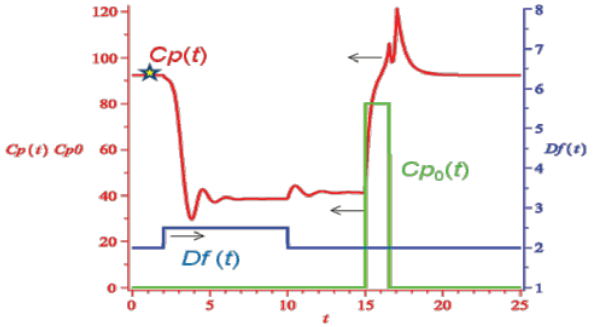


Fig. 3: Dynamic simulation to show how it is easy to move the operating point from the high to the low production branch.

2.3 Operating regions

The manifold that separate the region where occurs multiplicity in the ethanol concentration is defined by the saddle bifurcation points and is shown in Fig. 4. This manifold limits the operating region with possible high ethanol concentration (region A) from the region where only a low ethanol concentration is achievable (region B), where the operating conditions goes outside de operating region A, the ethanol concentration will fall down as it was depicted in Fig. 3.

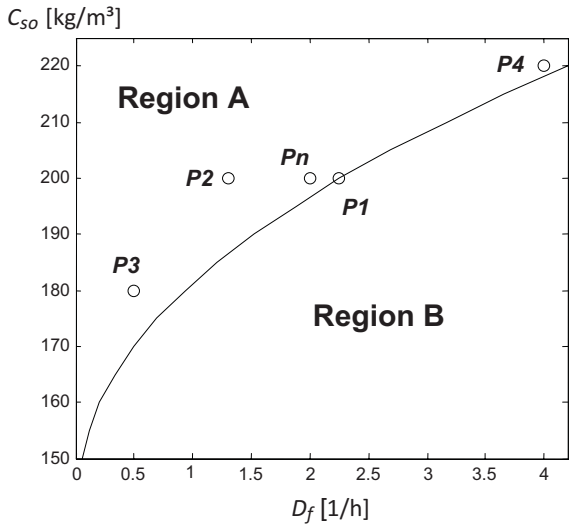


Fig. 4: Manifold of saddle points defining the operating region A (where exists multiplicity) and operating region B (where exits one solution only).

To characterize the differences between the high and low ethanol concentration branches, the system was linearized in five different operating points defined in Table 2 and placed in Fig. 4, where P_n is considered as the nominal/optimal operating point. The other four can occurs during a normal operation in the region A. Of course, P_4 has higher *EP* than P_n, but since the Jöbses model cannot describe higher C_{S0} correctly, we will just assume that P_n is the best operating point, but the same analysis could be performed considering P_4 as nominal model.

Table 2: Definition of Operating Region

	D_f $[h^{-1}]$	$C_{S0}\left[\frac{kg}{m^3}\right]$	OP1- high C_P $(C_P, C_S)_{SS}$*	OP2 - low C_P $(C_P, C_S)_{SS}$*
P_n	2.0	200	(92.57, 1.23)	(41.29, 111.34)
P_1	2.25	200	(91.83, 2.75)	(39.94,114,22)
P_2	1.3	200	(93.07, 0.40)	(45.56, 102.29)
P_3	0.5	180	(84.24, 0.31)	(52.09,68.90)
P_4	4.0	220	(101.38, 2.04)	(32.02, 151.16)

*$(C_P, C_S)_{SS}$ means steady-state values for ethanol and substrate concentration for a given D_f and C_{S0}.

Figures 5 and 6 respectively show the step response of the dynamic linearized models for high and low ethanol concentration branches. Essentially the high ethanol branch has an over-damped behavior, whereas for the low branch the system is under-damped.

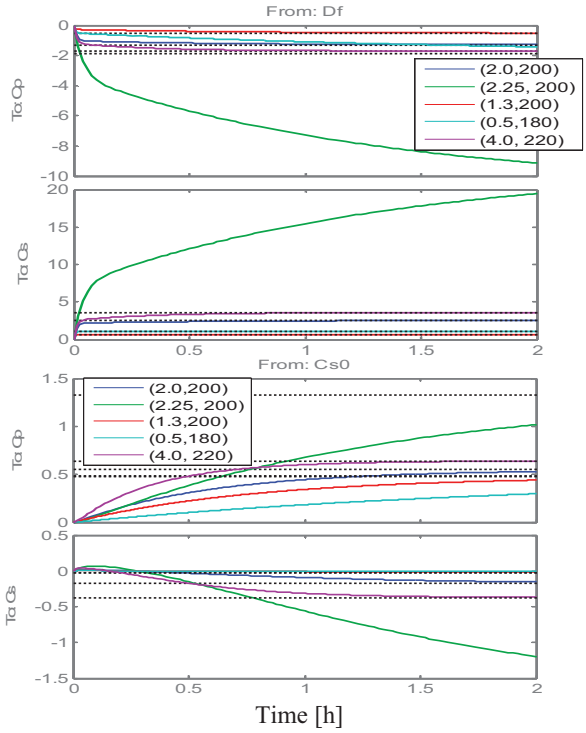


Fig. 5: Step response of the linearized models at the operating points defined in Table 2 corresponding to the high ethanol concentration branch for C_P and C_S outputs.

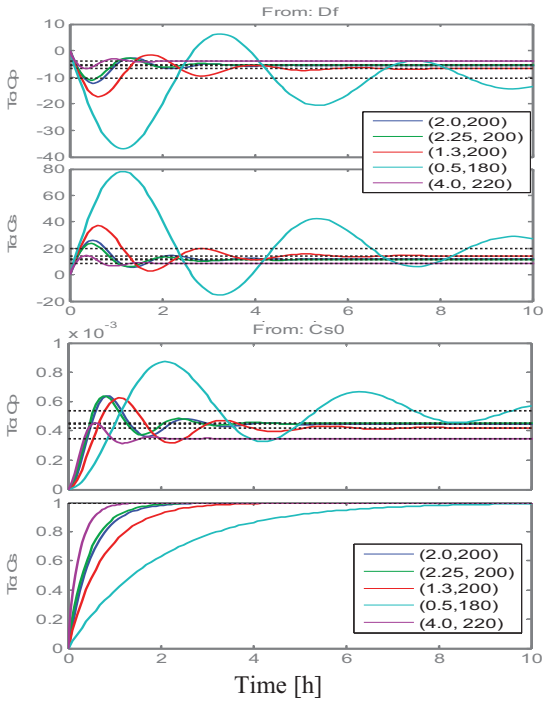


Fig. 6: Step response of the linearized models at the operating points defined in Table 2 corresponding to the low ethanol concentration branch for C_P and C_S outputs.

3 OPERATIONAL CONTROLLABILITY ANALYSIS

3.1 Manipulated and controlled variables

Ethanol and substrate concentrations are two natural controlled variables of the system and can be in principle measured on-line using 2D-fluorescence spectroscopy (Hantelmann *et al.*, 2006). As candidates for manipulated variables, we have the dilution rate (D_f), inlet substrate concentration (C_{S0}), and inlet ethanol concentration (C_{P0}). In principle, C_{P0} should be used only in critical situations, just to bring the system back to the higher production branch (as shown in Fig. 3). Therefore we will not consider it in our operational controllability analyses, where only the input-output pairs $(D_f, C_{S0}) \rightarrow (C_P, C_S)$ is further considered.

3.2 Nominal Operational Controllability

The determinant and the elements λ_{11} and λ_{12} of the RGA – relative gain array (Skogestad and Postlethwaite, 2005) – are calculated using the steady state gain matrix for each one of the linearized models defined in Table 2. These results are summarized in Table 3, where we can see that the determinant (Det) does not change its sign when the system goes from the operating region 1 (OR1) with high ethanol concentration to OR2 with low ethanol concentration. Nevertheless the recommended pairing using steady state RGA changes from OR1 to OR2. Usually, when the pairing recommendation is changed it is normally associated with a change in the determinant sign, what it does not happens for this system. When the determinant changes its sign it is equivalent to the change in the multivariable gain, what is quite critical for the success of any control strategy. The reason for this unusual behavior is related to the gain sign change of channel $C_{S0} \rightarrow C_S$ ($i.e., K_{22}$) as it can been seem by the gain matrix for the nominal operating point P_N for the high and low ethanol concentration branches, which are given by:

$$K_{PN\_high} = \begin{bmatrix} -1.30 & 0.55 \\ 2.55 & -0.18 \end{bmatrix} \quad (3)$$

$$K_{PN\_low} = \begin{bmatrix} -5.55 & 0.0004 \\ 11.82 & 0.999 \end{bmatrix} \quad (4)$$

Similar behavior occurs for all other OPs.

Table 3: Determinant and Steady State RGA

	Det OR1	Det OR2	RGA – OR1 $(\lambda_{11}, \lambda_{12})$*	RGA – OR2 $(\lambda_{11}, \lambda_{12})$*
P_n	-1.16	-5.55	(-0.20, 1.20)	(0.99, 0.01)
P_1	-10.58	-5.25	(-1.89, 2.89)	(0.99, 0.01)
P_2	-0.26	-6.78	(-0.05, 1.05)	(0.99, 0.01)
P_3	-0.48	-10.26	(-0.11, 1.11)	(0.99, 0.01)
P_4	-1.66	-3.98	(-0.38, 1.38)	(0.999, 0.001)

$(\lambda_{11}, \lambda_{12})$* it was calculated considering the pairing $D_f \rightarrow C_P$ and $C_{S0} \rightarrow C_S$.

3.3 RPN and rRPN Analysis

The Robust Performance Number (RPN) was introduced in (Trierweiler, 1997) and (Trierweiler and Engell, 1997) as a measure to characterize the operational controllability of a system. The RPN indicates how potentially difficult it is for a given system to achieve the desired performance robustly. The RPN is influenced by three terms: the desired closed loop performance, nonminimum phase behavior (i.e., RHP pole, zero, and pure time delays), and its degree of directionality.

The RPN is a measure of how potentially difficult it is for a given system to achieve the desired performance robustly. The easiest way to design a controller is to use the inverse of the process model. An inverse-based controller will have potentially good performance robustness only when the RPN is small. As inverse-based controllers are simple and effective, it can be concluded that a good control structure selection is one with a small (< 5) RPN (Trierweiler and Engell, 1997).

Table 4: RPN Analysis for the *high* ethanol concentration branch

	$t_R = 1.0\ h$		$t_R = 0.50\ h$		$t_R = 0.25\ h$	
	RPN	$rRPN$	RPN	$rRPN$	RPN	$rRPN$
P_n	1.66	0.088	1.61	0.085	1.58	0.086
P_1	2.27	0.152	1.90	0.1201	1.78	0.097
P_2	1.52	0.069	1.53	0.074	1.51	0.0790
P_3	1.50	0.068	1.53	0.0725	1.52	0.0781
P_4	1.91	0.124	1.80	0.110	1.69	0.101

Table 4 shows the RPN calculated by three different desired performance specified by the rise times $t_R = 1, 0.5,$ and $0.25\ h$ and 10% overshoot for both outputs, what makes the system 2, 4, and 8 times faster than open loop response for the high ethanol concentration branch and all operating points. The results shown in Table 4 allow us to conclude that considering each operating point independently they will be easily controllable. It is important to mention that the RPN does not give a clear idea of the control difficulties for that it is necessary to analyze the *relative* RPN ($rRPN$), which has been introduced by Trierweiler (2002) (see also (Trierweiler and Farina, 2002) for an additional discussion).

The $rRPN$ is the relative distance between the RPN curve and the minimum RPN curve and it is quantified by the areas under the curves. Values less than 1 and close to zero means that the desired performance is easily achievable. Again, since we are considering only the two stable branches, no nonminimum phase component occurs in the analyzed operating points. In this case, typically faster performance will usually reduce the $rRPN$. Table 4 depicts that it is possible to design a controller that achieves the desired performance for each one of the analyzed operating points. Similar analyses (not shown here) performed for the low ethanol concentration branch produce similar conclusions with a RPN and $rRPN$ in the order of 1.5 and 0.004, respect.

The local operational controllability analyses clearly conclude that for each one of the considered models it is possible and easy to design a controller with the desired performance. Although each point is easily controllable,

nothing can be said about all operating points together. Is there possible to design a controller that will produce a good performance for all operating points? Moreover, if a controller designed for OR1 will work in OR2? Based on the steps responses of the linearized models shown in Figures 5 and 6, it seems the responses are quite different, especially if we compare the under-damped behavior shown by OR2 and the over-damped of the OR1. In the next subsection, we answer these questions through the *nonlinear RPN* analysis.

3.4 Nonlinear degree – nRPN Analysis

In (Farenzena and Trierweiler, 2004) three novel indices were introduced to measure system's nonlinearity. These nonlinear measurements are derived from the Robust Performance Number (RPN) concept. The total system's nonlinearity can be measured by the *nonlinear RPN* ($nRPN$), while the purely static nonlinearity is captured by *nonlinear static RPN* ($nRPN_{STAT}$) and the dynamic component by the *nonlinear dynamic RPN* ($nRPN_{DYN}$). These indices do not require a nonlinear model, being enough a set of linear models. Therefore, they can easily be applied to quantify the nonlinearities of industrial plants and used to answer several practical important questions such as: how nonlinear is the system? Is it necessary to apply a nonlinear controller? What kind of nonlinear controller is required?

In the definition of the nonlinear RPN indices was introduced the logarithm function to make easier their interpretation. Values smaller than 1 indicate that the performance difference between nonlinear and linear controllers is not significant, so that a linear controller is recommended. Indices greater than 2 clearly indicate that a nonlinear controller is necessary. Between 1 and 2 is a transition zone, where in many times a robust controller can stabilize all possible plants, but the performance loss can be significant if the values are close to 2. This analysis is made for all three indices. For instance, if $nRPN$ and $nRPN_{STAT}$ are high and $nRPN_{DYN}$ is small, it indicates that the nonlinearity is essentially static and can be compensated by gain scheduling controller. If the all values are big (greater or close to 2), then a nonlinear model predictive controller is recommended.

Table 5: $nRPN$, $nRPN_{STAT}$, and $nRPN_{DYN}$. Analyses

	$t_R = 1.0\ h$			$t_R = 0.50\ h$		
	nRPN	*STAT.*	*DYN.*	*nRPN*	*STAT.*	*DYN.*
H	0.86	0.58	-0.052	0.79	0.58	-0.21
L	0.31	-0.28	0.455	0.27	-0.28	0.40
T	1.68	0.88	0.74	1.50	0.88	0.51
	$t_R = 0.25\ h$			$t_R = 0.10\ h$		
	nRPN	*STAT.*	*DYN.*	*nRPN*	*STAT.*	*DYN.*
H	0.74	0.58	-0.35	0.69	0.58	-0.54
L	0.27	-0.28	0.39	0.29	-0.28	0.42
T	1.32	0.88	0.24	1.12	0.88	-0.12

→ Stat. and Dyn. mean $nRPN_{STAT}$ and $nRPN_{DYN}$, respect. *H*: polytope model for the high C_P branch, *L*: polytope model for the low C_P branch, and $T = H \cup L$.

To quantify the nonlinearity using the *nonlinear RPN* it is necessary to construct a set of linearized models, called as polytope model. The operating points defined in Table 2 have been used to define the polytopes. Three polytope sets have being formed: *H* formed with the 5 linearized models at the *high* ethanol concentration branch; *L* formed from the linearized models corresponding to the *low* C_P branch; and *T* formed by the union of all 10 models.

Table 5 summarizes the results of the $nRPN$ analysis, which indicates that in general the polytope *H* is more nonlinear than the polytope *L* and the nonlinearity found in *H* is essentially static, whereas for *L* the nonlinearity is of the dynamic type. Of course, the combination of both polytopes (*T*) exhibits the highest nonlinearity, which has similar dynamic and static components for low performance, but for the highest desired performance (i. e., $t_R = 0.10$) it becomes essentially static. Moreover, for this performance it is expected that a linear controller will be able to control the system in both operating regions. To verify this prediction, we have design a multivariable PI controller, given by:

$$PI(s) = \begin{bmatrix} -0.161 \times \left(1 + \frac{1}{0.1128s}\right) & 1.6981 \times \left(1 + \frac{1}{0.0063s}\right) \\ 12.287 \times \left(1 + \frac{1}{0.6288s}\right) & 6.4204 \times \left(1 + \frac{1}{0.3606s}\right) \end{bmatrix} \quad (5)$$

This quite simple controller can control all 10 linearized models with a good performance as it is shown in Fig. 7, where it has been simulated a setpoint change in C_P of one unit at 1 h and a simultaneous load disturbance of $\Delta D_f = 0.2\ h^{-1}$ and $\Delta C_{S0} = 1\ kg/m^3$ at 6 h.

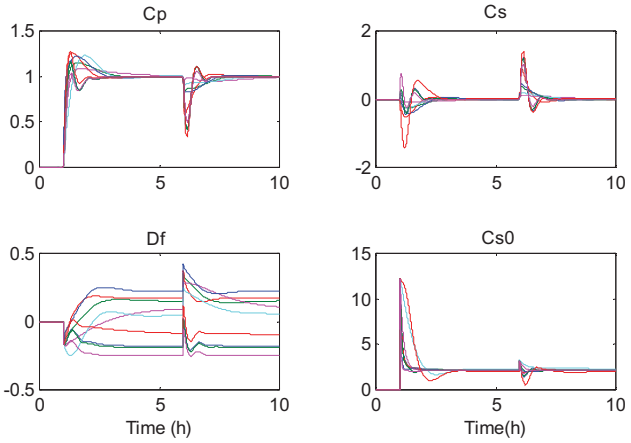


Fig. 7: Closed loop simulation with a multivariable *PI* controller for all 10 linearized models

4 CONTROL STRATEGY

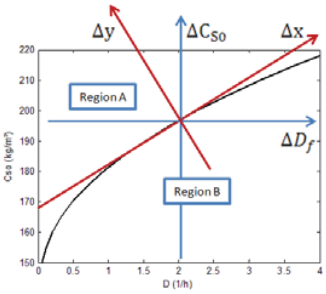
Note that the predictions made by nonlinear RPN indices were confirmed by the closed loop simulation using the linearized models. Similar results are obtained using the fully nonlinear model (Trierweiler and Diehl, 2009). In this section, it is shown the basic ideas of the recommended control strategy.

4.1 General control strategy

It is recommended to use an Extended Kalman Filter (EKF) technique to filter the measurements from the 2D-fluorescence spectroscopy. The EKF is important also to estimate the biomass concentration and to take the input and model uncertainties into account. A multivariable controller should be used with the pairing $(D_f, C_{S0}) \rightarrow (C_P, C_S)$ and the inlet ethanol concentration (C_{P0}) should be only used in exceptional cases to bring the system back to the high ethanol concentration branch as shown in Fig. 3.

4.2 Solving the constraint problem

A single linear controller can be used to control the bioreactor in all operating conditions. The only special problem is to constraint the range of the manipulated variables, which should be limited by the saddle point manifold. This can be easily guaranteed by a single variable transformation as shown in Fig. 8. Instead to use directly the physical variables $(\Delta C_{S0},\ \Delta D_f)$ the controller calculate the control actions for $(\Delta y,\ \Delta x)$ using the simple restriction $\Delta y > 0$. The conversion to $(\Delta C_{S0},\ \Delta D_f)$ is performed by the multiplication with the rotation matrix R, given by:



$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

where θ is the rotation angle.

Fig. 8: Simple variable transformation for guarantee the feasible region A

Note that the optimal operating point is close to $\Delta y = 0$. This idea is explored in (Trierweiler and Diehl, 2009) to maintain the system always close to the optimal economic conditions.

5 CONCLUSIONS

The paper has introduced several contributions, which are summarized as follows:

***Zymomonas mobilis* control problem** – as far as we know, this contribution is the first paper that discusses the control problem of a continuous bioreactor with *Z. mobilis*.

Control strategy – a simple and efficient control strategy has been proposed based on the control of (C_P, C_S) using the manipulated variables $(\Delta C_{S0},\ \Delta D_f)$. Since the branch where the high C_P occurs is restricted, it was proposed to use a variable transformation as shown in Fig. 8. This transformation makes simple to consider the restriction range and open a simple strategy for assuring the optimal economic conditions, which occurs at $\Delta y \approx 0.5$.

RGA pairing with negative diagonal elements – the *Z. mobilis* bioreactor with the Jöbses's model it is a interesting example to illustrate that the common pairing rule where the positive diagonal elements are recommended it is not always the best global pairing. For this system the best pairing changes when the system moves from the high to the low C_P branch. It is interesting that the paring change occurs without alteration in the sign of the gain matrix determinant.

The RPN analysis is reliable – the RPN analysis, especially using the nonlinear RPN indices made possible very easily to check and quantify the nonlinearity degree and based on these analyses prescribe the appropriated controller. It was shown that if a fast controller is designed a simple linear controller can be used. To confirm this prediction a simple multivariable PI controller has been tuned and simulated.

Procedure for analysis a bioprocess optimization and control problem – finally the paper illustrate the typical steps necessary to develop a control and optimization strategy for a bioprocess system. Before designing a controller, it is necessary to systematically analyze the system as have been done in the paper.

Jöbses's Model – we have assumed that the model could predict the system behavior at high dilution rate and inlet substrate concentration. Of course, this extrapolation is totally questionable considering a real application.

Acknowledgements: The authors are very grateful for the grants from CAPES / BRAZIL. The first author also thanks Prof. Dr. W. Marquardt for receiving him for his sabbatical at RWTH Aachen, where this work has been written.

REFERENCES

Daugulis, A. J., Mclellan, P. J. and Li, J. (1997) Experimental investigation and modeling of oscillatory behavior in the continuous culture of *Zymomonas mobilis*. *Biotechnology and Bioengineering,* 56**,** 99-105.

Elnashaie, S. S. E. H., Chen, Z., Garhyan, P., Prasad, P. and Mahecha-Botero, A. (2006) Practical Implications of Bifurcation and Chaos in Chemical and Biological Reaction Engineering. *International Journal of Chemical Reactor Engineering 4*

Farenzena, M. and Trierweiler, J. O. (2004) System Nonlinearity Measurement Based on the RPN Concept. *DYCOPS - 7TH IFAC Symp. on Dynamics and Control of Process Systems***,** 181-191.

Hantelmann, K., Kollecker, M., Hüll, D., Hitzmann, B. and Scheper, T. (2006) Two-dimensional fluorescence spectroscopy: A novel approach for controlling fed-batch cultivations. *Journal of Biotechnology,* 121**,** 410-417.

Jöbses, I. M. L., Egberts, G. T. C., Luyben, K. C. A. M. and Roels, J. A. (1986) Fermentation kinetics of *Zymomonas mobilis* at high ethanol concentrations: Oscillations in continuous cultures. *Biotechnology and Bioengineering,* 28**,** 868-877.

Rogers, P., Jeon, Y., Lee, K. and Lawford, H. (2007) Zymomonas mobilis for Fuel Ethanol and Higher Value Products. *Biofuels.* Springer Berlin / Heidelberg.

Skogestad, S. and Postlethwaite, I. (2005) *Multivariable feedback control : analysis and design,* Chichester, England ; Hoboken, NJ, John Wiley.

Trierweiler, J. O. (1997) A Systematic Approach to Control Structure Design. Dortmund, University of Dortmund.

Trierweiler, J. O. (2002) Application of the RPN Methodology for Quantification of the Operability of the Quadruple-Tank Process. *Brazilian Journal of Chemical Engineering,* 19**,** 195-206.

Trierweiler, J. O. and Diehl, F. C. (2009) Control Strategy for a *Zymomonas mobilis* Bioreactor used in the Ethanol Production. *Submitted to PSE 2009*

Trierweiler, J. O. and Engell, S. (1997) The Robust Performance Number: A New Tool for Control Structure Design. *Computer Chemical Engineering* 21**,** 237-243.

Trierweiler, J. O. and Farina, L. A. (2002) The RPN Methodology Applied to Quantify Process Operability and Controller Design. *Proc.15th IFAC World Congress, Barcelona***,** CDROM, paper 2663.

Zhang, Y. and Henson, M. A. (2001) Bifurcation Analysis of Continuous Biochemical Reactor Models. *Biotechnology Progress,* 17**,** 647-660.

Adaptive extremum-seeking control of fed-batch cultures of micro-organisms exhibiting overflow metabolism

L. Dewasme * **A. Vande Wouwer** * **B. Srinivasan** ** **M. Perrier** **

* *Service d'Automatique, Faculté Polytechnique de Mons, Boulevard Dolez 31, B-7000 Mons, Belgium (e-mails : Laurent.Dewasme, Alain.VandeWouwer@fpms.ac.be)*
** *Département de génie chimique, Ecole Polytechnique de Montréal, C.P.6079, Montréal, Que., Canada*

Abstract: *Overflow metabolism* characterizes cells strains that are likely to produce inhibiting metabolites resulting from an excess of substrate feeding and a saturated respiratory capacity. The critical substrate level separating the two different metabolic pathways is generally not well defined. This paper proposes two non-model based extremum-seeking strategies preventing a too important accumulation of inhibiting metabolites in fed-batch cultures, by estimating the critical substrate level on the basis of 3 simple measurements related to the feeding, oxygen and carbon dioxide. A simple substrate controller based on Lyapunov stability arguments is then designed and tested in combination with the two extremum-seeking schemes.

Keywords: Extremum seeking, nonlinear adaptive control, fermentation process, biotechnology.

1. INTRODUCTION

Industrial vaccine production is usually achieved using fed-batch cultures of genetically modified yeast or bacteria strains, which can express different kinds of recombinant proteins. From an operational point of view, it is necessary to determine an optimal feeding strategy (i.e. the time evolution of the input flow rate to the fed-batch culture) in order to maximize productivity.

The main encountered problem comes from the metabolic changes of such strains in presence of feeding overflow. This "overflow metabolism", also called "short-term Crabtree effect", is a metabolic phenomenon that is induced when the rate of glycolysis exceeds a critical value, leading to a generally inhibiting by-product formation from pyruvate (for not well understood reasons). It occurs for instance in *S. cerevisiae* cultures with aerobic ethanol formation, in *P. pastoris* with aerobic methanol formation, in *E. coli* cultures with aerobic acetate formation or in mammalian cell cultures with the aerobic lactate formation. To avoid this undesirable effect, a closed-loop optimizing strategy is required, which could take various forms (Pomerleau (1990), Chen et al. (1995), Akesson (1999), Renard (2006), Dewasme et al. (2007)).

In this study, a non-model based extremum-seeking strategy is chosen. Two original techniques are proposed and compared.

The first one is related to the work of Blackman in the 60's, revisited and improved in Ariyur and Krstic (2003) while the second one is based on a simple recursive least squares technique (RLS). Non-model based extremum-seeking has already been applied succesfully to dynamic optimization of continuous cultures in Wang et al. (1999).

Alternatively, model-based extremum-seeking strategy as presented in the works of Guay et al. (2004), Titica et al. (2003a) and Titica et al. (2003b) could also be considered for the on-line determination of the critical glucose concentration. However, the convergence of this adaptation scheme is slow and lacks robustness (Dewasme and Vande Wouwer (2008)).

2. MODEL AND CONTROL OBJECTIVES

2.1 Modeling cultures of micro-organisms exhibiting overflow metabolism

In this study, we consider a generic model that would, in principle, allow the representation of the culture of different strains presenting an overflow metabolism (yeasts, bacteria, animal cells, etc). This model describes therefore the cell catabolism through the following three main reactions:

$$\text{Substrate oxidation :} \quad \mathrm{S} + k_5\,\mathrm{O} \overset{r_1 X}{\rightarrow} k_1\,\mathrm{X} + k_8\,\mathrm{C} \tag{1a}$$

Overflow reaction (typically fermentation) :

$$\mathrm{S} + k_6\,\mathrm{O} \overset{r_2 X}{\rightarrow} k_2\,\mathrm{X} + k_4\,\mathrm{P} + k_9\,\mathrm{C} \tag{1b}$$

$$\text{Metabolite product oxidation :} \quad \mathrm{P} + k_7\,\mathrm{O} \overset{r_3 X}{\rightarrow} k_3\,\mathrm{X} + k_{10}\,\mathrm{C} \tag{1c}$$

where X, S, P, O and C are, respectively, the concentration in the culture medium of biomass, substrate (typically glucose or glycerol), product (i.e. ethanol or methanol in yeast cultures, acetate in bacteria cultures or lactate in animal cells cultures), dissolved oxygen and carbon dioxide. k_i are the yield coefficients and r_1, r_2 and r_3 are the nonlinear specific growth rates given by:

$$r_1 = \min\left(r_S, r_{Scrit}\right) \tag{2}$$

$$r_2 = \max\left(0, r_S - r_{Scrit}\right) \tag{3}$$

$$r_3 = \max\left(0, \min\left(r_P, \frac{k_5(r_{Scrit} - r_S)}{k_7}\right)\right) \tag{4}$$

where the kinetic terms associated with the substrate consumption r_S, the critical substrate consumption r_{Scrit} (generally de-

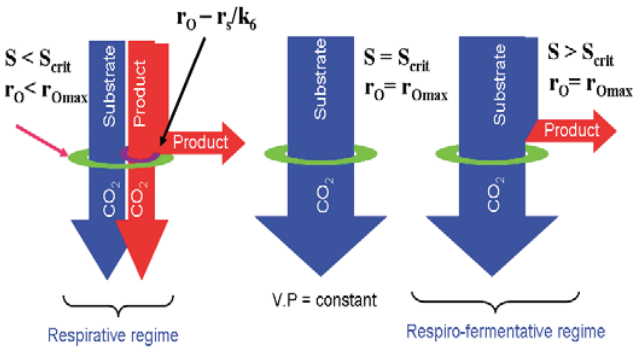


Fig. 1. Illustration of Sonnleitner's bottleneck assumption for cells limited respiratory capacity.

pendant on the cells oxidative or respiratory capacity r_O) and the product oxidative rate r_P are given by:

$$r_S = \mu_S \frac{\mathrm{S}}{\mathrm{S}+K_S} \tag{5a}$$

$$r_{Scrit} = \frac{r_O}{k_5} = \frac{\mu_O}{k_5}\frac{\mathrm{O}}{\mathrm{O}+K_O}\frac{Ki_P}{Ki_P+P} \tag{5b}$$

$$r_P = \mu_P \frac{\mathrm{P}}{\mathrm{P}+K_P} \tag{5c}$$

These expressions take the classical form of Monod laws where μ_S, μ_O and μ_P are the maximal values of specific growth rates, K_S, K_O and K_P are the saturation constants of the corresponding element, and Ki_P is the inhibition constant.

This kinetic model is based on Sonnleitner's bottleneck assumption (Sonnleitner and Käppeli (1986)) which was applied to a yeast strain *Saccharomyces cerevisiae* (Figure 1). During a culture, the cells are likely to change their metabolism because of their limited respiratory capacity. When the substrate is in excess (concentration $S > S_{crit}$), the cells produce a metabolite product P through fermentation, and the culture is said in respiro-fermentative (RF) regime. On the other hand, when the substrate becomes limiting (concentration $S < S_{crit}$), the available substrate (typically glucose), and possibly the metabolite P (as a substitute carbon source), if present in the culture medium, are oxidized. The culture is then said in respirative (R) regime.

Component-wise mass balances give the following differential equations :

$$\frac{dX}{dt} = (k_1 r_1 + k_2 r_2 + k_3 r_3)X - DX \tag{6a}$$

$$\frac{dS}{dt} = -(r_1 + r_2)X + DS_{in} - DS \tag{6b}$$

$$\frac{dP}{dt} = (k_4 r_2 - r_3)X - DP \tag{6c}$$

$$\frac{dO}{dt} = -(k_5 r_1 + k_6 r_2 + k_7 r_3)X - DO + OTR \tag{6d}$$

$$\frac{dC}{dt} = (k_8 r_1 + k_9 r_2 + k_{10} r_3)X - DC - CTR \tag{6e}$$

$$\frac{dV}{dt} = F_{in} \tag{6f}$$

where S_{in} is the substrate concentration in the feed, F_{in} is the inlet feed rate, V is the culture medium volume and D is the dilution rate ($D = F_{in}/V$). OTR and CTR represent respectively the oxygen transfer rate from the gas phase to the liquid phase and the carbon transfer rate from the liquid phase to the gas phase. Classical models of OTR and CTR are given by:

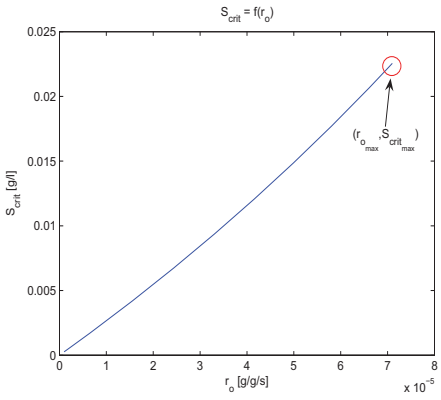


Fig. 2. S_{crit} as a function of r_o.

$$OTR = k_L a(O_{sat} - O) \tag{7a}$$

$$CTR = k_L a(P - P_{sat}) \tag{7b}$$

where $k_L a$ is the volumetric transfer coefficient and, O_{sat} and P_{sat} are respectively the dissolved oxygen and carbon dioxide concentrations at saturation.

2.2 Control objectives

First, we show that the respiratory capacity has an influence on the critical substrate concentration level. In the optimal operating conditions ($S = S_{crit}$), the fermentation and metabolite product oxidation rates are equal to zero and the substrate consumption rate r_S is equal to r_{Scrit} or $\frac{r_O}{k_5}$. Consequently, after a trivial mathematical manipulation of (5a), a relation between the critical substrate concentration level and the cell respiratory capacity is obtained as:

$$S_{crit} = \frac{K_S r_O}{k_5 \mu_S - r_O} \tag{8}$$

Figure 2 shows a plot of this relation where the point $[0,0]$ corresponds to a totally inhibited respiratory capacity, preventing any growth, and the point $[r_{O_{max}}, S_{crit_{max}}]$ corresponds to maximum productivity (i.e. absence of metabolite product in the culture medium and a sufficient level of oxygenation). Obviously, the presence of the product in the culture medium can decrease the respiratory capacity and in turn the value of the critical substrate concentration $S = S_{crit}$. In order to maintain the system at the edge between the respirative and respiro-fermentative regimes, it would be necessary to determine online the critical substrate concentration (S_{crit}) and to control the substrate concentration in the culture medium around this value (Dewasme and Vande Wouwer (2008)). Unfortunately, the substrate concentration measurement is a difficult task as typical concentration levels are below the resolution of currently available probes (or sensors).

An alternative solution is to reformulate the problem not as a maximazition of the respiratory capacity but as the maximization of the substrate consumption rate coupled to the minimization of the fermentation rate.

This can finally be formulated as follows:

$$max_{S_{crit}} Y = max_{S_{crit}} \varphi_1 - \varphi_2 \tag{9}$$

where:

- Y is the assumed measurable cost function;
- φ_1 and φ_2 correspond to the reaction rates $r_1 X$ and $r_2 X$, respectively.

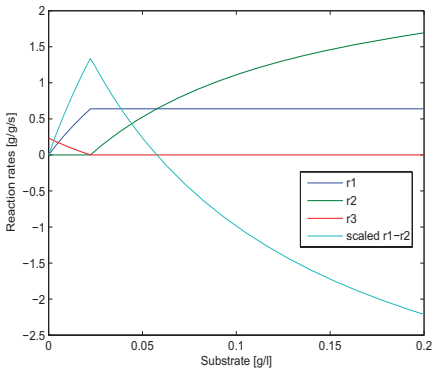


Fig. 3. Reaction rates and optimization criteria as a function of S.

In order to estimate the cost function Y online, we use a pseudo-steady state assumption. Indeed, assuming that the variations of substrate, oxygen and carbon dioxyde concentrations are equal to zero, we obtain from (6b), (6d) and (6e):

$$D(S_{in}-S) = (r_1+r_2)X \tag{10a}$$
$$-DO+OTR = (k_5r_1+k_6r_2+k_7r_3)X \tag{10b}$$
$$DC+CTR = (k_8r_1+k_9r_2+k_{10}r_3)X \tag{10c}$$

Dilution terms can be considered as negligible in comparison with OTR, CTR and DS_{in}. Replacing the reaction rates r_iX by φ_i $(i=1,2,3)$, (10) can be written:

$$DS_{in} = \varphi_1+\varphi_2 \tag{11a}$$
$$OTR = k_5\varphi_1+k_6\varphi_2+k_7\varphi_3 \tag{11b}$$
$$CTR = k_8\varphi_1+k_9\varphi_2+k_{10}\varphi_3 \tag{11c}$$

From this on, after some basic mathematical manipulations, it is possible to express a relation evolving proportionaly to $\varphi_1-\varphi_2$, as a function of the yield coefficients, OTR, CTR and DS_{in}. We decide to call DS_{in} the "substrate intake rate" (SIR) and we finally obtain:

$$Y = \varphi_1-\varphi_2 \propto y$$
$$y = 2\,k_{10}\,OTR - 2\,k_7\,CTR + (k_7k_9-k_5k_{10}+k_8k_7-k_6k_{10})SIR \tag{12}$$

This optimization criteria can thus be evaluated on the basis of 3 measurements (OTR, CTR and SIR) coupled to a sufficiently good identification of several yield coefficients. Figure 3 shows the evolution of the reaction rates and the criteria (scaled 20 times higher) as a function of the substrate concentration for a model of *S. cerevisiae* where the respiratory capacity is assumed to be constant (no oxygen limitation and no inhibition).

3. ADAPTIVE MODEL-FREE EXTREMUM-SEEKING STRATEGIES

Two adaptive extremum-seeking techniques are proposed in the following.

3.1 Extremum-seeking through a bank of filters

The objective of the extremum-seeking strategy is to determine on-line the parameter $\hat{\theta}$ (which in this case represents the critical glucose concentration). To this end, the system is excited by injection of a relatively slow sinusoidal dither signal $d = Asin(\omega t)$, as shown in Figure 4.

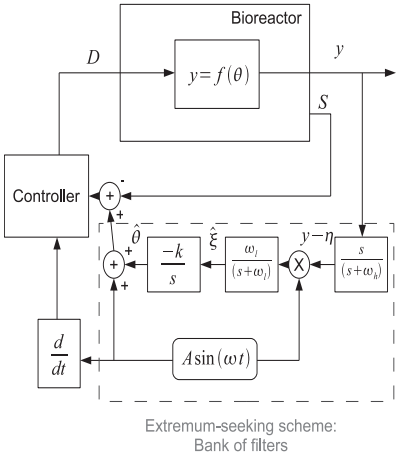


Fig. 4. Extremum-seeking scheme with a bank of filters.

The following equations describe the extremum-seeking method: The corresponding equations to Figure 4 are:

$$y = f(\hat{\theta}+Asin(\omega t)) \tag{13a}$$
$$\dot{\hat{\theta}} = k\xi \tag{13b}$$
$$\dot{\xi} = -\omega_l\xi+\omega_l(y-\eta)Asin(\omega t) \tag{13c}$$
$$\dot{\eta} = -\omega_h\eta+\omega_h y \tag{13d}$$

where:

- $y = f(\hat{\theta}+Asin(\omega t))$ is the measurable cost function;
- $\hat{\theta}$ is the estimation of the unknown parameter;
- k is the gain of the integrator;
- ξ can be seen as the gradient estimation ($\approx \frac{d\hat{\theta}}{dt}$);
- ω_l is the cut-off frequency of the low-pass filter;
- ω_h is the cut-off frequency of the high-pass filter;
- η is an intermediate variable explaining the absence of the low frequencies rejected from y in $y-\eta$ by the high-pass filter;

A first high-pass filter is used in order to reject the continuous component of y. The output is then multiplied by the dither signal in order to be "demodulated". As the dither signal and the output of the high-pass filter can only be in phase ($\hat{\theta}<\theta^*$) or out of phase ($\hat{\theta}>\theta^*$), there exists another continuous component inside the result of this demodulation. The second low-pass filter is used in order to isolate this new component containing the information of interest and sometimes residual mid-frequencies signals. This signal ξ is then filtered one last time by an integrator in order to attenuate the last "parasite" components and to recover the estimation of the unknown parameter from the integration of the continuous component. Following the theorem demonstrated in Krstic and Wang (1997), by choosing adequate values for all the parameters of the optimizing loop, the system should exponentially converge to an $O(\omega+A)$-neighborhood of the optimum.

3.2 Extremum-seeking through a RLS scheme

This second technique presents a scheme somewhat equivalent to the previous one where the bank of filters is actually replaced by a continuous recursive least squares (RLS) (Sastry and Bodson (1989)) scheme (see Figure 5) that computes the gradient ξ using a linear relationship, which is inspired from the shape of r_1-r_2 as a function of the substrate concentration:

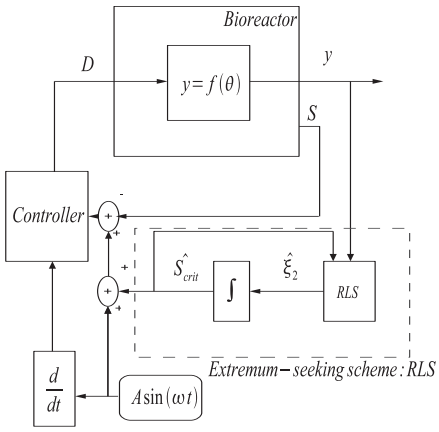


Fig. 5. Extremum-seeking scheme with RLS.

$$y = \hat{\xi}\Phi \tag{14}$$

where:

- $y = \varphi_1 - \varphi_2$
- $\hat{\xi} = [\hat{\xi}_1\ \hat{\xi}_2]$
- $\Phi = [1\ S_{crit}]$

In comparison with the previous extremum-seeking technique, (14) can be seen as the new relation replacing (12). The vector parameter $\hat{\xi}$ is then identified through the continuous RLS scheme that follows:

$$e = y - \hat{\xi}\Phi \tag{15a}$$

$$\dot{\hat{\xi}} = KR^{-1}\Phi^T e \tag{15b}$$

$$\dot{R} = K(\Phi^T\Phi - \lambda R) \tag{15c}$$

where:

- K is the strictly positive and constant adaptation gain;
- R is the inversed covariance matrix acting as a directional adaptation gain;
- λ is a forgetting factor used in order to avoid a "covariance wind-up problem" due to the absence of bounds in R growth (if $\lambda = 0$, $\dot{R} \geq 0$ (Sastry and Bodson (1989))).

$\hat{\xi}_2$ can be considered as the gradient estimation. This one is pushed to zero in average using an integral control of the form:

$$\dot{\hat{S}}_{crit} = k_i\hat{\xi}_2 \tag{16}$$

The conclusions about the convergence error are identical to the previous extremum-seeking technique.

3.3 Controller design

We derive adaptation and control laws from the consideration of a candidate Lyapunov function ensuring system stability.
First, equation (6b) can be rewritten as follows:

$$\frac{dS}{dt} = -\nu X - D(S - S_{in}) \tag{17}$$

where $\nu = r_1 + r_2$ is considered as an unknown kinetic parameter. Defining:

$$Z_s = S_{crit} + d - S \tag{18}$$

the control error variable, where $d = Asin(\omega t)$, is the periodical "dither signal".

$$\tilde{\nu} = \nu - \hat{\nu} \tag{19}$$

the estimation error on ν, we consider the following Lyapunov candidate function:

$$V = \frac{1}{2}Z_s^2 + \frac{1}{2\gamma}\tilde{\nu}^2 \tag{20}$$

where γ is a strictly positive tuning parameter.
A stabilizing controller is obtained if one can prove the strict negativity of the Lyapunov function derivative. Differentiating V and considering S_{crit} constant in order to decouple the control law from the extremum-seeking scheme (this can be done assuming that the controller converges significatively faster than the extremum-seeking scheme), we obtain:

$$\dot{V} = Z_s\left[\nu X + D(S - S_{in}) + \dot{d}\right] + \tilde{\nu}(-\frac{\dot{\hat{\nu}}}{\gamma}) \tag{21}$$

Replacing (18) and (19) in (21) and forcing $\dot{V}$ to be negative as in:

$$\dot{V} = -k_p Z_s^2 \tag{22}$$

where k_p is a strictly positive tuning parameter, we obtain:

$$-k_p Z_s = \hat{\nu}X + D(S - S_{in}) + \dot{d} \tag{23}$$

provided that:

$$\dot{\hat{\nu}} = -\gamma Z_s X \tag{24}$$

Finally, the control law is given by:

$$D = \frac{\left[k_p Z_s + \dot{d} + \hat{\nu}X\right]}{S_{in} - S} \tag{25}$$

This last expression explains the presence of the derivative $\dot{d}$ in the controller (Figure 4 and 5).

4. SIMULATION RESULTS

Coupling the controller designed in the subsection 3.3 with the extremum-seeking schemes, we apply the complete loop to a small-scale simulated yeasts culture (typically 20 l bioreactor). The initial and operating conditions are: $X_0 = 0.4g/l$, $S_0 = 0.5g/l, E_0 = 1g/l, O_0 = O_{sat} = 0.035g/l, C_0 = C_{sat} = 1.286g/l$, $V_0 = 5l$, $S_{in} = 350g/l$ where E_0 is the initial concentration of ethanol. For the kinetic and yield parameter values, the reader is referred to Sonnleitner and Käppeli (1986).

4.1 Application of the bank filters technique

The parameters for this extremum-seeking scheme are $A = 0.007$, $\omega = \frac{2\pi}{0.2}\ h^{-1}$, $\omega_h = 0.1\omega\ h^{-1}$, $\omega_l = 1.5\omega\ h^{-1}$, $k = 100$ and $k_p = 100$. The culture time is fixed to 24 h. Figures 6 and 7 show the results when no inhibition from ethanol accumulation is considered. This seems to be realistic as the ethanol concentration is below 4 g/l.
However, inhibition is an important phenomena that has to be taken into account. When included in our bioprocess model, the extremum-seeking results are as shown in Figure 8 and 9. It is apparent that the biomass level that can be achieved is significantly affected by the presence of ethanol, despite the set-point adaptation. Note that these results are very satisfactory in view of the situation where a constant substrate concentration is regulated. Indeed, a small error on S_{crit} would lead to a dramatical accumulation or reconsumption of ethanol and biomass growth would probably be affected beyond model prediction.
As it is explained in Ariyur and Krstic (2003), the output error of the extremum-seeking algorithm achieves a local exponential convergence to an $O(A^2)$-neighborhood of the origin if it is

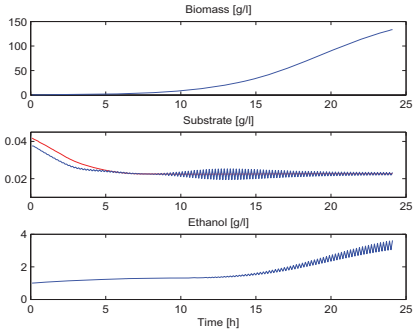


Fig. 6. Biomass (X), substrate (S in blue and $\hat{S}_{crit}$ in red), and ethanol (E) concentrations evolutions when no respiratory capacity inhibition is considered.

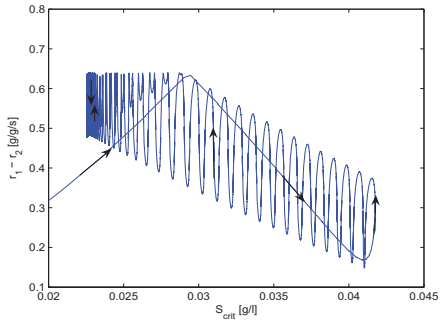


Fig. 7. Convergence of the optimization criteria $r_1 - r_2$ to the optimum when no respiratory capacity inhibition is considered.

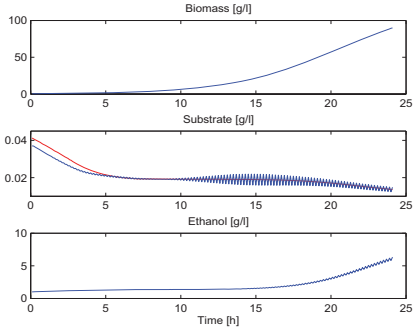


Fig. 8. Biomass (X), substrate (S in blue and $\hat{S}_{crit}$ in red), and ethanol (E) concentrations evolutions when inhibition is considered.

assumed that we are operating around a point of zero slope as it is typically the case for a convex function. As it can be observed in Figure 3, the criteria does not present a point of zero slope as the function has a discontinuous derivative at the optimum. Despite this difficulty, we see that the algorithm converges well and more particularly, the error is around an $O(A)$-neighborhood of the origin (Krstic and Wang (1997)).
This last remark, which won't be elaborated in this paper, is clearly a source of bias in the set-point when the ethanol inhibition is considered (cfr Figure 9). As A needs to be chosen sufficiently large to create a significant variation on the system dynamics, a small error cannot be avoided and the ethanol is

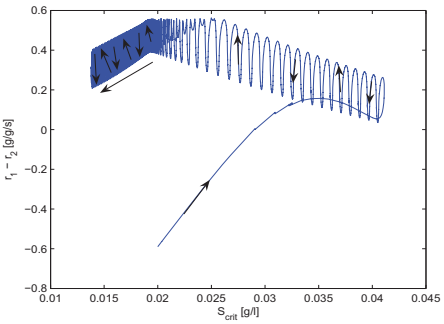


Fig. 9. Convergence of the optimization criteria $r_1 - r_2$ to the optimum when inhibition is considered.

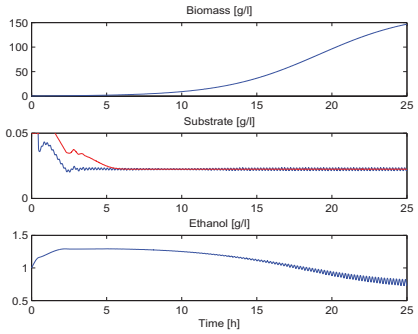


Fig. 10. Extremum-seeking with RLS: Biomass (X), substrate (S in blue and $\hat{S}_{crit}$ in red), and ethanol (E) concentrations evolutions when no respiratory capacity inhibition is considered.

accumulated while the algorithm goes on converging. As the ethanol concentration grows, the respiratory capacity slightly decreases and, following (8), S_{crit} does so.

4.2 Application of the RLS technique

The tuning parameters are defined as: $A = 0.001$, $K = 100$, $\lambda = 0.1$, $\omega = \frac{2\pi}{0.2}$, $k_i = 0.01$ and $k_p = 100$. The culture time is still fixed to 24 h. Figures 10 and 11 show the new results when no inhibition from ethanol accumulation is considered, and Figures 12 and 13 when the inhibition term is taken into account. The main observations are: (i) convergence is clearly faster. (ii) convergence is achieved around S_{crit} so that the ethanol concentration slowly decreases in the last hours.
When no inhibition is considered as in Figure 10, this set-point bias has no consequence on the extremum while, in Figure 12, when inhibition is taken into account, the set-point error is, by chance, playing a positive role so that ethanol is consumed.
In this application, the RLS algorithm is less computationally demanding, and easier to tune than the bank of filters strategy.

5. CONCLUSION

The high productivity of fed-batch cultures using genetically modified strains exhibiting overflow metabolism relies on a double condition: an optimal feeding strategy and the implied limitation of the inhibiting by-product formation. To this end, an adaptive controller using two different non-model based extremum-seeking strategies is designed for a general case of

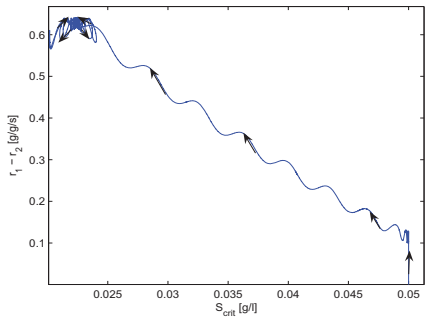


Fig. 11. Extremum-seeking with RLS: convergence of the optimization criteria $r_1 - r_2$ to the optimum when no respiratory capacity inhibition is considered.

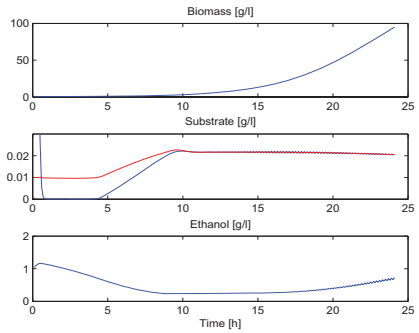


Fig. 12. Extremum-seeking with RLS: Biomass (X), substrate (S in blue and $\hat{S}_{crit}$ in red), and ethanol (E) concentrations evolutions when inhibition is considered.

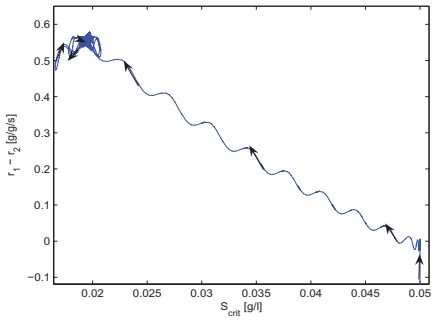


Fig. 13. Extremum-seeking with RLS: convergence of the optimization criteria $r_1 - r_2$ to the optimum when inhibition is considered.

overflow metabolized strain and is applied to the particular case of *S. cerevisiae*. The tracking of the critical substrate level (or, at least, its kinetic image), representing the optimum, is correctly performed by both extremum-seeking techniques, limiting the ethanol accumulation despite the considerations of an ethanol-inhibited respiratory capacity and discontinuous derivatives around the optimum.

6. ACKNOWLEDGMENT

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.
The authors also gratefully acknowledge the support of the Wallonie-Bruxelles-Québec commission in the framework of the research project between A. Vande Wouwer and M. Perrier.

REFERENCES

M. Akesson. *Probing control of glucose feeding in* Escherichia coli *cultivations*. PhD thesis, Lund Institute of Technology, 1999.

K. B. Ariyur and M. Krstic. *Real-Time Optimization by Extremum-Seeking Control*. John Wiley & Sons, INC., publication, wiley-interscience edition, 2003.

L. Chen, G. Bastin, and V. van Breusegem. A case study of adaptive nonlinear regulation of fed-batch biological reactors. *Automatica*, 31(1):55–65, 1995.

L. Dewasme and A. Vande Wouwer. Adaptive extremum-seeking control applied to productivity optimization in yeast fed-batch cultures. *IFAC 2008, Seoul, Korea*, July 2008.

L. Dewasme, F. Renard, and A. Vande Wouwer. Experimental investigations of a robust control strategy applied to cultures of S. cerevisiae. *ECC 2007 , Kos, Greece*, July 2007.

M. Guay, D. Dochain, and M. Perrier. Adaptive extremum seeking control of nonisothermal continuous stirred tank reactors. *Proc. Adchem 2003, Hong Kong, China*, pages 333–338, 2004.

M. Krstic and H. H. Wang. Design and stability analysis of extremum seeking feedback for general nonlinear systems. *Proceedings of the 1997 Conference on Decision and Control, San Diego, CA*, (TA02-3), 1997.

Y. Pomerleau. *Modélisation et commande d'un procédé fed-batch de culture des levures pain*. PhD thesis, Département de génie chimique. Ecole Polytechnique de Montréal., 1990.

F. Renard. *Commande robuste de bioprocédés opérés en mode fed-batch. Application industrielle à des cultures de* S. cerevisiae. PhD thesis, Faculty of Engineering, Mons, April 2006.

S. Sastry and M. Bodson. *Adaptive Control: Stability, Convergence and Robustness*. Prentice-Hall, Englewood Cliffs, New Jersey 07632, prentice-hall information and system sciences series, thomas kailath edition, 1989.

B. Sonnleitner and O. Käppeli. Growth of S*accharomyces cerevisiae* is controlled by its limited respiratory capacity : Formulation and verification of a hypothesis. *Biotechnol. Bioeng.*, 28:927–937, 1986.

M. Titica, D. Dochain, and M. Guay. Adaptive extremum seeking control of fed-batch bioreactors. *European Journal of Control*, 9(6):618–31, 2003a.

M. Titica, D. Dochain, and M. Guay. Real-time optimization of fed-batch bioreactors via adaptive extremum-seeking control. *Chemical Engineering Research and Design*, 81(A9): 1289–1295, October 2003b.

H. Wang, M. Krstic, and G. Bastin. Optimizing bioreactors by extremum seeking. *Int. J. Adaptive Control and Signal Processing*, 13:651–669, 1999.

Probing Protein Folding Dynamics Using Multivariate Statistical Techniques

Ahmet Palazoglu*, Yaman Arkun, Burak Erman*****
Attila Gursoy****

**Dept. of Chemical Engineering and Materials Science, University of California, CA 95616 USA (Tel: 530-752-8774; e-mail: anpalazoglu@ucdavis.edu)*
***Dept. of Chemical and Biological Engineering, Koç University, Turkey (e-mail: yarkun@ku.edu.tr)*
****Dept. of Chemical and Biological Engineering, Koç University, Turkey (e-mail: berman@ku.edu.tr)*
***** Dept. of Computer Engineering, Koç University, Turkey, (e-mail: agursoy@ku.edu.tr)}*

Abstract: The study of protein folding and its ramifications in biological contexts is at the heart of computational biology. In this paper, we discuss a number of tools in systems engineering that would provide an analysis framework to help explain the observed dynamic behavior of the protein, ultimately making the connection between protein structure and functionality. A case study of villin headpiece folding using principal components analysis as well as clustering demonstrates the potential of these tools in responding to this challenge.

Keywords: optimal folding trajectories, dynamic simulations, principal components analysis, clustering.

1. INTRODUCTION

The study of proteins is easily justified by the fact that they constitute an essential element of all living beings. Specifically, proteins are responsible for controlling gene expression, allow transmission of signals between cells and organs, transport and store other species and defend the body against microbes, among many other functions. Thus, due to their universal significance, understanding the relationships between their sequence, configurations and the vital functions they play in the body, can help development of new therapies and novel biomaterials. As such, uncovering the mysteries of proteins requires an interdisciplinary approach, enlisting not only biologists and medical professionals but also engineers, mathematicians and computer scientists. Recent studies include bioinformatics approaches that explore data mining (Brito, Dubitzky et al. 2004) and evolutionary algorithms (Pal, Bandyopadhyay et al. 2006), in addition to structure prediction problems (Krogh, Larsson et al. 2001; Floudas 2007) and computational techniques focusing on optimization (Krogh, Larsson et al. 2001; Greenberg, Hart et al. 2004).

All protein molecules are chains of amino acids and referred to as linear heteropolymers due to the unbranched nature of their monomeric units (amino acids) (Creighton 1993). The amino acid building blocks consist of a central α-carbon (C^{α}) atom surrounded by four groups: an amino group ($-NH_2$), a carboxyl group (-COOH), a hydrogen atom and a fourth group (-R) that can be one of twenty specific molecules, and is referred to as the side group. The specific side group gives the amino acid its unique characteristic. The sequence of amino acids (also called residues) as read from the amino (N-terminus) to the carboxyl (C-terminus) is referred to as the primary structure. Helices, β-strands, loops, etc. are the secondary structures. Organization of the secondary structures in space to form a stable 3-D structure leads to the tertiary structure. The lowest free energy tertiary structure is the unique native conformation with which the protein performs its function.

The type of a protein and its folding characteristics are determined by its primary structure, i.e., the sequence of amino acids (Dill, Bromberg et al. 1995). It is also noted that the folding process is often aided by molecular chaperons that help the protein fold correctly as it exits the ribosome by minimizing the influence of other nearby proteins as well as by binding to the protein to prevent misfolding (Shinde and Inouye 2000). This is especially important as incorrectly folded proteins resulting from errors during folding are responsible for illnesses such as Creutzfeldt-Jakob disease, Bovine spongiform encephalopathy, Parkinson's and Alzheimer's diseases. Due its implications in understanding such diseases, the dynamics of folding has received substantial attention in recent years (Karplus and Kuriyan 2005; Colombo and Micheletti 2006).

The dynamics of protein folding have been studied extensively in vitro, where the protein is denatured to assume an arbitrary initial configuration and then as the natural conditions are restored, folds into its native configuration. This refolding process has been explored both by molecular dynamics (MD) simulations (Duan and Kollman 1998; Pan and Daggett 2001; Mori, Colombo et al. 2005) and using mostly stop-flow experiments and NMR spectroscopy (Eaton, Thompson et al. 1996; Plaxco and Dobson 1996) and the results provided unique insight towards the folding

dynamics. During the refolding process, the simulations explore the conformational energy landscape accessible to the protein molecule and all-atom MD simulations with explicit solvent can only feasibly achieve time scales shorter than about 1μs for relatively small proteins which leaves out a number of phenomena inaccessible and poorly understood.

In this paper, we show how systems engineering tools can be used to probe the dynamics of protein folding to provide a better understanding of the key mechanisms. The next section discusses protein folding simulations and the type of information gathered as a result. Dynamic folding trajectories that result from such simulations can be interrogated by a number of analysis tools, and we focus on the use of Karhunen-Loeve and clustering to extract spatial and temporal features to help explain the folding dynamics.

2. SIMULATIONS OF PROTEIN FOLDING

Folding of a protein takes place in the form of a competition between the loss of configurational entropy and the decrease of energy due to the formation of inter-residue contacts. Consequently, a free energy barrier separates the unfolded and folded states. The energy surface, as a function of the variables active in folding, such as the $3N$ coordinates of an N atom protein and a multitude of additional dimensions describing the surrounding water molecules, is called the 'energy landscape'. The competition of entropy and energy results in a rugged landscape, and leads to transient trapping of structures that are either partially folded or misfolded. A comprehensive account of protein folding simulations can be found in a recent article (Scheraga, Khalili et al. 2007).

The protein can be modeled at different levels of complexity ranging from all-atom to coarse-grained representations. The all-atom visualization coupled with MD simulations gives the most detailed picture of folding but the computational time is a serious bottleneck. The only full-trajectory molecular dynamics simulation in the presence of explicit water up to this date is that of a 35-residue protein (Duan and Kollman 1998). In typical coarse-grained approximation approaches (Haliloglu and Bahar 1998; Doruker, Jernigan et al. 2002), the protein consists of N beads that represent the amino acids joined into a linear chain by virtual bonds analogous to the chemical bond. A virtual bond joins two consecutive alpha carbons, C^{α}, along the chain. The length of a virtual bond is fixed, a condition referred to as the 'fixed bond length condition'. Each bead has a finite volume. No bead shares its own volume with any other bead. This is called the 'excluded volume condition'. Folding of the protein progresses from a random initial state at $t = 0$ to the final state at $t = t_f$, subject to the fixed bond length and excluded volume conditions at all stages of folding. Folding to the native configuration requires the specification of interactions between pairs of amino acids. This information is based on empirical energy functions, chosen such that the unique native state corresponds to the minimum of total energy (Erman and Dill 2000).

The protein folding problem in its simplest form may be viewed as *a constrained optimization problem:* We are given an initial configuration of N beads connected in the form of a linear chain. The beads want to move towards their specified final destinations by spending minimal energy subject to the (i) connectivity between beads, (ii) fixed bond length and (iii) excluded volume constraints. Each bead obeys Newton's second law of motion throughout the folding trajectory. The forces acting on each bead are received either from the other beads or they are external interaction forces with the environment. Under these conditions, one needs to determine the optimal forces acting on the beads and the resulting optimal trajectory of the beads from their initial configuration to their final native states.

Here, we analyze the optimal pathways followed by the protein during folding. These pathways were generated using the optimal control framework proposed in our earlier work (Guner, Arkun et al. 2006). A coarse grained dynamic model based on Newton's equation of motion is used to make the dynamic optimization manageable.

3. INTERROGATION OF SIMULATION DATA

While simulations provide a wealth of data on the nature of protein motion, extraction of useful information that would shed light on the dominant folding/unfolding mechanisms, evolution of interactions among key residues such as those that determine the hydrophobic core, as well as understanding of the structural conformations and their relationship with biological function is non-trivial. The computational burden and complexity are significant barriers; thus, methods that help reduce dimensionality and provide analytical capabilities in a low-dimensional subspace are largely used. Here, we discuss two techniques. Karhunen-Loeve expansion (KLE) or Principal Components Analysis (PCA) can extract key coordinates (modes) that govern the global motion of the protein. Clustering helps classify large-scale correlated motions that can explain the presence of meta-stable states in which certain protein configurations exist and evolve.

3.1 Principal Components Analysis

The application of PCA to the study of macromolecular motion dates back many years where MD simulations were studied to identify fluctuation modes (Garcia 1992) and to extend simulation time scales (Amadei, Linssen et al. 1993). In the latter work, the conformational space is subdivided into an 'essential' subspace (Van Aalten, De Groot et al. 1997) which contains only a few degrees of freedom, exhibiting unharmonic motion and a 'residual' subspace where the fluctuations are Gaussian. Recent studies explore the energy landscape and the conformational states (Alakent, Doruker et al. 2005; Mu, Nguyen et al. 2005), identify modes contributing to protein fluctuations in MD simulation of apo-adenylate kinase (Lou and Cukier 2006), and extract key mechanistic features from simulations of chemotrypsin inhibitor 2 (Palazoglu, Gursoy et al. 2004). One can also refer to a comprehensive review article for further details (Stein, Loccisano et al. 2006).

The data matrix can be constructed in various forms depending on the information desired. For example, we can construct a matrix of spatial positions as they evolve in time.

The simulations yield the position vectors of N residues for M time intervals and, after subtracting the temporal mean, this results in a $M \times K$ array, where $K = 3N$. One can also use the fluctuation matrix which captures the jump dynamics governing protein folding. The fluctuation matrix becomes $M \times K$ with $K = N$ and has been previously studied (Palazoglu, Gursoy et al. 2004). Another possibility is to form a matrix in which temporal evolution of the magnitude of the distance between the contact pairs is captured. This matrix would have M time intervals and K would correspond to the total number of short- and long-range contact pairs.

The expansion has K modes (eigenvector directions) and each eigenvalue measures the mean energy of its corresponding mode. Among the class of all linear expansions, KLE is optimal in the sense that, on a subspace of lower dimension $L < K$, it retains the most energy possible. One can retain only the first few L modes that extract the important trends and filter the details deemed insignificant by the user.

3.2 Clustering

Cluster analysis (Everitt, Landau et al. 2001) is a class of statistical methods that seeks to partition a set of N observations (objects) into distinct groups. Each observation corresponds to a particular sampling interval (distinct period in time) for which corresponding measurements are available on the same set of parameters. One of the early applications of clustering to MD simulation data is by Karpen et al. (Karpen, Tobias et al. 1993) where feature vectors (dihedral angle time series) are clustered for a 2.2 ns trajectory of the small peptide YPGDV to identify conformational states during unfolding. When applied to protein models, clustering can classify ensembles of structural models based on their backbone structure, using C^{α} distances as the dissimilarity measure (Domingues, Rahnenfuhrer et al. 2004). The molecular motion of proteins can also be classified using clustering to identify functionally relevant structures (Pan, Dickson et al. 2005) and to gain insight towards the shape of the energy landscape (Plaku, Stamati et al. 2007).

Agglomerative hierarchical clustering is used to identify sampling intervals exhibiting similar 'behavior' based on a chosen metric. It accepts as input a symmetric matrix $\boldsymbol{D}$ whose elements D_{ij} indicate the relative dissimilarity between sampling intervals i and j. Matrix $\boldsymbol{D}$ can derive from various parameters in a given simulation, such as the dihedral angles, internal coordinates and potential energies, and must be properly defined for the cluster solution to be physically meaningful. The hierarchical clustering starts with all objects residing in their own cluster, and by using various linkage rules, proceeds by merging the closest objects, and subsequently, the closest clusters, finally terminating when all objects are collected under a single cluster. The output of the hierarchical clustering algorithm is graphical in nature (a dendrogram), and facilitates the visualization of recurring phenomena manifested in the data. Another popular method, k-means clustering (Everitt, Landau et al. 2001), creates clusters based on the maximization of between-cluster variance and minimization of the within-cluster variance, and often gets trapped in local extrema, requiring multiple initializations. The number of clusters needs to be specified a priori for the k-means algorithm and it starts by randomly populating these clusters and proceeding by the optimization step to reform the clusters. These shortcomings were overcome in a recently proposed aggregated k-means clustering strategy (Beaver and Palazoglu 2006) where an ensemble of cluster solutions, generated by performing many randomly initialized runs of the algorithm, can be aggregated to form a single, hierarchical solution. A recent study discusses the performance of different clustering algorithms applied to MD trajectories (Shao, Tanner et al. 2007).

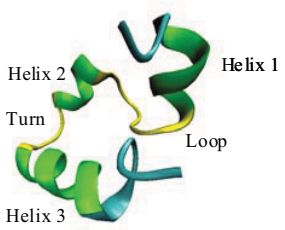


Fig. 1. The structure of villin headpiece.

4. CASE STUDY OF CHICKEN VILLIN HEADPIECE

We consider a 36-residue protein, (PDB code 1Vii.pdb), chicken villin headpiece that is the smallest protein that can fold autonomously. It has been shown through a landmark all-atom explicit water simulation of villin headpiece (Duan and Kollman 1998) that there is a sudden initial hydrophobic collapse followed by longer structural adjustment phase. Other simulation studies also agree with the folding events revealed by Duan and Kollman, e.g., the implicit-water simulation by Shen and Freed (Shen and Freed 2002) and MD scheme integrated with Monte Carlo search by Mori et al. (Mori, Colombo et al. 2005).

Chicken villin headpiece (Figure 1) has 3 short helices, Helix 1, 2 and 3 which contain the residues 4-8, 15-18, and 23-32, respectively. They are held together by a loop between residues 9-14, and a turn between residues 19-22.

Fig. 2. Snapshots from the folding process starting from an arbitrary initial configuration, $t = 0$, followed by $t = 30$, $t = 60$, and $t = 90$, and $t = 150$.

4.1 Folding Trajectories

The optimal folding trajectories were calculated starting from several random initial configurations (Guner, Arkun et al. 2006). Each simulation is performed for 301 time steps. Results include the optimal values for both the position of each bead and the force applied to each bead as a function of time. Figure 2 shows a representative result where the initial denatured configuration is significantly stretched out and the protein starts to establish the helices first. Once the helices form, the loop and the turn secondary structures begin to get established. Finally, the native 3-D structure is reached after refinement of the overall configuration.

The root-mean-square-distance (RMSD) is the distance between the native structure $S_1 = (s_{11} \quad s_{12} \quad \ldots \quad s_{1N})$, and a folding structure $S_2 = (s_{21} \quad s_{22} \quad \ldots \quad s_{2N})$, where s_{ij} is the position of the j^{th} bead in structure i:

$$RMSD = \sqrt{\frac{2}{N(N-1)} \sum_{i}^{N} \sum_{j>i}^{N} \left(\| s_{1i} - s_{1j} \| - \| s_{2i} - s_{2j} \| \right)^2}$$

Figure 3 shows the RMSD variation between the native structure and simulated structures with respect to time for the whole chain. On the average, initial configurations fold around time step 100, with an average final RMSD of 3Å.

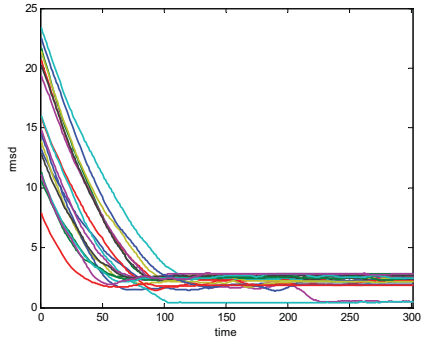
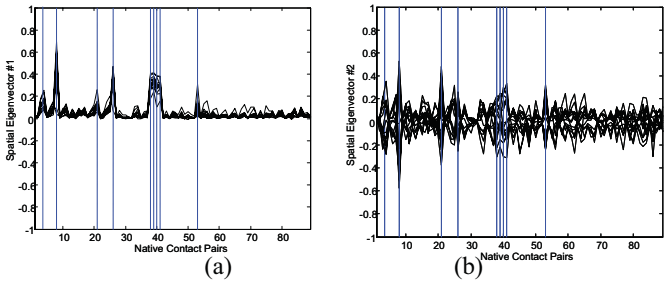


Fig. 3. The evolution of RMSD over all simulations.

4.2 KL Analysis

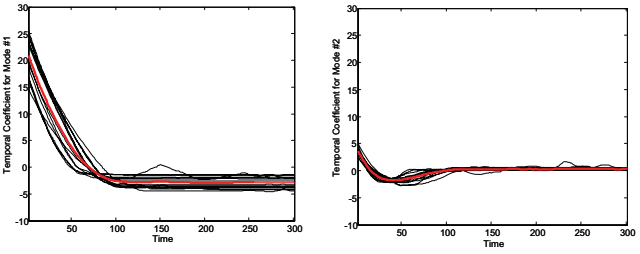
The long-range contact pairs are defined as those residues that are 5 and more residues apart. For chicken villin, there are a total of 89 native contact pairs and 8 are considered as long-range contacts: 2-34, 7-14, 7-34, 10-33, 10-34, 11-33, 11-34, 19-26. Here, we consider the temporal evolution of the magnitude of the contact pairs, r_{ij}'s. The matrix is 301×89 with 301 rows for the time steps and 89 columns for each pair of native contacts. For this analysis, we focused on 14 simulations. We found that, in general, 2 modes capture 99% of the variance in the simulation data. Figure 4 shows the first and second spatial eigenvectors indicating which native contact pairs contribute to these directions the most. The major contributions to the first come from the long-range contacts, as indicated by the vertical lines. The other contact pairs have generally minor contributions to this direction, perhaps with the exception of the pair 2-7 (position 4 on the plot). It is reported (Frank, Vardar et al. 2002) that three phenylalanine residues F47, F51, and F58 (residues 7, 11 and 18) make up the bulk of the hydrophobic core along with the hydrophobic residues L42, V50, and L69 (residues 2, 10 and 29). Thus, it is noteworthy that the first mode is significantly influenced by the interaction between residues 2 and 7 as the hydrophobic collapse occurs. Another observation is that all pairs load positively in this direction, indicating that all move in the same direction, effectively in the direction of reducing the distances among contact pairs. In fact this coordinated collapse is observed in Fig. 1. Another important observation is that this loading behavior is independent of the initial configuration, underscoring the fundamental nature of the collapse. In the second mode, the influence of the long-range contacts is attenuated (especially for 10-33, 10-34, 11-33 and 11-34) and short-range contacts become more important, almost across the board for all such contacts. The native contact pair 2-7 still retains its influence. The contact pairs load in both positive and negative directions, a key difference from the first mode, indicating a more complex motion. It is also important to note that this loading depends on the initial configuration, implying that the formation of secondary structures can follow different paths in time.



(a) (b)

Fig. 4. First (a) and second (b) spatial eigenvectors, vertical lines indicate the position of long-range native contact pairs, and vertical line at position 4 points to the native contact pair 2-7.

As shown in Fig. 5, the temporal coefficient of the first spatial eigenvector decays exponentially, indicating that the manner with which energy is minimized is common for all simulations regardless of the initial configuration. This also supports the all-positive loading directions of the contact pairs as shown in Figure 4a. On the other hand, the temporal coefficient of the second spatial eigenvector shows second-order behavior and is attenuated significantly, underscoring the lesser influence of the second mode. This mode explains the fast dynamics associated with the short-range contacts as secondary structures (helices in this case) are made quickly and then readjusted to conform to the overall formation of the protein structure. Folding dynamics exhibit two time-scales, which is consistent with the two-step folding mechanism of the hydrophobic collapse model (Baldwin 2002).



(a) (b)

Fig. 5. First (a) and second (b) temporal coefficients, bold red line indicates the averages.

4.3 Cluster Analysis

To demonstrate the potential of clustering, we focus on native contact distances analyzed before. For a given simulation, the expectation is to see if the dynamic signature of contact distances can be used to label each contact pair as belonging to a class, distinguished by its characteristic temporal evolution or contact distance. Thus, the feature vector is the

time series of contact distance magnitudes, r_{ij}'s. The simulation data from the 14 runs were stacked into a matrix of dimension 4214×89 and the data were normalized to [0, 1]. The data matrix is then transposed for clustering the 89 contact pairs. Thus, the scaling is in reality performed on the *rows* of the clustered data matrix, as opposed to for the columns as is more typical. A different scaling is used in this analysis because the variables have different mean levels although they are measured in the same units.

Using aggregated *k*-means clustering with average linkage, the dendrogram in Figure 6 is obtained. It shows two coarse and six relatively distinct fine clusters with a cophenetic coefficient (Beaver and Palazoglu 2006) of 0.96, which indicates that the dendrogram is a good representation of the relationships among the objects. The aggregated distances show that the cluster members have short merging distances while the main clusters merge at relatively large distances. This shows that the within-cluster variance is low while the between-cluster variance is high.

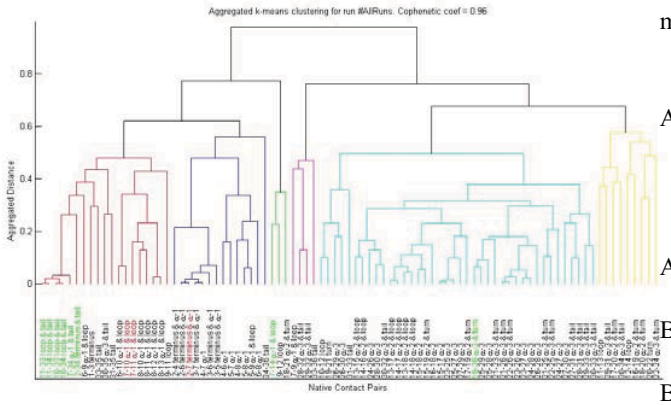


Fig. 6. The dendrogram for clustering the distances of all native contact pairs for 14 simulations. Contacts labelled as green are the long-range native contacts whereas the contacts labelled as red are the ones involving residues in the hydrophobic core.

In Figure 6, it is noted that a large number of the residues that form the long-range native contacts and the hydrophobic core fall in the same cluster. Indeed, first cluster (labeled as red) contains six of the eight long-range contact pairs along with two short-range contact pairs that contain hydrophobic core residues. This cluster captures the concerted motions of the loop and the tail, as well as the loop and helix-1 contacts. The second cluster (dark blue) is largely residues involved in forming the helix-1. It is noted that the long-range native contact pair 7-14 (PHE7-THR14) appears in a rather isolated third cluster (green) and its motion shows greatest similarity to contact pairs 9-12 and 18-21. The long-range interaction, i.e., 7-14, is the only long-range interaction close to the N-terminal, thus having dynamic characteristics different than the other long-range contacts is expected. The small fourth cluster (magenta) contains residues towards the end of the protein chain and also notably differs from the first three clusters. The fifth cluster (light blue) contains all the remaining short-range contacts associated with helix 2 and helix 3, including the long-range contact 19-26. This long-range contact pair shares characteristics common with the other helix 3 contact pairs, thus dynamically acts in concert with a large number of short-range contact pairs specific to the secondary structure with which it is associated. The final cluster (yellow) brings together short-range contacts primarily involving turn and tail secondary structures.

We must re-iterate that a coarse clustering decision indicates two main clusters, containing the subclusters (1, 2, 3) and (4, 5, 6), respectively. Such a grouping would suggest two classes of residues where the first mainly contains the long-range-contacts and the second, the short-range contacts. Yet, the fact that specific residues (such as long-range contacts) do not appear all in a single cluster and usually appear together with other residues is supported by (Larson, Ruczinski et al. 2002) who claim that both poorly and highly conserved residues are equally likely to participate in the protein folding nucleus. They also note, however, that there is an observable bias in the mean sequence conservation of the residues in the folding nuclei. This is especially consistent with the membership of the large cluster on the left.

REFERENCES

Alakent, B., P. Doruker, et al. (2005). "Application of time series analysis on molecular dynamics simulations of proteins: A study of different conformational spaces by principal component analysis " J. Chem. Phys. **121**(10): 4759-4769.

Amadei, A., A. B. Linssen, et al. (1993). "Essential dynamics of proteins." Proteins **17**: 412-425.

Baldwin, R. L. (2002). "Making a network of hydrophobic clusters." Science **295**(5560): 1657-8.

Beaver, S. and A. Palazoglu (2006). "A cluster aggregation scheme for ozone episode selection in the San Francisco, CA Bay Area." Atmospheric Environment **40**(4): 713-725.

Brito, R., W. Dubitzky, et al. (2004). "Protein folding and unfolding simulations: A new challenge for data mining." Omics - J Integrative Biology **8**(2): 153-166.

Colombo, G. and C. Micheletti (2006). "Protein folding simulations: combining coarse-grained models and all-atom simulations." Theor. Chem. Acc. **116**: 75-86.

Creighton, T. E. (1993). Proteins: Structures and Molecular Properties. New York, W.H. Freeman.

Dill, K. A., S. Bromberg, et al. (1995). "Principles of protein folding - a perspective from simple exact models." Protein Science **4**: 561-602.

Domingues, F. S., J. Rahnenfuhrer, et al. (2004). "Automated clustering of ensembles of alternative models in protein structure data." Protein Engineering **17**: 537-543.

Doruker, P., R. L. Jernigan, et al. (2002). "Dynamics of large proteins through hierarchical levels of coarse-grained structures." J. Comp. Chem. **23**: 119-127.

Duan, Y. and P. A. Kollman (1998). "Pathways to a protein folding intermediate observed in a 1-Microsecond simulation in aqueous solution." Science **282**(5389): 740-744.

Eaton, W. A., P. A. Thompson, et al. (1996). "Fast events in protein folding." Structure **4**: 1133-1139.

Erman, B. and K. A. Dill (2000). "Gaussian theory of protein folding." J. Chem. Phys. **112**: 1050-1056.

Everitt, B., S. Landau, et al. (2001). Cluster Analysis. New York, NY, Oxford.

Floudas, C. A. (2007). "Computational methods in protein structure prediction." Biotech. and Bioeng. **97**(2): 207-213.

Frank, B. S., D. Vardar, et al. (2002). "The role of aromatic residues in the hydrophobic core of the villin headpiece subdomain." Protein Science **11**(3): 680-687.

Garcia, A. E. (1992). "Large-amplitude nonlinear motions in proteins." Phys. Rev. Lett. **68**: 2696-2699.

Greenberg, H., W. Hart, et al. (2004). "Opportunities for combinatorial optimization in computational biology." Informs J Computing **16**(3): 211-231.

Guner, U., Y. Arkun, et al. (2006). "Optimum folding pathways of proteins. Their determination and properties." J. Chem. Phys. **124**: 134911.

Haliloglu, T. and I. Bahar (1998). "Coarse-grained simulations of conformational dynamics of proteins: application to apomyoglobin." Proteins: Structure, Function, and Genetics **31**: 271-281.

Karpen, M. E., D. J. Tobias, et al. (1993). "Statistical clustering technique for the analysis of long molecular dynamics trajectories: Analysis of 2.2-ns trajectories of YPGDV." Biochemistry **32**: 412-420.

Karplus, M. and J. Kuriyan (2005). "Molecular dynamics and protein function." PNAS **102**(19): 6679-6685.

Krogh, A., B. Larsson, et al. (2001). "Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes " J. Mol. Biology **305**(3): 567-580.

Larson, S. M., I. Ruczinski, et al. (2002). "Residues participating in the protein folding nucleus do not exhibit preferential evolutionary conservation." J Mol Biol **316**(2): 225-33.

Lou, H. and R. I. Cukier (2006). "Molecular dynamics of apo-adenylate kinase: A principal component analysis." J. Phys. Chem. B **110**: 12796-12808.

Mori, G. M. S., G. Colombo, et al. (2005). "Study of the villin headpiece folding dynamics by combining coarse-grained Monte Carlo evolution and all-atom molecular dynamics." Proteins: Structure, Function, and Bioinformatics **58**: 459-471.

Mu, Y., P. H. Nguyen, et al. (2005). "Energy landscape of a small peptide revealed by dihedral angle principal component analysis." Proteins: Structure, Function, and Bioinformatics **58**: 45-52.

Pal, S. K., S. Bandyopadhyay, et al. (2006). "Evolutionary computation in bioinformatics: A review." IEEE Trans. Systems Man and Cybernetics C - Applications and Reviews **36**(5): 601-615.

Palazoglu, A., A. Gursoy, et al. (2004). "Folding dynamics of proteins from denatured to native state: principal component analysis." J. Comp. Biology **11**: 1149-1168.

Pan, P. W., R. J. Dickson, et al. (2005). "Functionally relevant protein motions: Extracting basin specific collective coordinates from molecular dynamics trajectories." J. Chem. Phys. **122**: 034904.

Pan, Y. P. and V. Daggett (2001). "Direct comparison of experimental and calculated folding free energies for hydrophobic deletion mutants of chymotrypsin inhibitor 2: Free energy perturbation calculations using transition and denatured states from molecular dynamics simulations of unfolding." Biochemistry **40**(9): 2723-2731.

Plaku, E., H. Stamati, et al. (2007). "Fast and reliable analysis of molecular motion using proximity relations and dimensionality reduction." Proteins: Structure, Function, and Bioinformatics **67**: 897-907.

Plaxco, K. W. and C. M. Dobson (1996). "Time-resolved biophysical methods in the study of protein folding." Current Opinion in Structural Biology **6**: 630-636.

Scheraga, H. A., M. Khalili, et al. (2007). "Protein-folding dynamics: Overview of molecular simulation techniques." Annual Review of Physical Chemistry **58**: 57-83.

Shao, J. Y., S. W. Tanner, et al. (2007). "Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms." Journal of Chemical Theory and Computation **3**(6): 2312-2334.

Shen, M. Y. and K. F. Freed (2002). "All-atom fast protein folding simulations: the villin headpiece." Proteins: Structure, Function, and Genetics **49**: 439-445.

Shinde, U. and M. Inouye (2000). "Intramolecular chaperones: polypeptide extensions that modulate protein folding." Cell and Developmental Biology **11**: 35-44.

Stein, S. A. M., A. E. Loccisano, et al. (2006). "Principal components analysis: A review of its application on molecular dynamics data." Ann. Rep. in Comp. Chemistry **2**: 233-261.

Van Aalten, D. M. F., B. L. De Groot, et al. (1997). "A comparison of techniques for calculating protein essential dynamics." J. Comp. Chemistry **18**(2): 169-181.

Applied Advanced Process Analytics in Biopharmaceutical Manufacturing: Challenges and Prospects in Real-time Monitoring and Control

Cenk Ündey*, Sinem Ertunç, Thomas Mistretta, Manuj Pathak

Amgen Inc.,
Process Development
Process and Systems Analysis
40 Technology Way, West Greenwich, RI 02817 USA.

Abstract: Biopharmaceutical manufacturing processes are inherently complex due to their nonlinear bioprocess dynamics, variability in batch operations and manufacturing schedule, raw materials involved, and automatic process control. A typical processed lot generates large amounts of data that need to be analyzed and interpreted for process troubleshooting and continuous improvement purposes in addition to product release. Multivariate Batch Process Modeling, Monitoring and Control approaches in real-time are elaborated by providing industrial examples from the commercial manufacturing processes. Examples and opportunities in cell culture (e.g., bioreactor applications) and purification (e.g., large-scale chromatography) operations are summarized. Impact of Process Analytical Technologies (PAT), soft-sensor development, first principles modeling applications and commercial-scale examples are presented.

Keywords: Process Analytical Technologies, Quality-by-Design, Biopharmaceutical Manufacturing, Real-Time Multivariate Process Monitoring, Soft-sensors.

1. INTRODUCTION

Biotechnology-based products have become increasingly important in recent years in treating chronic diseases such as cancer and arthritis. Although biopharmaceutical drugs constitute a small portion of (about 8% in 2004) of the pharmaceutical market, approximately 27% of new medicines in active development are now biotech products (Business Insights, 2005). Process development and commercial manufacturing of these products, such as therapeutic proteins; require a good understanding of chemistry, manufacturing and controls. Advanced process analytical technologies can be incorporated in process development as well as into commercial-scale manufacturing for advanced monitoring, control, continual process improvement, cost reduction and risk management.

Biopharmaceutical processes are typically comprised of a series of unit procedures operated in batch mode to produce therapeutic proteins, and have complex biological mechanisms that result in non-linear and time-variant process dynamics. This makes their modeling, monitoring and control challenging. In a typical commercial-scale biopharmaceutical manufacturing process there are multiple batch processing unit operations where off-line samples are taken to ensure in-process control and quality objectives are met and real-time measurements are made for open and closed-loop control and monitoring. A number of measurements are also made for raw material release testing towards use in manufacturing. When this is looked at from a holistic perspective; there are many batches, variables and operational characteristics to analyze in a meaningful and proactive manner. While this goal can be achieved via post-mortem analysis, it is more desirable to monitor and control these multivariable multi-stage biopharmaceutical processes in real-time. Advanced process analytics in the form of real-time multivariate process monitoring and control provides an efficient means of identifying/reducing variation, managing process risks, relating process information to critical quality attributes (CQAs) and determining process improvement opportunities such as increasing yields and decreasing impurities. Due to long lead times and inflexibility in the regulatory approval process, applying process changes for improvement has not been very straightforward in the conventional regulatory paradigm. The United States Food and Drug Administration has published a series of new guidelines to address this issue and provide more flexibility for innovation while appropriately managing the risk around proposed process changes for continual improvement via process analytics and quality-by-design (QbD) principles (US FDA, 2002, 2004, 2006, Chirino and Mire-Slius, 2004). *Process Analytical Technology* (PAT) is defined as a system for designing, analyzing, and controlling manufacturing through real-time measurements of critical quality and performance attributes of the process as well as raw materials and other process inputs. QbD promotes improved process understanding

* Corresponding author. *E-mail*: *Cenk.Undey@ amgen.com*

during process and product development and building quality in the design instead of testing for quality. This can be achieved via correlative, causal, or mechanistic knowledge and at the highest level via first principle models (Cinar, et al., 2003, Rathore and Winkle, 2009).

Process analytical approaches including deploying multivariate techniques along with on-line/at-line/in-line sensors and analyzers and the use of chemometrics in biotechnology have received significant attention in recent years (Junker and Wang, 2006). Applications included monitoring and control of microbial fermentations (Albert and Kinley, 2001, Lopes, et al., 2004, Undey et al., 2004, Gnoth, et al., 2007), cell culture (Arnold, et al., 2003, Undey, et al, 2006) and purification processes (Larson, et al., 2003, Lee et al., 2006, Rathore, et al., 2008). Development of new measurement technologies for real-time bioprocess monitoring and control is also crucial (Becker, et al., 2007).

In this paper, a generic approach for developing further process understanding, modelling, monitoring and control is summarized. Specific applications and case studies are provided from commercial manufacturing experience in the use of process analytics. The potential of soft-sensors and first principles modeling in biopharmaceutical manufacturing is discussed with industrial examples. Use of process analytics and real-time multivariate monitoring technology in operational success is also demonstrated. Challenges and prospects of adaptive process control are discussed.

2. *PROCESS UNDERSTANDING VIA MODELING AND ADVANCED MONITORING*

Aforementioned QbD approach demands a high-level of process understanding for ensuring control of CQAs. Levels of process understanding and knowledge can be categorized in increasing order (lowest to highest) as descriptive, correlative, causal, mechanistic and first principles-based. While mechanistic and first principles models provide the highest level of understanding and predictability, their development and adaptation may not be very straightforward in biopharmaceutical process development and manufacturing that use design of experiments (DOE)-based approaches along with other heuristic knowledge about the process and product. During commercialization of a product an essential level of process understanding is demonstrated to ensure process consistency, product safety, efficacy and purity. However, in order to expedite time-to-market (while the drug is meeting the efficacy, purity and safety requirements) and make the product available to the patients, its manufacturing process may not be sufficiently optimized early in the product lifecycle. Additionally, there may be scale-up effects, raw material lot-to-lot variability (i.e., as an unmeasured load disturbance to the process) and other operational aspects such as maintenance schedules and human factors collectively driving the overall process variability. Real-time multivariate statistical process monitoring (RT-MSPM) provides a means to proactively monitor this overall process variability and build the necessary foundation towards predictive monitoring and multivariable control. In the generic methodology proposed in this paper, a multivariate model for each unit operation is developed for advanced monitoring and prediction purposes where applicable. Based on the frequency of data availability, models are used in real-time and/or via post-mortem batch analytical purposes. Making the data available and establishing required databases, connections to source systems, data pre-treatment and reconciliation are practical considerations that need to be addressed in industrial setting prior to enabling RT-MSPM (Undey, 2008).

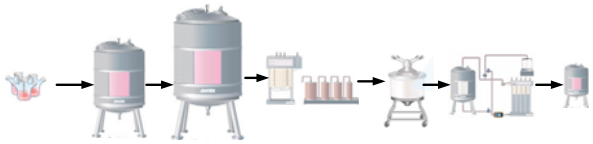


Fig. 1. Simplified process flow of a biologics manufacturing process (inoculation, scale-up bioreactors, clarification and purification train).

As shown in Fig. 1 typical biopharmaceutical manufacturing processes involve multiple unit operations including bioreactors for scale-up, cell growth/protein production, clarification ultrafiltration and chromatography columns and skids. There may be more steps based on each product's requirement and several parallel trains for plant throughput maximization.

3. *MULTIVARIATE BATCH MODELING AND MONITORING*

Multivariate (MV) modeling techniques such as Principal Components Analysis (PCA) and Partial Least Squares (PLS) are used to handle batch process data issues such as large number of variables, colinearity and missing data while summarizing the overall variability in the principal component and latent variable space. Historical in-control batches are used for MV model development. Control limits for MV statistics are calculated and MSPM charts are used for efficient monitoring (Nomikos and MacGregor, 1995).

In this study batch process data are analyzed and monitored in two hierarchical levels. The first level is called the observation level and used to monitor the batch evolution with respect to a maturity variable in real-time. The second level is the batch level which is used for monitoring the batch fingerprint and can be used in predicting a final performance variable (Wold et al., 1998). Since early in the progress of a batch the confidence in the prediction is typically low, care needs to be taken when to start using batch level MV charts for real-time monitoring after a certain amount of data is available. All the real-time batch level cases presented in this paper starts computing the batch level MV statistics when 50% of the batch is completed and it has provided a good predictive performance.

Umetrics' Simca-P+ and SBOL (Simca-Batch On-Line) modules are used for the MSPM cases presented in this paper. Simca-P+ is the tool used for offline MV model development. SBOL is the tool used for real-time MSPM and uses the MV models developed by Simca-P+.

4. *DESIGN SPACE MONITORING AND POSTMORTEM ANALYSES OF BATCH PERFORMANCE*

During initial bioprocess design and characterization the full variability due to raw material lots, scale-up parameterization and large-scale operations cannot be estimated and therefore the process is monitored to ensure consistency. Process design space is constructed to understand the ranges on key and critical operating variables (process inputs) and their impact on the key/critical performance variables (process outputs) at bench and pilot-scale manufacturing and later scaled-up to commercial operation. Multivariate models can be helpful in also comparing the bench and/or pilot-scale experience against large-scale manufacturing to study similarities and identify any differences in performance. It can also be used to improve bench/pilot-scale model representation against large-scale so that process development can improve on the scale up parameterization. Representative scaled-down process models (i.e., actual scaled-down equipment) are crucial for troubleshooting and process improvement experimentation. In the following example (Fig. 2), PCA-based multivariate models were applied to process performance variables from batches performed at both bench and large-scale manufacturing.

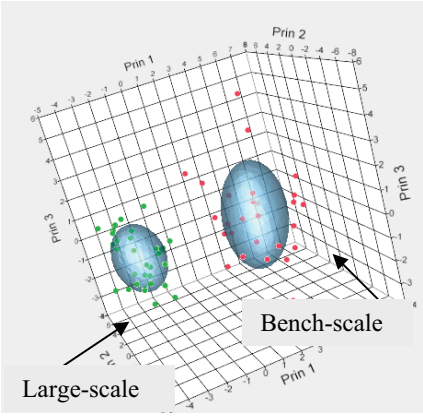


Fig. 2. Multivariate comparison and monitoring of different manufacturing scales against design space

With this approach, variability in process data from more than twenty performance variables is summarized with only three principal components. Ellipsoids represent 95% confidence volume. In this case, large-scale batches seem to be more tightly controlled compared with bench-scale batches (partly due to a wider range of input space explored in bench-scale and scale-up differences). Multivariate monitoring of both scales provide a means of comparison as well as identifying improvement areas where changes can be made to move the two spaces closer to each other in terms of expected correlations, variability and means.

As a postmortem case study, multivariate modeling is performed retrospectively on commercial manufacturing batches to support technical investigations, identify process and/or operational improvement opportunities, the sources of process variation to increase process understanding. Data from historical batches is used to develop PLS-based models for key process performance variables. Information from process characterization studies can also be leveraged to further refine the PLS models. As an example, a low product titer trend was observed in manufacturing of a commercial biologic. A PLS-model was developed for product titer using multiple process inputs from 36 historical batches. Based on a review of exceptional batches in the MV charts (fingerprint in Fig. 3 and variable contribution chart in Fig. 4), it was determined that the cell specific productivity was significantly lower in the decreased product titer batches. This information guided process analysts to focus on operational parameters likely to adversely impact cell specific productivity. This analysis revealed that a shift in induction timing potentially contributed to the low product titer trend. After the induction timing was adjusted back to target, higher product titers were obtained.

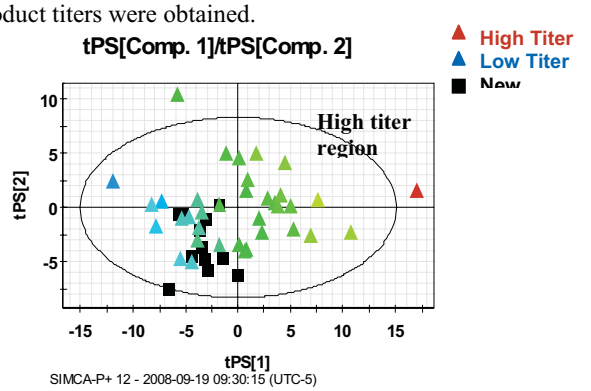


Fig. 3. A score scatter plot showing new manufacturing batches relative to historical ones.

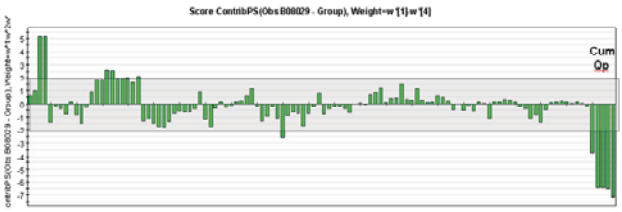


Fig. 4. Variable contribution plot to Hotelling's T^2 showing low cell specific productivity as a significant contributor to low titer (horizontal band indicates +/- 2 standard deviation about the mean batch).

With introduction of new on-line technologies such as cell density probes, pCO_2 probes and Glucose/Lactate sensors comes the possibility of applying multivariate process modeling to predict and control biologic manufacturing processes towards achieving target productivity and quality end points. Several opportunities exist to improve the process models using new and available characterization data and increase model sensitivity and predictability. While it is very informative to use MSPM for postmortem exploratory analysis of process upsets, it is more desirable to monitor the process in real-time to detect and diagnose those upsets and take preventive actions where possible. In addition to monitoring, prediction of key end points while the process is in progress is also possible (a.k.a. soft-sensors) and provides

various opportunities towards more efficient process operations and advanced process control and optimization.

5. *REAL-TIME MULTIVARIATE BIOPROCESS MONITORING CASE STUDIES*

5.1. *CASE STUDY-1: REAL-TIME LARGE-SCALE BIOREACTOR CELL CULTURE MONITORING*

The first example is from a bioreactor monitoring and shows how real-time monitoring technology is used towards operational excellence, hence identifying equipment and mechanical related issues as well. A transient decline (~3%) is observed in Final Viabilities (measured offline) in a Perfusion Bioreactor across batches. Deviations in real-time MV charts are detected for those low viability batches (*Hotelling's* T^2 chart for one of the low viability batches is shown in upper Fig. 5). Variable contribution plots identified that the low Final Viability batches that are run on particular bioreactor and its skid had higher Perfusion Feed and Retentate Temperature compared to the historical batches. Further investigation revealed that the wrong size gaskets were installed in the pump seal flush line for these batches. This made the control of temperature and pressure on the seal flush line very difficult. The correct size gaskets were replaced prior to the next batch and the temperature profile and final viability were within their normal ranges.

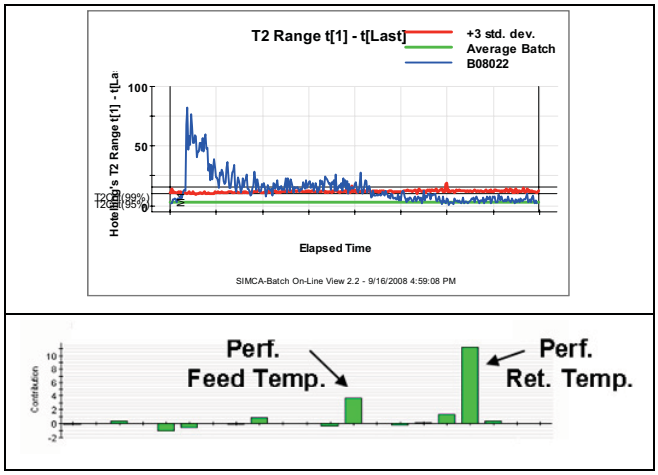


Fig. 5. Hotelling's T^2 chart (top) for one of the low final viability batches and contributions to the deviation (bottom)

5.2. *CASE STUDY-2: REAL-TIME MONITORING OF LARGE-SCALE CHROMATOGRAPHY PROCESS*

PCA models were developed for each phase of the Protein A affinity chromatography unit operation (e.g., pre-equilibration, elution) from historical process data. In general, only two or three principal components were needed for a given phase to summarize the batch process. Future batches could then be projected onto the model to allow rapid detection of deviations from the normal operating space and corrective actions would be taken where possible. Model has detected failure of a pre-column pH probe during purification of a batch through Protein A chromatography. The online pH probes are commonly used to verify proper equilibration of a column or end-point of a titration for a viral inactivation step prior to taking a confirmatory offline sample. In this example, the *Hotelling's* T^2 plot identifies the out of trend batch (on the left at Fig. 6) in real-time and the variable contribution plot (on the right hand side) identifies the pH probe having a significant effect. This information was used to replace the probe prior to further processing and avoided compromising any offline verification samples.

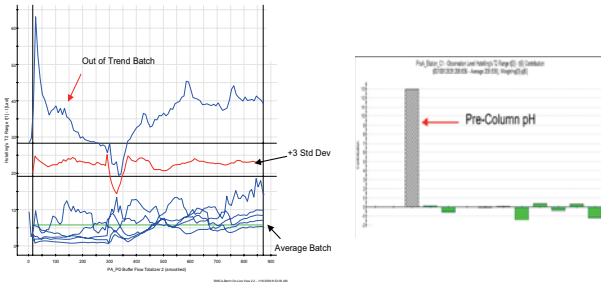


Fig. 6. Hotelling's T2 chart (on the left) indicating an out-of-control batch and the contribution plot diagnosing the faulty pH probe (on the right)

6. *REAL-TIME SOFT (VIRTUAL)-SENSOR APPLICATIONS IN BIOPHARMACEUTICAL MANUFACTURING*

There are many on-line/at-line/in-line probes and analyzers available for measuring bioprocess variables and quality/performance indicators. In a typical setting, for instance a bioreactor has temperature, level, pH, pressure, agitator speed, aeration rate, and dissolved oxygen measurement systems and data acquisition. Additional measurement systems such as cell density, dissolved carbon dioxide, mass spectroscopy in the off-gas, on-line HPLC, and fluorometric sensors are among the available and desirable technology. Research has shown that there are mathematical means via empirical and first principles modeling to generate predictions on the performance end-points in real-time (Cheruy, 1997, Undey, et al., 2006). A generic framework is depicted in Fig. 7, where more frequently measured and readily available process variables are used for developing a soft-sensor to generate an on-line quality estimate. Industrial examples are given for a cell culture production bioreactor case where product final titer is predicted in real-time and a first-principles model for monitoring a chromatography column performance in real-time.

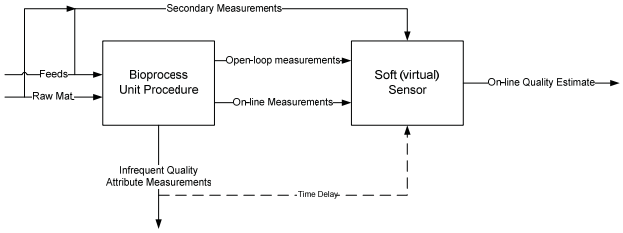


Fig. 7. Soft (virtual)-sensors for bioprocesses.

6.1. *CASE STUDY-3: REAL-TIME SOFT-SENSOR FOR BIOREACTOR FINAL TITER PREDICTION*

The production bioreactor is a critical step for biopharmaceutical processes since the target protein is expressed during this step. It is very important to closely monitor this unit operation in real-time. Any deviations from the normal operation during this step may lead to low productivity and out-of-spec product.

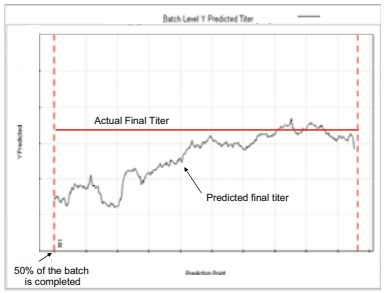


Fig. 8. Real-time Final Titer prediction starts when 50% of the batch is completed.

Final Titer is used as the performance parameter that is predicted in the batch level while monitoring the process in real-time. Accurately predicting Final Titer several days in advance of harvesting provides many PAT opportunities such as offline assay elimination (titer is typically determined via an HPLC method offline), titer optimization, schedule optimization and real-time control. Fig. 8 illustrates real-time final titer prediction for a batch by only using the continuously measured inputs and outputs such as pH, O_2 flow and bioreactor volume as predictors. Final titer is predicted within 0.5 % of the actual offline testing result.

6.2. *CASE STUDY-4: REAL-TIME SOFT-SENSORS FOR CHROMATOGRAPHY OPERATION*

In this case study for biologics purification has been real-time determination of unit operation step yields. These are typically calculated by measuring offline samples of the load and product pools, and results are often not available until well after batch completion. UV detection unit is typically available by the production unit, Beer's Law (A=ϵBC, i.e., absorbance is proportional to concentration of a solution, where ϵ is molar absorptivity, B, path length of the sample and C, concentration of the solution) can be used as a basis to correlate the absorbance to Protein Concentration UV in real time (Fig. 9). Estimated concentration can then be used to determine unit operation step yield.

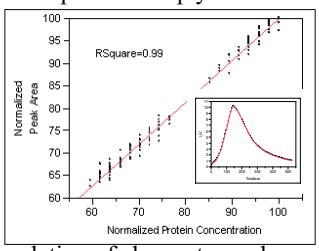


Fig. 9. Correlation of chromatography peak area (shown as inset) to protein concentration.

For process chromatography operating in bind and elute mode, the eluate peak can be integrated numerically. The same can be done for phases of the chromatography operation, such as regeneration, to complete the mass balance. Incorporating these mass balance terms into the PLS model significantly improves the predictive power.

Another area of exploration for advanced monitoring has been the use of first principles models to predict chromatography resin binding capacity. The resin capacity is typically determined during characterization studies by measuring the percentage breakthrough of protein during the loading phase. Due to the large volume requirements, this is typically not measured routinely once the process is implemented at manufacturing scale. This capacity measurement is useful to understand the changes in resin characteristics during the operational lifetime of the column and make decisions about when to change out the resin to ensure high performance.

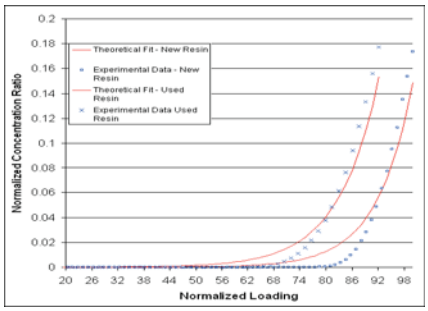


Fig. 10. Capacity changes of resin used in 90 batches of biopharmaceutical purification process.

Using the analytical solution to the differential mass-balance equation developed by Ghose et al. (2004), actual resin capacity at different points in the column lifetime was modeled. The model inputs were determined from literature and via process knowledge. This capacity information can then be compared against the expected column loading for a given batch to determine if any losses would be expected (Fig. 10). It can also be incorporated as an input variable to the soft-sensor (multivariate model in this case) to provide a more accurate prediction of step yield.

7. *CONCLUSIONS*

Process analytical technology (PAT) within the Quality-by-design framework provided by regulatory agencies define guidelines towards demonstrating good process understanding in product development and therefore requires improved monitoring and control for pharmaceutical manufacturing processes. Real-time multivariable monitoring framework for bioprocesses as shown in the examples not only provides proactive process supervision but also serves as an operational success management tool, hence, it has the potential to reduce production costs since it is also used to monitor and detect the equipment-related issues.

Soft (virtual)-sensors were also studied in this paper and promising examples were provided from cell culture and

purification areas to show how existing process and its data can be utilized for prediction of key performance end points. While advanced monitoring capabilities are demonstrated, it is important to understand how process control (e.g., feedback and/or feed forward) can be achieved to ensure consistent process outputs based on biological performance end points. This paper summarized the necessary preliminary steps of process understanding and advanced monitoring applied in biopharmaceutical manufacturing environment. New measurement technologies will be necessary to enable closed-loop control of key performance variables and CQAs to meet biopharmaceutical manufacturing needs. It is necessary that these on-line quality measurement systems that are available or being developed to have faster turn around times than the process time (i.e., the on-line test results can be available before the relevant biological phase is completed) so that a control action can be taken in a timely manner. Advanced multivariable monitoring with robust, accurate and improved on-line/at-line/in-line probes and analyzers with fully-automated data management and multivariable control capabilities hold great possibilities for 21st century PAT-enabled biopharmaceutical manufacturing.

REFERENCES

Albert, S., and Kinley, R. (2001). Multivariate statistical monitoring of batch processes: an industrial case study of fermentation supervision. *Trends in Biotechnology*, 19(2), 53-62.

Arnold, S.A., Crowley, J., Woods, N., Harvey, L.M., and McNeil, B. (2003). In-situ near infrared spectroscopy to monitor key analytes in mammalian cell cultivation, *Biotechnology and Bioengineering*, 84(1), 13-19.

Becker, T., Hitzmann, B., Muffler, K., Pörtner, R., Reardon, K.F., Stahl, F. and Ulber, R. (2007). Future aspects of bioprocess monitoring, *Advances in Biochemical Engineering and Biotechnology*, 105, 249-293.

Business Insights (2005). *The biotechnology market outlook: growth opportunities and effective strategies for licensing and collaborations*, MindBranch, Inc.

Cheruy, A. (1997). Software sensors in bioprocess engineering, *Journal of Biotechnology*, 52, 193-199.

Chirino, A.J., and Mire-Slius, A., 2004. Characterizing biological products and assessing comparability following manufacturing changes. *Nature Biotechnology*, 22(11), 1383–1391.

Cinar, A., Parulekar, S. J., Undey, C., and Birol, G. (2003). *Batch Fermentation: Modeling, Monitoring and Control*, CRC Press, New York, NY.

Ghose, S., Nagarath, D., Hubbard B., Brooks, C., and Cramer, S.M. (2004). Use and optimization of a dual-flow rate loading strategy to maximize throughput in protein-A affinity chromatography. *Biotechnology Progress,* 20(3), 830-840.

Gnoth, S., Jenzsch, M., Simutis, R., and Lubbert, A. (2007). Process Analytical Technology (PAT): Batch-to-batch reproducibility of fermentation processes by robust process operational design and control, *Journal of Biotechnology*, 232, 180-186.

Junker, B.H., and Wang, H.Y. (2006). Bioprocess monitoring and computer control: Key roots of the current PAT initiative, *Biotechnology and Bioengng*, 95(2), 226-261.

Larson, T.M., Davis J., Lam, H., and Cacia, J. (2003). Use of process data to assess chromatographic performance in production-scale protein purification columns. *Biotechnology Progress*, 19, 485–492.

Lee, W., Pasumarthy, M., Undey, C., Kundu, S., and Stout, J. (2006). Assessment of production-scale protein purification column packing through transition analysis, *RXII–Recovery of Biological Products XII,* Phoenix, AZ.

Lopes, J.A., Costa, P.F., Alves, T.P., and Menezes, J.C. (2004). Chemometrics in bioprocess engineering: Process analytical technology (PAT) applications, *Chemometrics and Intelligent Lab. Systems,* 74, 269-275.

Nomikos, P., and MacGregor, J.F. (1995). Multivariate SPC charts for monitoring batch processes, *Technometrics*, 37, 41-59.

Rathore, A.S., Yu, M., Yeboah, S., and Sharma A. (2008). Case study and application of process analytical technology (PAT) towards bioprocessing: use of on-line high-performance liquid chromatography (HPLC) for making real-time pooling decisions for process chromatography, *Biotechno.l Bioeng.*, 100(2), 306–316.

Rathore, A.S., and Winkle, H. (2009). Quality by design for biopharmaceuticals, *Nature Biotechnology*, 27(1), 26-34.

Undey, C., Tatara, E. and Cinar, A. (2004). Intelligent Real-time Performance Monitoring and Quality Prediction for Batch/fed-batch Cultivations, *J Biotechnol.*, 108, 61-77.

Undey, C., Geldart, R.W., Collins, E., Kundu, S., Humphreys, L., Johnson, K., and Frohlich, B. (2006). A case study for PAT initiative project: seeding density variability reduction approach combining an on-line cell density probe and predictive modeling, *IBC's 3rd Intl.on Implementing PAT for Bioprocessing*, Reston, VA.

Undey, C. (2008). Are we there yet? An Industrial Perspective of Evolution from Post-mortem Data Analysis towards Real-time Multivariate Monitoring and Control of Biologics Manufacturing Processes, *IBC's BioProcess International Analytical and Quality Summit*, Cambridge, MA, June 2-4.

US Food and Drug Administration (August 2002). *Pharmaceutical cGMPs for the 21st century: A risk-based approach (FDA, Rockville, MD).* [Web] [cited 2009 01/06/2009]; Available from: http://www.fda.gov/oc/guidance/gmp.html.

US Food and Drug Administration (September, 2004). *Guidance for Industry: PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance* (FDA, Rockville, MD). [Web] [cited 2009 01/06/2009]; Available from: http://www.fda.gov/Cder/guidance/6419fnl.pdf.

US Food and Drug Administration (September 2006). *Guidance for industry: Q10 quality systems approach to pharmaceutical cGMP regulations* (FDA, Rockville, MD). [Web] [cited 2009 01/06/2009]; Available from: http://www.fda.gov/cber/gdlns/qualsystem.pdf.

Wold, S., Kettaneh, N., Fridén, H., and Holmberg, A. (1998). Modelling and diagnostics of batch processes and analogous kinetic experiments, *Chemometrics and Intelligent Laboratory Systems*, 44, 331-340.

Cascade Hybrid Control for Anaerobic Digestion Systems ⋆

J.P. García-Sandoval *,

* *Chemical Engineering Department, University of Guadalajara, Calz. Gral. Marcelino García Barragán 1451, Guadalajara, Jalisco 44430, México (e-mail: paulo.garcia@cucei.udg.mx)*

Abstract: A cascade hybrid controller is proposed to guarantee the operational stability of anaerobic digestion systems and at the same time to assure a satisfactory remotion. The effluent chemical oxygen demand (COD) concentration and the dilution rate are taken respectively as the regulated and the manipulated variables. This controller is composed by a continuous inner loop which controls the concentration of volatile fatty acids and updates its reference at each sampling period according to the available discrete information of the COD, enhancing the robustness of the closed-loop with respect of influent disturbances. The performance of the proposed control scheme is illustrated via numerical simulations and compared with a discrete controller which only uses the COD information.

Keywords: Cascade control, anaerobic digestion, continuous and discrete measurements

1. INTRODUCTION

The stringent regulation policies imposed in chemical and biological processes has brought abut the need of more efficient monitoring, decision and control systems. Nowadays, it is not enough to regulate readily available variables such as pH or temperature, to guarantee both product quality and process safety (Alcaraz-González and González-Álvarez, 2007). In the particular case of anaerobic digestion, one must pay attention to certain operating conditions that may lead the system to the eventual crash even under tightly controlled pH and temperature conditions (Dochain, 2008). The last two decades have seen an increasing interest to improve the operation of bioprocesses by applying advanced control schemes. In particular, Anaerobic digestion (AD) has regained the interest of the wastewater treatment scientific and industrial community to reduce and transform the organic matter from industrial and municipal effluents into a high-energy gas (Henze et al., 1997). Nevertheless, its widespread application has been limited, because of the difficulties involved in achieving the stable operation of the AD process, which cannot be guaranteed by regulating temperature and pH, because the microbial community within the AD process is quite complex (Méndez-Acosta et al., 2008). In addition, the behavior of such a process may be affected by the substrate composition and inhibition by substrates or products, for example the accumulation of volatile fatty acids (VFA). Moreover, it is well-known that, to guarantee the so-called operational stability (Hill et al., 1987) and to avoid the eventual crash of the anaerobic digester, the organic matter in the liquid phase must be kept in a set of predetermined values.

Over the past decade, the regulation of the organic matter has been addressed by proposing many control techniques to keep certain operating variables which are readily available (such as the chemical oxygen demand (COD) and the biogas production) at a predetermined value (Schügerl, 2001; Alvarez-Ramírez et al., 2002; Puñal et al., 2002; Ahring and Angelidaki, 1997). For example, Alvarez-Ramírez et al. (2002) designed a continuous PI feedback controller to regulate COD concentration which uses the VFA as a secondary measurement that is incorporated into the feedback loop scheme to enhance the robustness of the control scheme with respect of influent disturbances. Nevertheless, the problem of the operational instability due to the accumulation of volatile fatty acids remains open (Méndez-Acosta et al., 2008). In this context, the regulation of the VFA concentration as a controlled variable seems to be very promising, because the operational stability of the AD process is largely dependent on the accumulation of VFA. For instance, some authors recommend a VFA concentration below 1.5 g/l (25 mmol/l) (Angelidaki et al., 2004). For this reason, it is necessary to design multiobjective controls which fulfill environmental regulations about COD effluents and at the same time guarantee the operational stability.

In this work we propose the regulation of COD concentration using a cascade scheme which assures the robustness and the operational stability via VFA continuous regulation, where the VFA reference is updated at each sampling period according to the available discrete information of the COD, in contrast with the continuous information used by Alvarez-Ramírez et al. (2002). The proposed scheme is a cascade hybrid control since is composed by a continuous and a discrete part. The paper is organized as follows. In section 2 the AD model is presented. Then, in section 3 the controller is designed and in section 4 it is implemented

⋆ Supported by PROMEP under proyect 103.5/08/2919 and CONACYT under grant 43658.

and tested via numerical simulations. Finally, the paper is closed with some concluding remarks.

2. MODEL PRESENTATION

There are hundreds of dynamical models available to describe anaerobic digestion, from the basic ones considering only one biomass (Andrews, 1968) to detailed models including several bacterial populations and several substrates. Among the complex models available in the literature, the IWA Anaerobic Digestion Model 1 (Batstone et al., 2002) has imposed itself as an useful tool to describe the behavior of a digestion plant with more insight into the process dynamics. However, its excessive complexity makes any advanced mathematical analysis of the model critical (Hess and Bernard, 2008). Thus it is considered a simplified macroscopic model of the anaerobic process based on 2 main reactions (Bernard et al., 2001), where the organic substrate (S_1) is degraded into volatile fatty acids (S_2) by acidogenic bacteria (X_1), and then the VFA are degraded into methane CH_4 and CO_2 by methanogenic bacteria (X_2). A mass balance model in a continuous stirred tank reactor (CSTR) can straightforwardly be derived (Bastin and Dochain, 1990; Bernard et al., 2001):

$$\dot{X}_1 = \mu_1(S_1) X_1 - \alpha D X_1 \tag{1a}$$
$$\dot{S}_1 = -k_1 \mu_1(S_1) X_1 + D(S_{1in} - S_1) \tag{1b}$$
$$\dot{X}_2 = \mu_2(S_2) X_2 - \alpha D X_2 \tag{1c}$$
$$\dot{S}_2 = k_2 \mu_1(S_1) X_1 - k_3 \mu_2(S_2) X_2 + D(S_{2in} - S_2) \tag{1d}$$

where D is the dilution rate, S_{1in} and S_{2in} are respectively, the concentrations of influent organic substrate and of influent VFA. The k_is are pseudo-stoichiometric coefficients associated to the bioreactions. Parameter $\alpha \in (0, 1]$ represents the fraction of the biomass which is not retained in the digester. The specific bacterial growth rates, $\mu_1(S_1)$ and $\mu_2(S_2)$, are given by the nonlinear equations (Dochain and Vanrolleghem, 2001; Van-Impe et al., 1998):

$$\mu_1(S_1) = \mu_{\max 1} \frac{S_1}{S_1 + K_{S1}}$$
$$\mu_2(S_2) = \mu_{\max 2} \frac{S_2}{S_2 + K_{S2} + (S_2/K_{I2})^2}$$

where $\mu_{1\max}$, K_{S1}, $\mu_{2\max}$, K_{S2} and K_{I2} are the maximum bacterial growth rate and the half-saturation constant associated to the substrate S_1, the maximum bacterial growth rate in the absence of inhibition, and the saturation and inhibition constants associated to substrate S_2, respectively.

Hess and Bernard (2008) presented an extensive steady state analysis for system (1). They have found that, for a given dilution rate, the VFA concentration have three equilibrium points mathematically stables, but two of them are operationally unstable since one drives to substrate inhibition which can produce methanogenic bacterial extinction, which is the second operational unstable equilibrium. For this reason volatile fatty acids must be controlled, but at the same time, it should be guaranteed the COD remotion. In the next section we propose a control algorithm to achieve these goals.

3. CONTROL DESIGN

3.1 Problem formulation

Consider a nonlinear system describe by

$$\dot{x}(t) = f(x(t), u(t), \mu) \tag{2}$$
$$y_1(t) = C_1 x(t) \tag{3}$$
$$y_2(k\delta) = C_2 x(k\delta) \tag{4}$$

where $x \in \mathbb{R}^n$ represents the state vector, $u \in \mathbb{R}$ describes the input vector, $y_1 \in \mathbb{R}$ is a continuous measurement, while $y_2 \in \mathbb{R}$ is a discrete measurement obtained at each sampling period δ. Finally, $\mu \in \mathbb{R}^p$ is a parameter vector which may take values in a neighborhood $\mathcal{P} \in \mathbb{R}^p$ of the nominal ones, μ_0.

Given a constant value y_{2r}, it is desirable to regulate the output y_2 around y_{2r}, then, it can be defined the regulation error

$$e_2 = y_2 - y_{2r} \tag{5}$$

and the proposed *Regulation Problem* consists in finding, if possible, a controller which guarantees that $\lim_{t\to\infty} e_2(t) = 0$. However this controller must take into account both, continuous and discrete outputs (3)-(4) in order to guarantee the stability of the closed-loop system. The next assumption about the steady state around $e_2 = 0$ is instrumental in the controller design.

Assumption 1. Given the constant reference y_{2r} there exists at least one linear vector $x_{ss}(\mu)$ and one scalar $u_{ss}(\mu)$ such that equations

$$0 = f(x_{ss}, u_{ss}, \mu), \tag{6}$$
$$0 = C_2 x_{ss} - y_{2r}, \tag{7}$$

hold. Additionally, the scalar y_{1ss} is defined such that

$$0 = C_1 x_{ss} - y_{1ss}. \tag{8}$$

Remark 2. x_{ss} represents the steady state vector, u_{ss} is the input necessary to achieve this steady state, while y_{1ss} denotes the constant value which reaches the output, y_1. Notice that, by using the central manifold theory (Isidori, 1995), it is evident that x_{ss}, u_{ss} and y_{1ss} will depend on both y_{2r} and the unknown parameter vector, μ. Additionally, assumption 1 can be reformulated in order to fix the value of y_{1ss} and find the solution of (6), (7) and (8) for x_{ss}, u_{ss} and y_{2r}. Finally, if for a given y_{2r}, the solution of (6) and (7) is not unique, neither is y_{1ss}.

In order to design a controller which solves the proposed regulation problem, some basic concepts will be presented in the next subsection.

3.2 Basic facts about jump observers

Consider the linear system

$$\dot{x}(t) = Ax(t) + Bu(t) \qquad \forall t \in [0, \infty) \tag{9}$$
$$y(k\delta) = Cx(k\delta) \qquad k = 1, 2, 3, \ldots, \tag{10}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $y \in \mathbb{R}^q$ are the state, input and output vectors, respectively. In this case the outputs are obtained at each sampling time δ.

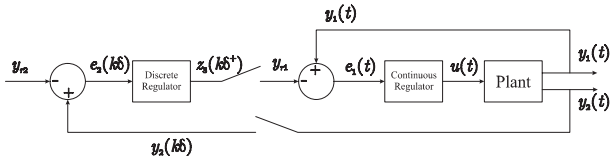


Fig. 1. Cascade proposed control scheme.

The usual way to estimate the unknown states of system (9) from output (10) consists in discretizing and designing a discrete observer. However, the observer thus obtained only provides information at each sampling period. Additionally, to obtain a discrete version of (9) it is necessary to have a well defined input in order to place the appropriate holder (for example a zero holder or a exponential holder), hence unexpected input variations during intersampling periods may produce a failure in the discrete observer (García-Sandoval, 2006). For this reason, an interesting problem would be to construct an observer of the form

$$\dot{z}(t) = Az(t) + Bu(t) \qquad \forall t \neq k\delta \tag{11}$$

$$z\left(k\delta^+\right) = z(k\delta) - G\left[y(k\delta) - Cz(k\delta)\right] \quad t = k\delta \tag{12}$$

where $z \in \mathbb{R}^n$ are the observer states and $z\left(k\delta^+\right)$ denotes the updated observer states at each sampling instant, i.e this is a continuous observer which updates its states at each sampling instant. The next lemma establishes conditions for the existence of such observer.

Lemma 3. Consider system (9)-(10) and suppose the pair $\left(e^{A\delta}, C\right)$ is observable, then an observer of the form (11)-(12), where the matrix gain G is such that matrix $(I + GC)\, e^{A\delta}$ is Schur, guarantees that $\lim_{t\to\infty}\left[x(t) - z(t)\right] = 0$.

Proof. See Appendix.

Remark 4. The main feature of observer (11)-(12) is that the intersampling state information is available at any time and it is not necessary to have a pre-established dynamic behavior for the input. Equation (11) can be seen as an continuous open loop observer in the intersampling period, whose states, according to (12), are updated at the sampling instant.

Observer (11)-(12) can be used to design a cascade controller which uses the continuous information in the inner loop and the discrete information in the external loop, as presented in the next subsection.

3.3 Proposed controller

In order to solve the regulation problem, it is proposed to design a cascade control scheme where the inner loop consists of a continuous controller which uses the input $u(t)$ to regulate the continuous output, y_1, around y_{1r} (notice that y_{1r} may not necessarily be equal to y_{1ss}, since y_{1ss} is initially unknown). Then, by using a jump observer, it is possible to devise an external loop controller which estimates the reference y_{1r} necessary to regulate the discrete output, y_2, around y_{2r}. The scheme is illustrated in Figure 1.

The next Theorem presents the proposed solution to the regulation problem.

Theorem 5. Consider assumption 1 holds. Assume the pairs

$$[A_0, B_0] \quad \text{and} \quad [A, C_{T1}]$$

with

$$A = \begin{pmatrix} A_0 & -B_0 \\ 0 & 0 \end{pmatrix}, \quad C_{T1} = (C_1 \;\; 0), \tag{13}$$

$$A_0 = \left[\frac{\partial f}{\partial x}\right]_{(x,u,\mu)=(0,0,0)} \quad \text{and} \quad B_0 = \left[\frac{\partial f}{\partial u}\right]_{(x,u,\mu)=(0,0,0)}$$

are stabilizable and detectable, respectively. Then a controller which solves the regulation problem for system (2) using outputs (3) and (4) is

$$\dot{z}_1(t) = (A_0 - G_1C_{01})\, z_1(t) - B_0 z_2(t) + B_0 u(t) + G_1 e_1(t), \qquad \forall t \neq k\delta, \tag{14a}$$

$$\dot{z}_2(t) = -G_2C_{01}z_1(t) + G_2 e_1(t), \quad \forall t \neq k\delta, \tag{14b}$$

$$z_3\left(k\delta^+\right) = G_dC_{02}z_1(k\delta) + z_3(k\delta) - G_d e_2(k\delta), \qquad \begin{matrix} \forall t = k\delta, \\ k = 1, 2, 3, \ldots \end{matrix} \tag{14c}$$

$$u(t) = Kz_1(t) + z_2(t) \tag{14d}$$

with

$$e_1(t) = y_1(t) - z_3\left(k\delta^+\right), \tag{15}$$

$$e_2(k\delta) = y_2(k\delta) - y_{2r}, \tag{16}$$

where K and $G = (G_1, G_2)^T$ are such that matrices $(A_0 + B_0K)$ and $(A - GC_{T1})$ are Hurwitz, while given matrices

$$\overline{A}_d = \begin{pmatrix} A_d & -M_d \\ 0 & 1 \end{pmatrix} \tag{17}$$

with

$$A_d = e^{(A - GC_{1T})\delta} \quad \text{and} \quad M_d = \int_0^\delta e^{(A-GC_{1T})\lambda} G d\lambda, \tag{18}$$

the scalar G_d is such that $\left(I + \overline{G}_d\overline{C}_2\right)\overline{A}_d$ is Schur, where

$$\overline{G}_d = \begin{pmatrix} 0 \\ G_d \end{pmatrix}, \quad \overline{C}_2 = (C_{T2} \;\; 0) \quad \text{and} \quad C_{T2} = (C_2 \;\; 0), \tag{19}$$

obviously the pair $\left(\overline{A}_d, \overline{C}_2\right)$ must be detectable.

Proof. Given a reference y_{2r}, if the solution of (6), (7) and (8) is x_{ss}, u_{ss} and y_{1ss}, the linear version of system (2) around the pair $(x, u) = (x_{ss}, u_{ss})$ and the nominal values μ_0 is

$$\dot{\underline{x}}(t) = A_0\underline{x}(t) + B_0u(t) - B_0u_{ss} + \widehat{f}\left(\underline{x}(t), u(t), \mu\right) \tag{20a}$$

$$y_1(t) = C_1\underline{x}(t) + y_{1ss} \tag{20b}$$

where $\underline{x} = x - x_{ss}$, while

$$A_0 = \left.\frac{\partial f}{\partial x}\right|_{\substack{x=x_{ss},\\ u=u_{ss}}}, \quad B_0 = \left.\frac{\partial f}{\partial u}\right|_{\substack{x=x_{ss},\\ u=u_{ss}}} \quad \text{and} \quad C_{01} = \left.\frac{\partial h_1}{\partial x}\right|_{x=x_{ss}},$$

are the linear approximations and $\widehat{f}(\underline{x}, u, \mu)$ contains the second or higher order terms. Since z_3 remains constant along each sampling instant, it can be considered that $\dot{z}_3 = 0\ \forall t \neq k\delta$. Defining

$$\xi_1 = \underline{x} - z_1$$

$$\xi_2 = u_{ss} - z_2$$

$$\xi_3 = y_{1ss} - z_3$$

the linear approximation of the closed-loop system can be written as

$$\dot{\xi}(t) = \overline{A}\xi(t), \qquad \forall t \neq k\delta \tag{21}$$

$$\xi\left(k\delta^{+}\right) = \left(I + \overline{G}_d\overline{C}_2\right)\xi(k\delta), \quad \forall t = k\delta,\ k = 1, 2, .. \tag{22}$$

where

$$\overline{A} = \begin{pmatrix} A - GC_{T1} & -G \\ 0 & 0 \end{pmatrix}$$

while A and C_{T1} are given in (13). Notice that $\underline{x}(k\delta) = \underline{x}\left(k\delta^{+}\right)$. Integrating (21) for $t \in \left(k\delta^{+}, (k+1)\delta\right)$,

$$\xi\left((k+1)\delta\right) = \overline{A}_d\xi\left(k\delta^{+}\right), \tag{23}$$

where $\overline{A}_d$ is described in (17), while A_d and M_d are defined in (18). Taking an increment in (22) and replacing (23),

$$\xi\left((k+1)\delta^{+}\right) = \left(I + \overline{G}_d\overline{C}_2\right)\overline{A}_d\xi\left(k\delta^{+}\right).$$

Since by hypothesis $(A - GC_{T1})$ is Hurwitz, then A_d is Schur and $\overline{A}_d$ have only one eigenvalue in the unitary circle. However, if G_d is such that $\left(I + \overline{G}_d\overline{C}_2\right)\overline{A}_d$ is Schur, it is guaranteed that $\xi\left(k\delta^{+}\right)$ is asymptotically stable. Therefore $\lim_{k\to\infty} z_3(k\delta) = y_{1ss}$, $\lim_{k\to\infty} z_2(k\delta) = u_{ss}$ and $\lim_{k\to\infty} z_1(k\delta) = \underline{x}(k\delta)$. Then by continuity $\lim_{t\to\infty} z_2(t) = u_{ss}$ and $\lim_{t\to\infty} z_1(t) = \underline{x}(t)$ and the closed-loop dynamics of (20) approach to

$$\dot{\underline{x}}(t) = (A_0 + B_0K)\underline{x}(t) + \widehat{f}\left(\underline{x}(t), K\underline{x}(t) + u_{ss}, \mu\right)$$

which is asymptotically stable in a neighborhood of the origin since $(A_0 + B_0K)$ is Hurwitz. Then $\lim_{t\to\infty}[x(t) - x_{ss}] = 0$ and this guarantees that $\lim_{t\to\infty}[y(t) - y_{2r}] = 0$, concluding the proof.

Remark 6. z_1 and z_2 represent the dynamic of the continuous regulator, while z_3 is the state of the discrete regulator. Notice that z_2 which is present in the input (14d) can be considered as an integral action since integrating (14b), which is precisely the dynamic of z_2, appears an integral of $e_1(t)$. Analogously, z_3 is also an discrete integral action which consider $e_2(k\delta)$. These variables give the robustness to the cascade controller.

4. CONTROL IMPLEMENTATION

Considering that the VFA concentration is measured continuously, while the COD concentration is measured at each sampling period δ, the outputs of system (1) are

$$y_1(t) = C_1x(t) \qquad \text{and} \qquad y_2(k\delta) = C_2x(k\delta),$$

where $x = \mathrm{col}(X_1, S_1, X_2, S_2)$ and

$$C_1 = (0\ 0\ 0\ 1) \qquad , \qquad C_2 = (0\ 1\ 0\ 0).$$

System (1) is similar to (2), hence controller (14) presented in Theorem 5 can be designed. The linear matrix approximations of system (1) are

$$A_0 = \mu_1(S_{1ss}) \begin{pmatrix} 0 & \Psi_1 & 0 & 0 \\ -k_1 & -k_1\Psi_1 - \dfrac{1}{\alpha} & 0 & 0 \\ 0 & 0 & 0 & \Psi_2 \\ 0 & k_2\Psi_1 & -k_3\Psi_2 & -k_3\Psi_2 - \dfrac{1}{\alpha} \end{pmatrix},$$

$$B_0 = \left(-\alpha X_{1ss}\ \ (S_{1in} - S_{1ss})\ \ -\alpha X_{2ss}\ \ (S_{2in} - S_{2ss})\right)^T,$$

where $\Psi_1 = \frac{K_{S1}}{\mu_{\max 1}}\frac{\mu_1(S_{1ss})}{S_{1ss}^2}X_{1ss}$, $\Psi_2 = \frac{\mu_2(S_{2ss})}{S_{2ss}^2}X_{2ss}\cdot\left[\frac{K_{S2} - (S_{2ss}/K_{I2})^2}{\mu_{\max 2}}\right]$, while the steady state is described by

Table 1. Nominal parameter values and variations.

Parameter	Nominal value	% of variation from nominal value		
		$t < 10$	$10 \leq t < 30$	$t \geq 30$
$\mu_{\max 1}$	$1.2\,\mathrm{d}^{-1}$	−10	−20	−20
$\mu_{\max 2}$	$0.744\,\mathrm{d}^{-1}$	−7	−10	−10
K_{S1}	7.1 g/l	−5	10	10
K_{S2}	9.28 mmol/l	10	10	10
K_{I2}	16 mmol/l	−15	−15	−15
k_1	42.14 g/g	−12	−12	−12
k_2	116.5 mmol/g	15	15	15
k_3	268 mmol/g	13	13	13
α	0.5	0	15	15
S_{1i}	30 g/l	25	10	10
S_{2i}	750 mmol/l	−5	−5	50

$$\alpha D_{ss} = \mu_1(S_{1ss}) = \mu_2(S_{2ss}),$$

$$X_{1ss} = \alpha^{-1}k_1^{-1}(S_{1in} - S_{1ss}),$$

$$X_{2ss} = \alpha^{-1}k_3^{-1}\left[k_1^{-1}k_2(S_{1in} - S_{1ss}) + (S_{2in} - S_{2ss})\right],$$

where it must be fixed one variable (which might be any of this: D_{ss}, S_{1ss} or S_{2ss}) in order to estimate the others.

In order to analyze the closed-loop dynamic behavior, the parameters reported in Table 1 where taken from (Bernard et al., 2001). Following Theorem 5 and considering a sample period of 1day, the next matrices were obtained

$$A_0 = \begin{pmatrix} 0 & 0.1207 & 0 & 0 \\ -13.1687 & -5.7096 & 0 & 0 \\ 0 & 0 & 0 & 0.1554 \\ 36.4063 & 14.0570 & -83.7500 & -42.2704 \end{pmatrix},$$

$$B_0 = (-0.6526\ \ 27.5000\ \ -3.0566\ \ 743.1464)^T,$$

while using LQR techniques the next feedback parameters were obtained

$$K = (-0.0007\ \ -0.0044\ \ -0.0162\ \ -0.0083),$$

$$G = (-0.1132\ \ 5.3922\ \ -0.8990\ \ 88.2809\ \ -10.0000)^T,$$

$$G_d = 2.5.$$

Since it is desirable to guarantee the operational stability, it was set that VFA reference must be bounded by the next domain: $y_{1r} \in (0, 12)$ mmol/l, while $D \in (0.05, 1)$ d^{-1}.

4.1 Simulation results

To verify if the proposed controller enhances the robustness of the closed-loop with respect of influent disturbances, a discrete controller was designed with the same sampling period to regulate the COD concentration. Figure 2 presents the dynamic response for both controllers. The simulation was carried out using steady state initial conditions and nominal parameter values (see Table 1). At time $t = 5$ d a step in the reference from 2.5 to 2 COD g/l was induced. As can be seen in Figure 2a both controllers can handle this step approximately with the same dynamic response, however, at time $t = 15$ d a drastic overload in the VFA influent concentration was induced. Since this overload does not affect the acidogenesis phase, the discrete controller remains unaltered as well as its COD concentration, while the cascade controller detects a change in the VFA concentration and modify the dilution

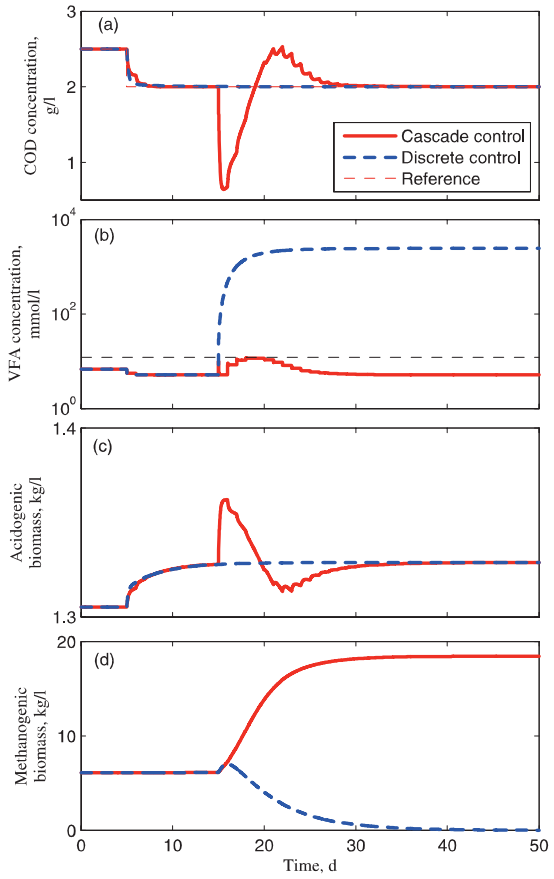


Fig. 2. Comparison of the cascade control with a traditional COD regulator. (a) COD concentration. (b) VFA concentration. (c) Acidogenic biomass. (d) Methanogenic biomass.

rate (see Figure 3). This induced a variation on the COD concentration which is corrected by the cascade controller in approximately 10 days. As can be seen in Figure 2b the VFA concentration increases in both cases, however, since for the cascade controller y_{1r} is bonded, at approximately $t = 25\,\mathrm{d}$, this reference reached the upper bound and the inner loop did not allow it to keep increasing, while for the discrete controller it keeps increasing, producing an acidification and consequently, a substrate inhibition which finally causes the methanogenic bacterial death, as shown in Figure 2d. This simulation remarks the capabilty of the cascade controller to guarantee the operational stability at the expense of a small temporary variation on the COD concentration.

Another simulation for the cascade controller was carried out, in order to verify the robustness to parametric and load variations as reported in Table 1, notice that variations up to 50% were induced. Figures 4 and 5 present the dynamic response. It is evident the robustness of the closed-loop with respect to load disturbances and several parametric variations.

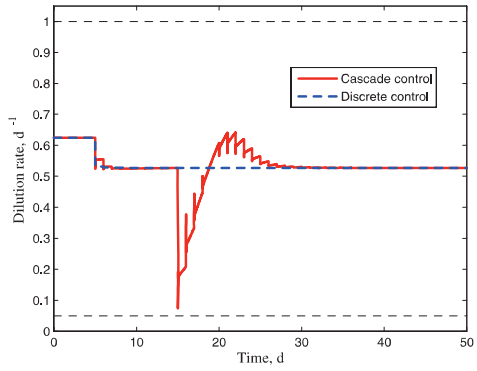


Fig. 3. Dilution rate for the cascade control and a traditional COD regulation.

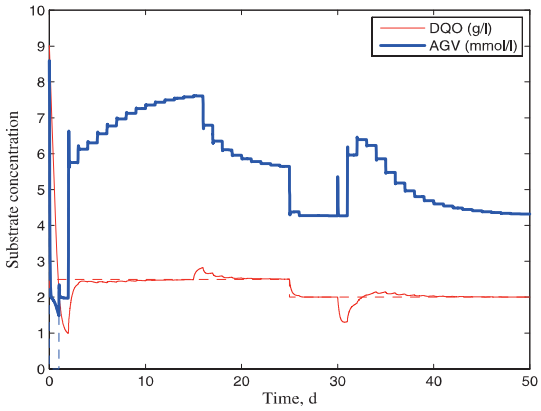


Fig. 4. Substrate concentrations.

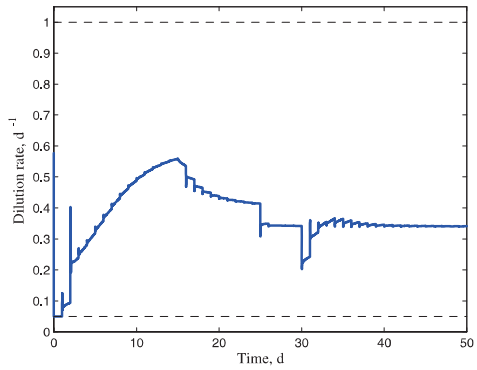


Fig. 5. Dilution rate.

5. CONCLUSION

A cascade scheme was proposed to regulate the COD concentration of AD processes. Although the control law is based on a linear approximation its robustness is conferred by integral actions of the controller. The cascade scheme also presents advantages since, additionally to the COD regulation, can be used to guarantee the operational stability, i.e. to avoid the AGV inhibition, at the expense of a small temporary variation on the COD concentration. Real time implementation are been carried out and results

will appear soon. Although the design of the control law (14) was motivated by the control of AD systems, it can be applied to another systems. As future work this theory will be extended to the case of not constant references.

(Chapter head:)*

Bibliography

Ahring, B. and Angelidaki, I. (1997). Monitoring and controlling the biogas process. In *Proceedings of the 8th International Conference on Anaerobic Digestion*, volume 1.

Alcaraz-González, V. and González-Álvarez, V. (2007). *Selected Topics in Dynamics and Control of Chemical and Biological Processes*, volume 361 of *Lecture Notes in Control and Information Sciences*, chapter Robust Nonlinear Observers for Bioprocesses Application to Wastewater Treatment, 119–164. Springer, Germany.

Alvarez-Ramírez, J., Meraz, M., Monroy, M., and Velasco, A. (2002). Feedback control design for an anaerobic digestor process. *J. Chem. Technol. & Biotechnol.*, 77(6), 725–734.

Andrews, J. (1968). A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates. *Biotechnology and Bioengineering*, 10, 707–723.

Angelidaki, I., Boe, K., and Ellegaard, L. (2004). Efect of operating conditions and reactor configuration on effciency of full-scale biogas plants. In *10th World Congress of Anaerobic Digestion*, 275–280. Canada.

Bastin, G. and Dochain, D. (1990). *On-Line Estimation and Adaptive Control of Bioreactors*. Elsevier.

Batstone, D., Keller, J., Angelidaki, I., Kalyuzhnyi, S., Pavlostathis, S., Rozzi, A., Sanders, W., Siegrist, H., and Vavilin, V. (2002). *Anaerobic Digestion Model No. 1 (ADM1)*, volume 13 of *Scientific and Technical Report*. IWA Publishing, London.

Bernard, O., Hadj-Sadok, Z., Dochain, D., Genovesi, A., and Steyer, J. (2001). Dynamical model development and parameter identification for anaerobic wastewater treatment process. *Biotechnology & Bioengineering*, 75(4), 424–439.

Dochain, D. and Vanrolleghem, P. (2001). *Dynamical Modelling and Estimation in Wastewater Treatment Processes*. IWA Publishing, Colchester.

Dochain, D. (2008). *Bioprocess Control*. Control Systems, Robotics and Manufacturing. Wiley.

García-Sandoval, J. (2006). *The Robust Regulation Problem Using Immersions: Reactors Applications*. Ph.D. thesis, CINVESTAV.

Henze, M., Harremoes, P., LA-Cour-Jansen, J., and Arvin, E. (1997). *Wastewater Treatment: Biological and Chemical Processes*. Springer Verlag, Berlin, 2nd edition.

Hess, J. and Bernard, O. (2008). Design and study of a risk management criterion for an unstable anaerobic wastewater treatment process. *Journal of Process Control*, 18(1), 71–79.

Hill, D., Cobbs, S., and Bolte, J. (1987). Using volatile fatty acid relationships to predict anaerobic digester failure. *Trans. ASAE*, 30(2), 496–501.

Isidori, A. (1995). *Nonlinear Control Systems*. Springer, London, 3rd edition.

Kailath, T. (1980). *Linear Systems*. Prentice Hall.

Méndez-Acosta, H., Palacios-Ruiz, B., Alcaraz-González, V., Steyer, J., González-Álvarez, V., and Latrille, E. (2008). Robust control of volatile fatty acids in a anaerobic digesters. *Industrial and Engineering Chemical Research*, 47(20), 7715–7720.

Puñal, A., Palazzotto, L., Bouvier, J., Conte, T., and Steyer, J. (2002). Automatic control of VFA in anaerobic digestion using a fuzzy logic based approach. In *IWA VII Latin America Workshop and Symposium on Anaerobic Digestion*, 126–133.

Schügerl, K. (2001). Progress in monitoring, modeling and control of bioprocesses during the last 20 years. *J. Biotechnol.*, 85(2), 149–173.

Van-Impe, J., Vanrolleghem, P., and Iserentant, D. (1998). *Advanced Instrumentation, Data Interpretation, and Control of Biotechnological Processes*. Kluwer Academic Publishers.

Appendix A. OUTLINE OF PROOF

Proof. [Lemma 3] Let us define

$$\xi(t) = x(t) - z(t), \quad \text{and } \xi\left(k\delta^+\right) = x(k\delta) - z\left(k\delta^+\right),$$

where $\xi(t)$ represents the continuous error and $\xi\left(k\delta^+\right)$ is the updated error for each sampling period. Note that $x\left(k\delta^+\right) = x(k\delta)$ since system (9) is continuous. Now

$$\dot{\xi}(t) = A\xi(t) \qquad \forall t \neq k\delta \tag{A.1}$$

$$\xi\left(k\delta^+\right) = (I + GC)\,\xi(k\delta) \qquad t = k\delta. \tag{A.2}$$

Solving (A.1) for $t \in \left[k\delta^+, (k+1)\,\delta\right]$, it follows that

$$\xi(k+1) = A_d \xi\left(k\delta^+\right), \tag{A.3}$$

where $A_d = e^{A\delta}$. From (A.2) and (A.3) it is obtained

$$\xi\left((k+1)\,\delta^+\right) = (I + GC)\,\xi(k+1) = (I + GC)\,A_d \xi\left(k\delta^+\right),$$

and thus, if the pair (A_d, CA_d) is observable, then a matrix G can be calculated such that $A_d + GCA_d$ is Schur and the error $\xi\left(k\delta^+\right)$ will converge to zero, hence $\lim_{k\to\infty}\left[x(k\delta) - z\left(k\delta^+\right)\right] = 0$; then for $k\delta < t \leq (k+1)\,\delta$ the solution $z(t)$ converges to $x(t)$, that is $\lim_{t\to\infty}\left[x(t) - z(t)\right] = 0$. On the other hand, to prove that the pair (A_d, CA_d) is observable if the pair (A_d, C) is observable, consider its observability matrix

$$\mathcal{O} = \begin{pmatrix} CA_d \\ CA_d^2 \\ \vdots \\ CA_d^n \end{pmatrix},$$

where $A_d \in \mathbb{R}^{n\times n}$, then using the Hamilton-Cailey theorem (Kailath, 1980)

$$A_d^n = a_0 I + a_1 A_d + \cdots + a_{n-1} A_d^{n-1},$$

the observability matrix becomes

$$\mathcal{O} = \begin{pmatrix} CA_d \\ CA_d^2 \\ \vdots \\ a_0 C + a_1 CA_d + \cdots + a_{n-1} CA_d^{n-1} \end{pmatrix}.$$

Since A_d is obtained through a discretization of matrix A then $a_0 \neq 0$ and $\mathcal{O}$ has full rank if the pair (A_d, C) is observable.

Analysis and Control of Crystallization Processes

Oral Session

A Stochastic Approach for Anti-Solvent Addition Policy in Crystallization Operations: An Application to a Bench-Scale Fed-Batch Crystallizer

Omar GALAN*, Massimiliano GROSSO, Roberto BARATTI**, José ROMAGNOLI***

** Department of Chemical Engineering, Louisiana State University, USA (e-mail: galan, jose@lsu.edu)*

*** Dipartimento di Ingegneria Chimica e Materiali, Università di Cagliari, Cagliari, Italy (e-mail: grosso, baratti@dicm.unica.it)*

Abstract: This work aims a stochastic approach for the calculation of robust anti-solvent addition policies for controlling the mean crystal size (MCS) in fed-batch crystallization operations. The proposed strategy is based-on a non-structured population balance where uncertainties associated with the start-up condition and random fluctuations along the fed-batch operation can be taken into account in a very natural fashion. We include and quantify the effect of the uncertainties by embedding a deterministic crystal growth model into a Fokker-Planck equation (FPE) resulting in a stochastic model for the MCS dynamics. This approach uses the Generalized Logistic equation (GLE) that has an adequate mathematical structure that suits the dynamic characteristic of the crystal growth. Thus, the numerical solution of the FPE provides the most likely MCS evolution for a given anti-solvent flow-rate. The effect of the anti-solvent is incorporated into the parameters of the FPE. The parameters of the FPE are computed as linear piece-wise interpolating functions of the anti-solvent flow-rate. The strategy uses a PID-like regulator in closed-loop fashion with the FPE to compute the anti-solvent addition flow-rates for different set-point targets in the MCS. In order to validate the stochastic model and assess the merits of the proposed strategy, the crystallization of sodium chloride in water using ethanol as anti-solvent is performed in a bench-scale fed-batch crystallizer. The implementation of the calculated anti-solvent policies resulted in a good control of the MCS despite modelling mismatch and uncertainties present during the crystallization operation.

Keywords: Anti-Solvent; Crystallization; Fokker-Planck Equation; Mean Crystal Size; Stochastic.

1. INTRODUCTION

The design of chemical plants endeavors to build equipment that preferably content hazards and make possible the transformation and separation of materials. It also attempts to harness the impact of apparently disordered and erratic phenomena (*e.g.* turbulent flow, pressure and temperature fluctuations, measurement noise, etc.). Fluctuations are a very common element in a large number of chemical, biological and physical phenomena. Practically, all systems are subjected to complicated external or internal influences that are not fully known and that are often termed noise or fluctuations. However, if a sufficiently long record of noisy measurement is analyzed, it may admit a statistical description. This means that it is possible to estimate the probability or likelihood that the process variable will attain in some specified range of values (Feigenbaum, 1980; Risken, 1984).

The study of stochastic system as the Brownian motion resulted in the Fokker-Planck equation (FPE). The FPE is just an equation of motion for the distribution function of fluctuating macroscopic variables. The FPE deals with those fluctuations of systems which stem from many tiny disturbances, each of which changes the variables of the system in an unpredictable but small way. The FPE provides a powerful tool with which the effects of fluctuations close to transition points can be adequately treated and that the approaches based on FPE are superior to other approaches based on Langevin equations (LE). The FPE plays an important role in chemical and biological processes that involve noise.

For many practical applications it is required to have simplified models that group the complexity behind a natural phenomenon and its interactions with its surroundings. For a dynamic system, it means of a set of deterministic differential equations with semi-empirical parameters. When studying chemical processes, these models are the core element for the design of all model-based control and optimization strategies. However, extra care is needed to take into account the no modeled dynamics and unknown exogenous disturbances acting on the process. The FPE is an interesting approach to introduce the robustness feature to the design of prediction, control and optimization tools.

This work describes a novel stochastic approach for the robust prediction of the mean crystal size (MCS) in a bench-scale fed-batch crystallization unit where anti-solvent is added to speed-up the crystal formation process. The crystal growth is modeled by a classic logistic equation of common use in theoretical ecology (May and McLean, 2007; Grosso et al., 2007). In a different fashion, the use of FPE for a monomer particle growth can be found in the literature (Matsoukas and Yulan, 2006). Unknown dynamics, internal and external fluctuations and sensitivity to initial conditions can be taken into account by embedding the logistic equation in the FPE.

2. Mean Crystal Size Estimation for an Anti-Solvent Aided Crystallization Process

Crystallization is a physical process for solid-liquid separation where the solid (solute) is dissolved in the solvent (liquid). The driving force in crystal formation is the super-saturation. The super-saturation condition establishes the thermodynamic equilibrium for the solid-liquid separation and it can be affected by cooling and evaporation. The super-saturation can be also induced by addition of precipitant or anti-solvent to the solution. The anti-solvent reduces the solubility of the solute in the original solvent resulting in super-saturation. The anti-solvent aided crystallization is an advantageous technique of separation where the solute is highly soluble or heat sensitive.

2.1 Mathematical Model

The development of rigorous mathematical models describing the dynamic of crystal growth in crystallization processes are based-on population balances. The idea of population balances has been widely used in theoretical ecology and extended to the modeling of particulate systems in chemical engineering. The population balances can be either structured or unstructured models.

At the core of the structured population dynamics, the number of crystals in a fed-batch crystallizer is increased by nucleation and decreased by dissolution or breakage. Structured population balances models provide detailed information regarding the crystal size distribution in the crystallization unit. However, they demand a great deal of knowledge on the complex thermodynamic associated with the solute and solvent properties to be adequately incorporated in the population balances. Some important contributions in this subject have been reported in the literature (Worlitschek and Mazzotti, 2004; Nagy et al. 2007; Nowee et al., 2007).

Here, we introduce a simple unstructured population model, where the crystals are classified by their size, *L*. The growth of each individual crystal is supposed to be independent by the other crystals and is governed by the same deterministic model. In order to take into account the growth fluctuations and the unknown dynamics not captured by the deterministic term, a random component can be introduced (Gelb, 1988). The stochastic model can thus be written as a Langevin equation of the following type:

$$\dot{L} = f(L;\vartheta) + \eta(t) \tag{1}$$

In Equation 1, $f(L;\vartheta)$ is the expected rate of growth of L (the deterministic model introduced below), L is the size of the single crystal, t is the time, ϑ is the vector parameter defined in the model, and $\eta(t)$ is a random term assumed as Gaussian additive white noise:

$$\begin{aligned} &E[\eta(t)] = 0 \\ &E[\eta(t)\eta(t+\tau)] = 2D\delta(\tau) \end{aligned} \tag{2}$$

Where D is the additive noise intensity. Equation 1 implies that the crystal size L behaves as a random variable, characterized by a certain probability density function (PDF) $w(L,t)$ depending on the state variables of the system, i.e. the size L and time t. Incidentally, it should be noted that one can regard the probability density $w(L,t)$ as the relative ratio of crystals having a given dimension L, in the limit of infinite observations. Thus, from a practical point of view, it coincides with the Particle Size Distribution experimentally observed.

The new random variable thus can be described in terms of its probability density distribution, $w(L,t)$, at any instant of time t and should follow the linear Fokker-Planck Equation, FPE:

$$\partial_t w + \partial_L [f(L;\vartheta(u))w] = D(u)\partial_{LL} w \tag{3}$$

along with the boundary conditions:

$$f(0)w(0,t) - D(u)\partial_L w(0,t) = 0 \tag{4a}$$

$$\partial_L w(\infty,t) = 0 \tag{4b}$$

The reflecting boundary condition in Equation 4a ensures that the elements of the population will never assume negative values, whereas Equation 4b ensures the decay condition on $w(L,t)$ as L goes to infinity, for any time.

The diffusion coefficient D determines the random motion of the variable L that takes into account the fluctuation in the particle growth process (Randolph and Larson, 1988; Olesen et al, 2005).

Regarding the deterministic part of the model, our purpose is to choose a model as simple as possible, with a parsimonious number of adjustable parameters. To this end, the Generalized Logistic equation (Tsoularis and Wallace, 2002), is possibly the best-known simple sigmoidal asymptotic function used to describe the time dependence of growth processes in an unstructured fashion:

$$f(L;\vartheta) = r L^{\alpha} [1 - (L/K)^{\beta}]^{\gamma} \tag{5}$$

In Equation 4, L is the size of the single crystal, the crystal growth rate r and the equilibrium mean crystal size K are considered constant for each experimental condition and they are supposed to be only dependent on the anti-solvent flow rate. Moreover, α, β and γ are positive real numbers that regulate the shape of the growing curve. Hereafter we will consider the simple case with $\alpha = \beta = \gamma = 1$. With these assumptions, the present growth model can be regarded as the simplest model taking into account mild nonlinearities. In spite of this simplicity, this model provides the main qualitative features of a typical growth process: the growth follows a linear law at low crystal size values and saturates at a higher equilibrium value.

Finally the evolution in time of the probability density is described in terms of a linear, partial differential equation depending on the parameters r (linear Malthusian growth

rate), K (crystal size asymptotic value) and D (diffusivity), that are assumed to depend on the feeding conditions. This functionality is achieved by linear interpolation of the parameters as a function of the different anti-solvent flow rates. This parameterization with the anti-solvent flow rate allows the merger of multiple models for different operating regimes to a single model in the all operating envelop.

It is worth to stress out that the synergy between unstructured population balances and the Fokker-Planck equation results in structured-like population balances.

3. Experimental Validation

For parameter estimation and model validation purposes, three different anti-solvent flow-rates were implemented. All experiments were carried in a bench scale crystallizer which was kept at a fix temperature. Only purified water, regent grade sodium chloride (99.5%) and absolute ethanol (99.9%) were used. The experimental set-up and procedure are described as follows.

3.1 Experimental Set-up

The experimental rig is made up of one litre glass, cylindrical crystallizer submerged in a temperature controlled bath. The temperature in the bath is measured using an RTD probe which is wired up to a slave temperature control system capable of heating and cooling. In similar fashion, the anti-solvent addition is carried out by a slave peristaltic pump. The master control is performed by a Distributed Control System (Honeywell® DCS) which is wired up to the slave temperature and flow-rate controllers respectively. The desired set-points are calculated at the master controller. All relevant process variables are archived in the DCS system. In addition there is a particle size analyzer (Mastersizer® 2000) for the determination of the mean and crystal size distribution along the experiment. The mean crystal sizes and their distribution are also verified by visual inspection using a digital camera mounted in a microscope. The captured images are then processed by means of sizing computer software (Amscope®).

3.2 Experimental Procedure

At the start-up condition, the crystallizer is loaded with an aqueous solution of $NaCl$ made up of $34\,g$ of NaCl in $100\,g$ of water. The temperature is kept at $25^{\circ}C$. Then ethanol was added to the aqueous NaCl solution using a calibrated peristaltic pump. Along the operation, $5\,mL$ samples were taken in an infrequent fashion. The samples are then measured off-line using the particle size analyzer. Also, part of the sample was filtered over filter paper and then dried in an oven at $50^{\circ}C$ for farther visual inspection.

3.3 Parameter Estimation

The three parameters for the Fokker-Planck equation were estimated based-on experimental data obtained by monitoring the mean size distribution for different flow rates of anti-solvent. For every operating condition, that is, anti-solvent flow rate, a set of parameters $\vartheta=\{r,K,D\}$ is estimated using a nonlinear least-square algorithm. The values for the estimated parameters are given in the Table 1.

u (mL/min)	r (1/h)	K (μm)	D (μm^2 /h)
0.82	1.83	131.51	568.12
1.64	1.15	132.03	287.48
3.23	62.58	105.45	291.27

Table 1: Estimated parameters for the different operating conditions for the constant values $\alpha=1$, $\beta=1$ and $\gamma=1$.

3.4 Model Validation

In order to assess the prediction capabilities of the mathematical model based on the Fokker-Planck equation, the model predictions are compared with the reported experimental data within a valid range for the different operating conditions. It implies that the parameterization of the crystal growth rate, the free crystal size and the diffusivity coefficient with the anti-solvent are only reliable within the experimental range. The numerical solution of the mathematical model based on the Fokker-Planck equation requires an initial condition for the crystal size distribution. However, the initial condition requires information on crystal sizing. Observing the experimental data (Figure 1), the first available data value is at $t_0=0.1h$ (new origin), where the mean crystal size is around $L_0=92\,\mu m$ for anti-solvent flow rate. In order to take into account the uncertainty associated with this condition, a standard deviation $\sigma_0=30\,\mu m$ typical for this measurement is then assumed. Once the initial and boundary conditions are posed, the partial differential equation is then solved using a collocation method. The number of collocation knots $n=300$ and they are chosen as the roots of the Chebyshev polynomial of degree "n" where the solution domain is $L\in[0,\ell]$ with $\ell=210\,\mu m$ a sufficient high value. Note that a different set of data values were used for parameter estimation. Thus, the numerical solution of the Fokker-Planck equation (2-3) by the collocation methods provides the predictions for the mean crystal size by computing the first moment of the distribution (see Figure 1).

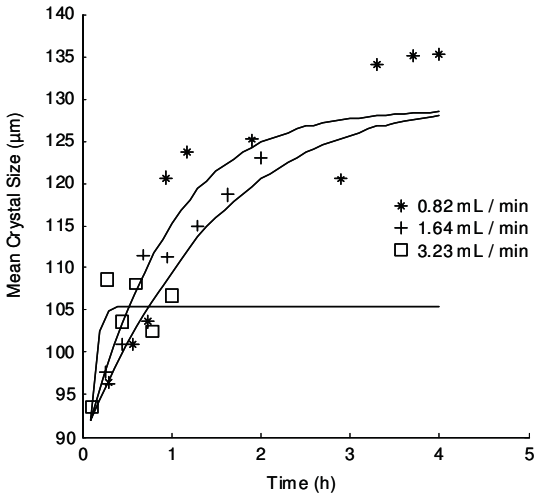


Figure 1: Model validation of the mean crystal size for different anti-solvent flow rates. Experimental results at low (*), medium (+) and high (□) anti-solvent flow rates. The three-parameter model based-on the Fokker-Planck equation predicts the dynamic behaviour of the mean crystal size growth for all anti-solvent flow rates (solid lines).

4. Calculation of the Robust Anti-Solvent Flow Rate Profile

4.1 Problem Formulation

Let us consider without loss of generality that the mean crystal size growth dynamics in an anti-solvent aided process is given by the following forced deterministic equation:

$$\dot{\mu}_L = F(\mu_L, u) \tag{6}$$

Where the forcing input (anti-solvent flow rate) is constrained, that is, $u \in [u^-, u^+]$. We say that the solutions of (6) in the time interval $t \in [0, t_0]$ for any input u and any given uncertain initial condition $\mu_L(0) \neq 0$ generates trajectories which at $t_0 > 0$ lie around a nominal value L_0, that is,

$$\mu_L(t_0) = L_0 \pm \Delta_0 \,, \forall t_0 > 0 \tag{7}$$

It means that regardless what the input is, the effect of the uncertainty in the initial condition is propagated to another point $\mu_L(t_0)$ in the trajectory (7), where a nominal mean crystal size L_0 is associated with an uncertainty Δ_0. It is advantageous since it is then possible to choose a new initial time origin at the point t_0, which coincides with experimental data value that helps to quantify the values for L_0 and Δ_0. Thus, the new initial condition at t_0 is still uncertain but it can be characterized and incorporated to our mathematical description of the problem.

The robust anti-solvent flow rate profile for the mean crystal size distribution is calculated using a hybrid strategy that requires engineering insight and process knowledge. The strategy involves a piece-wise function (8), a regulator-like (PI) algorithm (9), and a saturation function (10).

$$u = \begin{cases} u^+, \, if \;\; t \le t_0 \\ \zeta(\upsilon), \, otherwise \end{cases} \tag{8}$$

$$\dot{\upsilon} = \kappa_1 \dot{\mu}_L + \kappa_2 (\mu_L - L_r), \upsilon(t_0) = u^+ \tag{9}$$

$$\zeta(\upsilon) = \begin{cases} u^-, \, if \;\; \upsilon \le u^- \\ u^+, \, if \;\; \upsilon \ge u^+ \\ \upsilon, \, otherwise \end{cases} \tag{10}$$

From the actual starting-up condition at $t = 0$ to a chosen origin $t_0 > 0$, the anti-solvent flow rate is maintained constant at its maximum value u^+. It has been experimentally observed that the anti-solvent addition at the initial stage of the crystallization process promotes nucleation and speeds up the mean crystal size growth rate. Then, from time $t > t_0$ onwards the anti-solvent addition must be controlled in order to reach a desirable mean crystal size L_r. It is achieved by using a dynamic velocity-like regulator algorithm which is tuned selecting arbitrary values for the constants κ_1 and κ_2. Since the anti-solvent flow rate is constrained, it is necessary to include a saturation function to limit the control action within its physical range. It is important to highlight that velocity-like regulator with initial condition $\upsilon(t_0) = u^+$ introduces a bump-less transition from the saturation state. Also, the volume of spent anti-solvent at any time can be calculated using (11). The constant value "60" is the conversion factor for the anti-solvent flow rate from mL/min to mL/h.

$$V = 60[u^+ t_0 + \int_{t_0}^{t} u(\xi) d\xi] \tag{11}$$

It is important to remark that the strategy can be used either as a size-regulator, if the mean crystal size μ_L is a real-time available measurement or as tool to calculate an *off-line robust trajectory* for the anti-solvent addition. Due to its simplicity, this strategy can be easily implemented.

4.2 Simulation and Experimental Validation of Anti-Solvent Addition Policy

We next use numerical simulations and experimental validation to assess the closed loop performance of the anti-solvent addition strategy proposed for this specific crystallization process. The simulation parameters are given in Table 2. Since the experimental data is constrained to a certain operating window, the calculated profile is also limited to this operating range. The anti-solvent addition strategy was tested within the validity range of the experimental data and targeting a medium mean crystal size. The initial condition is then represented as a Gaussian –like distribution with mean value $L_0 = 92\,\mu m$ and standard deviation $\sigma_0 = 30\,\mu m$ based-on the experimental information.

Parameter	Value
u^- (mL/min)	0.82
u^+ (mL/min)	3.23
κ_1 (mL/h ·µm)	0.0052
κ_2 (mL/h^2· µm)	-0.0015
L_0 (µm)	92.0
σ_0 (µm)	30.0
ℓ (µm)	210.0
t_0 (h)	0.1
Δt (h)	0.1

Table 2: Simulation parameters for the assessment of the anti-solvent addition strategy.

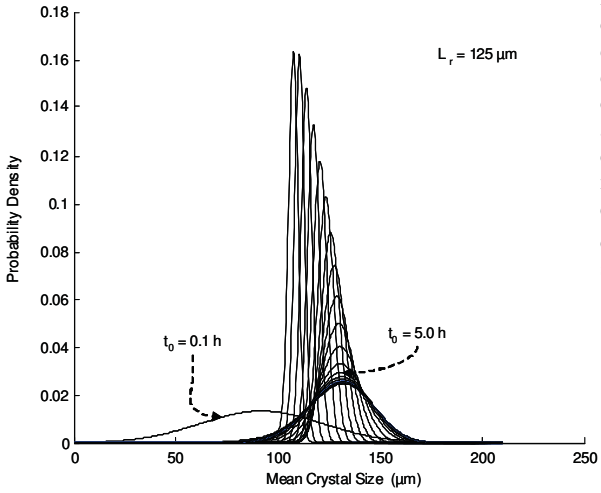


Figure 2: Propagation of the probability density.

Figure 2 depicts the time evolution of the time evolution of the probability density function as function of time for $L_r = 125 \mu m$. For sake of brevity only one value is reported here. For a set point of $L_r = 125 \mu m$ (medium size), the crystal growth rate is alike the previous case, however the anti-solvent flow rate is cut down to an even lower flow rate. It makes clear that the transition between high and low concentration of anti-solvent in the solution promotes the crystal growth instead of the new crystal formation. Therefore, it results in larger crystal size. The volume of anti-solvent added to the crystallizer for a desired mean crystal sizes are calculated using (11) and it is 669.48 *mL* for medium crystal size. The calculation is based on 5.0 hours of operation. Note that the upper and lower prediction bounds for the mean size distribution (Figure 3) are wider at the start-up condition and its narrows when reaching the desired set-point value. The steady-state crystal size distribution suggests that at controlled conditions it is possible to have a very narrow mean size distribution for a given anti-solvent addition policy (Figure 4).

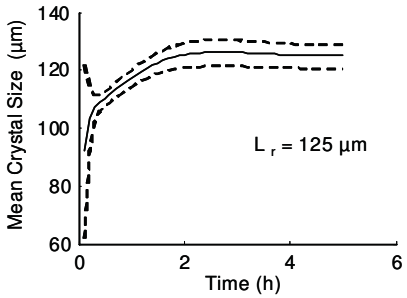


Figure 3: Predicted mean crystal size and its lower and upper bound.

The distribution indicates how disperse the experimental data is. The observed dispersion is incorporated in the diffusivity coefficient of the Fokker-Planck equation. It is envisaged that depending on the amount and quality of the experimental data, the diffusivity coefficient can be clearly related with the crystal size distribution and predicted by this modeling approach. Figure 5 shows the time evolution of the mean crystal size when the anti-solvent addition policy is implemented experimentally for a desired value of time evolution of the $L_r = 125 \mu m$. The quantitative growth of crystal size can be seen in the Figures 6a to 6c. Note the size scale at the corner of every picture.

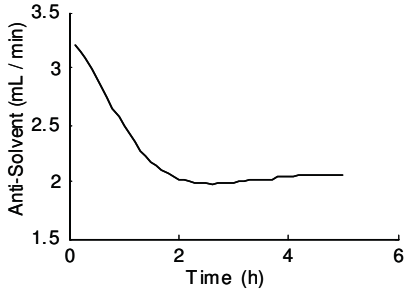


Figure 4: Anti-solvent addition policy for a desired mean crystal size of $L_r = 125 \mu m$.

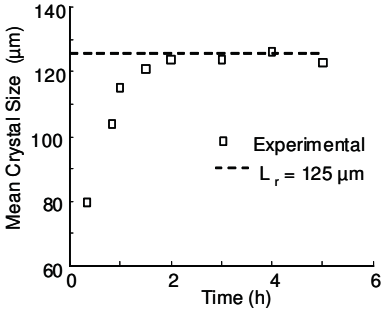


Figure 5: Mean crystal size evolution when the anti-solvent addition policy is implemented.

5. CONCLUSIONS

The hybrid strategy of using the Fokker-Planck Equation (FPE) and the PI-like regulator for the calculation of the anti-solvent addition policy has proved to be simple and still a powerful way to control the mean crystal size in a crystallization operation. It is envisaged that the FPE is a rather useful fashion to study systems with uncertain initial condition in dynamic systems. The initial uncertainty can be quantified and naturally included in the structure of the solution. Future work will be devoted to exploit the FPE approach for the determination of the mean crystal size and its distribution in cooling and anti-solvent aided crystallization.

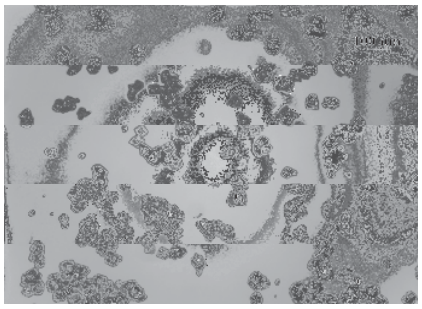


Figure 6.a $NaCl$ crystals at $t = 0.16\ h$

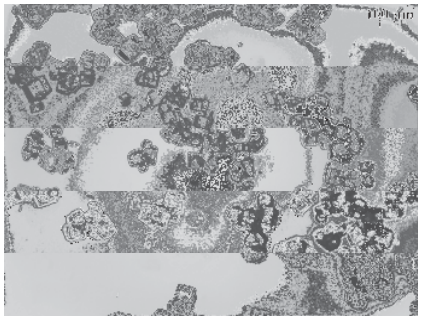


Figure 6.b $NaCl$ crystals at $t = 1\ h$

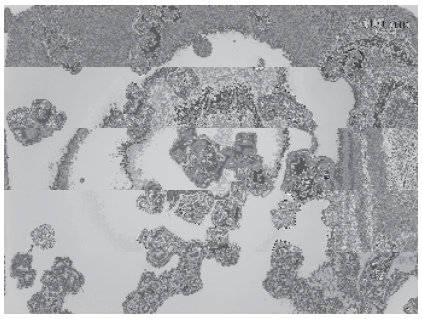


Figure 6.c $NaCl$ crystals at $t = 2\ h$

REFERENCES

Feigenbaum, M. J. (1980), "Universal Behaviour in Nonlinear Systems", Los Alamos Science, 1, pp. 4 – 27

Gelb A., Applied Optimal Estimation (1988) M.I.T. Press, Cambridge Massachusetts, USA

Grosso, M., Cella, Baratti R. (2007), "Development of a Probabilistic Model for the Plant Growth in the Nursery Phase", Food and Agriculture Products: Processing and Innovations, September 24-26, Naples, Italy.

Lo, C.F. (2007) "Stochastic Gompertz Model of Tumor Cell Growth." Journal of Theoretical Biology, 248, pp. 317-321.

Matsoukas T., Yulan, L. (2006), "Fokker-Plank Equation for Particle Growth by Monomer Attachment", Phys. Rev. E, 74, pp. 031122-1-031122-9.

May R., McLean, A. (2007), "Theoretical Ecology" Principles and Applications", Oxford University Press, Great Britain

Nagy Z., Fujiwara, M., Braatz, R. (2007). "Recent Advances in the Modelling and Control of Cooling and Anti-Solvent Crystallization of Pharmaceuticals", Dynamics and Control Process Systems, Vol. 2, pp. 29-38, June 6-8, 2007, Cancún, Mexico

Nowee S.M., Abbas A., Romagnoli, J.A. (2007), "Optimization in Seeded Cooling Crystallization: A Parameter Estimation and Dynamic Optimization Study", Chem. Eng. Proc., 46, pp. 1096-1106.

Olesen P., Ferkinghoff-Borg J., Jensen M.H., Mathiesen J., Phys. Rev. E (2005), 72, 031103

Randolph A.D., Larson M. (1988), Theory of Particulate Processes: Analyses and Techniques of Continuous Crystallization, Academic Press, San Diego, USA.

Risken H., (1984), "The Fokker-Planck Equation: Methods of Solution and Applications", Springer –Verlag, Germany.

Tsoularis A., Wallace J. (2002), "Analysis of logistic growth models", Math. Biosci., 179, pp. 21-55

Worlitschek J., Mazzotti M. (2004), "Model-Based Optimization of Particle Size Distribution in Batch-Cooling Crystallization of Paracetamol", Crystal Growth & Design, Vol. 4, No. 5, pp. 891-903.

Model Based Robust Batch-to-Batch Control of Particle Size and Shape in Pharmaceutical Crystallisation

Zoltan K. Nagy

Chemical Engineering Department, Loughborough University,
Loughborough, LE11 3TU, United Kingdom, (e-mail: Z.K.Nagy@lboro.ac.uk)

Abstract: The paper presents a novel batch-to-batch control approach for crystallization processes, which can be used for designing the shape of the crystal size distribution (CSD) and the habit of the crystals, to robustly achieve desired product properties. The method is implemented in a hierarchical structure. On the lower level a supersaturation control approach is used that drives the system in the phase diagram according to a concentration versus temperature trajectory, providing the within batch control. On the higher level a robust model-based optimization algorithm adapts the setpoint of the supersaturation controller, after each batch, using the adapted model parameters by applying a batch-to-batch moving horizon estimation (MHE) approach, to counteract the effects of changing operating conditions, and parameter uncertainties. The process is modelled using the two dimensional population balance equation (PBE), which is solved using the method of characteristics (MOC). The control approach is corroborated through simulations studies. The results illustrate the benefits of the robust iterative learning approach which is able to control both the entire CSD and the habit of the product crystals, reducing significantly the variability in the product properties after only four batches.

1. INTRODUCTION

Crystallization is one of the key unit operations in the pharmaceutical, food and fine chemicals industries. Despite the long history and widespread application of batch crystallization, there remains a disproportionate number of problems associated with its control (Braatz, 2002), mainly related the complex nonlinear dynamics with nonideal mixing, and various disturbances characteristic to these systems. The operating conditions of the crystallization process determine the physical properties of the products which are directly related to the crystal size distribution (CSD), shape or polymorphic form. These properties determine the efficiency of downstream operations, such as filtration, drying, and tablet formation, and the product effectiveness, such as bioavailability and shelf-life. With the recent change of industrial procedures from Quality-by-Testing (QbT) to Quality-by-Design (QbD) and the advent of process analytical technology (PAT) initiative, especially in the pharmaceutical industries, approaches which can be used to design desired product properties are of great interest (Fujiwara *et al.*, 2005). The classical control objectives expressed in characteristics of the size distribution (e.g. maximize average size, minimize coefficient of variation) can lead to conservative and economically inefficient designs of the crystallization systems, and they most often neglect the shape of the crystals. The paper presents a batch-to-batch iterative control (ILC) approach which can be used to directly design the shape of the crystal size distribution and the crystal habit at the same time, to robustly achieve desired product properties. The method is able for example to minimize filtration time without generating unnecessarily large crystals, or minimize breakage by controlling the aspect ratio. Since dissolution rate depends on the shape of the CSD, when the resulting crystals represent the final product (e.g. drugs for inhalers) controlling the shape of the CSD can provide novel applications in the area of drug delivery, or environmentally friendly dosage of pesticides, where particular multimodal distributions can be designed to achieve desired concentration level of the active ingredient. The crystallization system in this study is modelled via a two-dimensional population balance equation (2D-PBE) which is directly used in the optimization procedure where the objective function is expressed in terms of the shape of the entire CSD, and the aspect ratio is added as constraints. The control of crystal size and shape has been considered previously by Lee *at al.* (2002), however in their approach the authors used a computationally more demanding high resolution finite volume method to solve the 2D population balance equation, and parameter uncertainties were not considered directly in the control problem formulation. In this paper the population balance equation (PBE) is solved using a generic 2D-method of characteristics (LeVeque, 1992). Crystallization models are generally subject to significant uncertainties. A robust model based optimization approach (Nagy and Braatz, 2003) is evaluated and it is shown that taking parametric uncertainties into account in the problem formulation can lead to significant improvement in the robustness of the product quality after a few batches only. The control approach is implemented in a hierarchical

structure where on the lower level a model-free crystallization control methodology, the supersaturation controller, drives the system in the phase diagram, rather then in the time domain, providing the within batch control methodology, whereas on the higher level a robust batch-to-batch model based optimization algorithm, the adapts the setpoint of the supersaturation controller after each batch to counteract the effects of changing operating conditions and parameter uncertainties. The method adapts the uncertain kinetic parameters using the information available over the past batches, applying a moving horizon estimation (MHE) scheme (Rawlings *et al.*, 1993), which also provides the uncertainty description used in the robust optimal control problem. The optimization problem is solved using an efficient multistage approach implemented in the optimization package OptCon (Nagy *et al.*, 2004). The proposed approach is corroborated in the case of a simulated crystallization system.

2. 2D POPULATION BALANCE MODELLING OF BATCH CRYSTALLIZATION PROCESSES

Considering a single growth direction with two characteristic lengths L_1 and L_2, and a well-mixed crystallizer with growth and nucleation as the only dominating phenomena the crystal size distribution (CSD) expressed in the number density function $f_n(L_1, L_2, t)$, is given by the population balance equation (PBE) with the form

$$\frac{\partial f_n(L_1,L_2,t)}{\partial t} + \frac{\partial\{G_1(S,L_1;\theta_{g1}) f_n(L_1,L_2,t)\}}{\partial L_1} + \frac{\partial\{G_2(S,L_2;\theta_{g2}) f_n(L_1,L_2,t)\}}{\partial L_2} = B(S;\theta_b)\delta(L_1 - r_{01}, L_2 - r_{02}) \tag{1}$$

with initial condition given by the size distribution of seed, $f(L_1, L_2, 0) = f_{seed}(L_{1,0}, L_{2,0})$, t is time, $G_1(S, L_1; \theta_{g1})$ and $G_2(S, L_2; \theta_{g2})$ are the generic size dependent growth rates of crystal in the two characteristic directions. $B(S;\theta_b)$ is the nucleation rate, $\delta(\cdot,\cdot)$ is the two dimensional Dirac delta function, r_{01} and r_{02} are the characteristic sizes of the nucleai, $S = C - C_{sat}$ is the supersaturation, C is the solute concentration, $C_{sat}(T)$ is the saturation concentration at the temperature T, and θ_{g1}, θ_{g2} and θ_b are vectors of growth and nucleation kinetic parameters, respectively. The partial differential equation can be reduced to a system of ODEs by applying the method of characteristics (MOC). The aim of the MOC is to solve the PBE by finding characteristic planes in the $L_1 - L_2 - t$ space that reduce the PBE to a system of ODEs. The $L_1 - L_2 - t$ space is expressed in a parametric form by $L_1 = L_1(\mathcal{Z})$, $L_2 = L_2(\mathcal{Z})$ and $t = t(\mathcal{Z})$, where the parameter $\mathcal{Z}$ gives the measure of the distance along the characteristic curve. Hence, applying the chain rule with $f_n(L_1, L_2, t) = f_n(L_1(\mathcal{Z}), L_2(\mathcal{Z}), t(\mathcal{Z}))$ gives,

$$\frac{df_n}{d\mathcal{Z}} = \frac{dL_1}{d\mathcal{Z}}\frac{\partial f_n}{\partial L_1} + \frac{dL_2}{d\mathcal{Z}}\frac{\partial f_n}{\partial L_2} + \frac{dt}{d\mathcal{Z}}\frac{\partial f_n}{\partial t}. \tag{2}$$

Comparing (2) with (1) we find $\mathcal{Z} = t$ and the characteristic equations can be derived. Solving these together with the system of equations which results by applying the method of moments (MOM), we can calculate the dynamic evolution of $f_n(L_1, L_2, t)$ by the following ODEs,

$$\begin{aligned}
\frac{d\mu_{0,0}}{dt} &= B,\\
\frac{d\mu_{i,j}}{dt} &= iG_1\mu_{i-1,j} + jG_2\mu_{i,j-1},\\
\frac{dL_1}{dt} &= G_1,\\
\frac{dL_2}{dt} &= G_2,\\
\frac{\partial f_n}{\partial t} &= -f_n(L_1,L_2,t)\left(\frac{dG_1}{dL_1} + \frac{dG_2}{dL_2}\right) + B\delta(L_1 - r_{01}, L_2 - r_{02}),
\end{aligned} \tag{3}$$

where the system was solved for zeroth moment $\mu_{0,0}$, first order moments $\mu_{1,0}$ and $\mu_{0,1}$, second order moments $\mu_{1,1}$, $\mu_{2,0}$ and $\mu_{0,2}$, and third order moments $\mu_{2,1}$, $\mu_{1,2}$, $\mu_{3,0}$, $\mu_{0,3}$, defined by

$$\mu_{i,j} = \int_0^\infty \int_0^\infty L_1^{\,i} L_2^{\,j} f_n(L_1, L_2, t)\, dL_1 dL_2, \tag{4}$$

with initial conditions given by the following vector $x_0 = [\mu_{0,0}(0), \ldots, \mu_{03}(0), L_{1,0}, L_{2,0}, f_{seed}(L_{1,0}, L_{2,0})]$. The solute concentration is given by

$$C(t) = C(0) - \rho_c\left(\mu_{21}(t) - \mu_{21}(0)\right), \tag{5}$$

where ρ_c is the density of crystals. In (5) it is considered that the shape of the crystal is rectangular with L_1 and L_2 being width (and depth) and length, respectively. The nucleation kinetics is given by

$$B = k_b S^b, \tag{6}$$

with nucleation parameters $\theta_b = [k_b, \; b,]$. In the case when the growth rates are independent of size and are expressed by

$$G_1 = k_{g1} S^{g1}, \;\; G_2 = k_{g2} S^{g2}, \tag{7}$$

with growth parameters $\theta_{g1} = [k_{g1,} \; g1]$ and $\theta_{g2} = [k_{g2,} \; g2]$. In the case of size independent growth $dG_1 / dL_1 = 0$ and $dG_2 / dL_2 = 0$ and the system of ODEs (3) can be solved analytically (Rusli *et al.*, 2006) and the CSD can be constructed in any time step using different initial conditions obtained by varying $L_{1,0}$ and $L_{2,0}$, the shape of the distribution can be obtained with desired resolution. Additionally the analytical solution can be simplified by the assumption of growth dominated process, and constant supersaturation, which can be considered valid for supersaturation controlled processes. The dissolution process

was modelled similarly, with expressions similar to (7), but different constants, which apply when $S < 0$.

Several approaches have been proposed for designing the operating curves for crystallization systems. Generally speaking, two main categories can be distinguished, (i) the model-based approach (Rawlings *et al.*, 1993) and (ii) the direct design (Fujiwara *et al.*, 2005). In the model-based design approach the detailed model (4) is used together with optimization techniques to determine temperature versus time trajectories, which optimize desired product properties, usually expressed as functions of the moments of the CSD. The direct design approach is based on the understanding of the basic concept of crystallization, to operate the system within the metastable zone bounded by the nucleation and solubility curves. In this technique a supersaturation setpoint profile is chosen experimentally and it is followed in the phase diagram using a supersaturation controller based on concentration measurement. The approach proposed in the paper combines the two methods in a hierarchical control algorithm, in which a model-based robust optimization determines the operating profile in the phase diagram, which is used then as the setpoint for the supersaturation controller.

3. DISTRIBUTIONAL BATCH-TO-BATCH NMPC

The main feature of the batch-to-batch control is that variations on two time-scales must be considered. As shown on Figure 1, within and between batches variations can be considered leading to an optimization problem on two time-scales. In batch process operation often within batch measurements are not available or adjustment to the operating conditions cannot be made. In these cases batch-to-batch improvement is practically easier to implement, by learning from the information obtained usually from after-batch laboratory analyses. In this framework the within batch measurements (if available) can be used for model based parameter and state estimation/adaptation, and the updated model then can be used in an iterative learning control framework for the improvement of the future operating recipe (Figure 2).

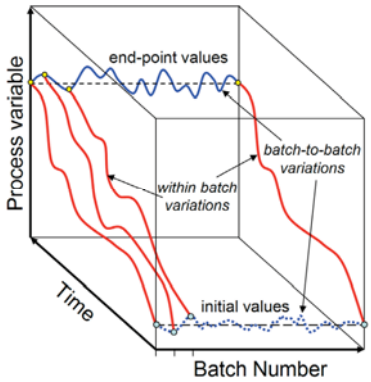


Fig. 1. Schematic representation of the dynamic two-time scale variations in batch control.

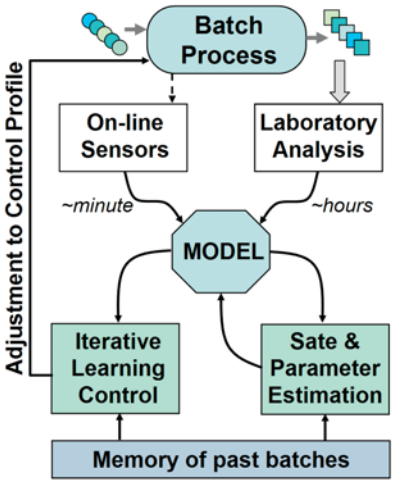


Fig. 2. Structure of the iterative learning control framework.

The optimal control problem for the iterative learning scheme can be formulated as the alternative application of a moving horizon estimation (MHE) and a robust optimal control problem. The MHE problem solved after each batch k is,

$$\min_{\hat{x}_{0,k},\theta_k} \sum_{i=k-1}^{k-N_b} W_i (y_i^{\text{model}}(\hat{x}_{0,k};\theta_k) - y_i^{\text{meas.}})^2 \tag{8}$$

subject to:

$$\dot{x}_i = f(x_i, u_i; \theta_k); \quad x_{0,i} = x_{0,k}; \; x_i \in \mathcal{X} \tag{9}$$

$$y_i = g(x_i, u_i; \theta_k); \; i = k-1, \ldots, k-N_b \tag{10}$$

where N_b is the number of past batches used in the estimation and W_i is a weighting matrix with exponential forgetting factor over the past batches. The model output used in the parameter estimation in the case of crystallization processes can be the entire CSD or properties of the CSD (e.g. number or weight average mean size). From the parameter estimation problem estimates of the uncertainty bounds on the parameters are also calculated, which can then be used in the robust optimization problem for control. The model based optimal control problem is expressed as

$$\min_{\Delta u_k} \mathcal{H}(x_k, u_k; \theta_k) \tag{11}$$

subject to:

$$\dot{x}_k(t) = f(x_k(t), u_k(t); \theta_k); \quad x_k(t_0) = \hat{x}_{0,k} \tag{12}$$

$$y_k(t) = g(x_k(t), u_k(t); \theta_k) \tag{13}$$

$$u_k = u_{k-1} + \Delta u_k \tag{14}$$

$$h(x_k(t), u_k(t); \theta_k) \le 0, \quad t \in [t_0, t_F] \tag{15}$$

where $\mathcal{H}$ is the performance objective, t is the time, t_F is the final time at the end of prediction (end of batch), $x(t) \in \mathbb{R}^{n_x}$ is the vector of states, $u(t) \in \mathcal{U} \subset \mathbb{R}^{n_u}$ is the set of input vectors, $y(t) \in \mathbb{R}^{n_y}$ is the n_y vector of measured variables used to compute the estimated states $\hat{x}(t_k)$,

$\theta \in \Theta \subset \mathbb{R}^{n_\theta}$ is the n_θ vector of uncertain parameters, where the set Θ can be either defined by hard bounds or probabilistic, characterized by a multivariate probability density function. The function $f : \mathbb{R}^{n_x} \times \mathcal{U} \times \Theta \rightarrow \mathbb{R}^{n_x}$ is the twice continuously differentiable vector function of the dynamic equations of the system, $g : \mathbb{R}^{n_x} \times \mathcal{U} \times \Theta \rightarrow \mathbb{R}^{n_y}$ is the measurement equations function, and $h : \mathbb{R}^{n_x} \times \mathcal{U} \times \Theta \rightarrow \mathbb{R}^{c}$ is the vector of functions that describe all linear and nonlinear, time-varying or end-time algebraic constraints for the system, where c denotes the number of these constraints. The repeated optimization problem is solved by dividing the batch time $t \in [0, t_F]$ into N equally spaced time intervals Δt (stages), with discrete time steps $t_k = k\Delta t$, and $k = 0, 1, \ldots, N$ (Biegler and Rawlings, 1991). The model based control approach is implemented in the Matlab toolbox, OptCon (Nagy *et al.*, 2007), which is based on a state-of-the-art large-scale nonlinear optimization solver (HQP) (Franke *et al.*), which uses a multiple shooting algorithm (Diehl *et al.*, 2002).

Consider the case of parameter uncertainty, with $\delta\theta \in \mathbb{R}^{n_\theta}$ defined as the perturbation about the nominal parameter vector $\hat{\theta}$. The real uncertain parameter vector is then given by $\theta = \hat{\theta} + \delta\theta$. Assuming zero mean, normal measurement errors, and known covariance matrix, the set of possible parameter values is given by the hyperellipsoidal confidence region, defined as

$$\Theta(\alpha) \triangleq \{\theta : (\theta - \hat{\theta})^T \mathbf{V}_\theta^{-1} (\theta - \hat{\theta}) \leq \chi^2_{n_\theta}(\alpha)\}, \qquad (16)$$

where α is the confidence level, $\chi^2_{n_\theta}(\alpha)$ is a quantile of the chi-squared distribution with n_θ degrees of freedom, and $\mathbf{V}_\theta \in \mathbb{R}^{n_\theta \times n_\theta}$ is the parameter covariance matrix. Uncertainty description (16) results most commonly from typical least-squares identification procedures from experimental data, and it is provided by the MHE estimation approach applied in this paper. The parameter covariance matrix is updated after each batch when a new set of parameters are also computed using the MHE algorithms. We denote with $\psi(x(t_f);\theta)$ the end-point property of interest. Considering the mean-variance approach the following objective function is used to account for parameter uncertainties in the NMPC:

$$\mathcal{H} = (1 - w)\mathcal{E}[\psi(x(t_f), \theta)] + w V_\psi(t_f), \qquad (17)$$

where $\mathcal{E}$ and $V_\psi \in \mathbb{R}$ is the expected value and variance, respectively, of the property at the end of the batch, and $w \in [0,1]$ is a weighting coefficient that quantifies the tradeoff between nominal and robust performance. The main advantage of this approach compared to the classical minmax optimizations is that the tradeoff between nominal and robust performance can be controlled by appropriately weighting the two objectives. Expected value and variance can be computed efficiently using a second order power series expansion,

$$\delta\psi = L\delta\theta + \frac{1}{2}\delta\theta^T \mathbf{M}\delta\theta + \ldots, \qquad (18)$$

where $L = (d\psi / d\theta)_{\hat{\theta},u} \in \mathbb{R}^{n_y}$, and $\mathbf{M} = (d^2\psi / d\theta^2)_{\hat{\theta},u} \in \mathbb{R}^{n_y \times n_\theta}$ are the first and second order sensitivities, respectively. Assuming zero mean, normally distributed parameters $\delta\theta$, deriving the expected value and variance of $\delta\psi$ based on (18) gives the analytical expressions:

$$\mathcal{E}[\delta\psi] = \frac{1}{2}\mathrm{tr}(\mathbf{M}\mathbf{V}_\theta) \qquad (19)$$

$$V_\psi = L\mathbf{V}_\theta L^T + \frac{1}{2}[\mathrm{tr}(\mathbf{M}\mathbf{V}_\theta)]^2 \qquad (20)$$

where $\mathrm{tr}(\mathbf{A})$ is the trace of matrix $\mathbf{A}$. The feasibility of the optimization under parametric uncertainty is achieved by reformulating the constraints in a probabilistic sense:

$$\mathbb{P}(h_i(x, u;\theta) \leq 0) \geq \alpha_i, \qquad (21)$$

where $\mathbb{P}$ is the probability and α_i is the desired confidence level for the satisfaction of constraint i. The robust formulation of (21) can be written using the t-test in the form:

$$\mathcal{E}[h_i] + t_{\alpha/2,n_\theta}\sqrt{V_{h_i}} \leq 0, \quad i = 1, \ldots, c. \qquad (22)$$

The expected value ($\mathcal{E}[h_i]$) and covariance (V_{h_i}) of the constraint h_i can be evaluated using first or second order approximations. For first order approximation $\mathcal{E}[h_i(x,u;\theta)] = h_i(x,u;\hat{\theta})$ and $V_{h_i} = L_{h_i}\mathbf{V}_\theta L_{h_i}^T$, whereas for second order approximation expressions similar to (19) and (20) can be used, with $L_{h_i} = (dh_i / d\theta)_{\hat{\theta},u} \in \mathbb{R}^{n_\theta}$, and $\mathbf{M}_{h_i} = (d^2 h_i / d\theta^2)_{\hat{\theta},u} \in \mathbb{R}^{n_\theta \times n_\theta}$. In this formulation the algorithm shows robust performance in the sense of constraint satisfaction and decreased variance of the performance index (Nagy and Braatz, 2004).

4. APPLICATION OF THE ROBUST BATCH-TO-BATCH NMPC FOR CRYSTALLIZATION PRODUCT DESIGN

For the case studies the crystallization of a pharmaceutical was considered as the model system, for which nucleation and growth kinetics were determined experimentally using image analysis using the Sympatec Qicpic equipment. It was found that a 1D-PBM was not able to describe accurately the variation of the CSD, since a time-varying volumetric shape factor was found to be necessary to capture the significant change in aspect ratio during the batch. Using the 2D-PBE the volume of the individual crystals are directly calculated hence the volumetric shape factor is not needed. Although, the 2D-PBM provides the full 2D CSD, in this paper the target distribution was given as a 1D CSD based on the equivalent spherical diameter, since this is the most commonly used characterisation mode of CSD used in practice. Additionally the *in situ* CSD measurement during the experiments was based on focused beam reflectance measurement (FBRM), which provides 1D information. The equivalent spherical diameter (r) is calculated by

$$r = \frac{6L_1^2 L_2}{\pi}. \qquad (23)$$

The 2D-PBM also allows the incorporation of the shape information in the optimization problem. Different product design problems were considered, when various objective functions expressed as desired shapes of the CSD and limits of the crystal aspect ratio ($\alpha = L_1 / L_2$) were optimized and the required temperature profiles were determined. The novel feature of the proposed approach is that the optimization is performed in the phase diagram, and a concentration trajectory for batch k is obtained as a function of temperature $C_{set,k} = f(T)$. This allows the direct application of the widely accepted supersaturation control in conjunction with the NMPC. The optimization problem can be expressed by the generic robust formulation:

$$\min_{C_{set,k}(T)} \{(1-w)\sum_i (f_n(r_i,t_f;\theta_k) - f_n^{desired}(r_i,t_f))^2 + wV[f_n(r,t_f;\theta_k]\} \quad (24)$$

$$\begin{aligned} s.t. \quad & T_{\min} \le T(t) \le T_{\max} \\ & R_{\min} \le \frac{dT}{dt} \le R_{\max} \\ & C(t_f) \le C_{\max} \\ & \alpha_{\min} \le \alpha \le \alpha_{\max} \end{aligned} \quad (25)$$

where $f_n^{desired}(r_i,t_f)$ is the desired (setpoint) CSD with a given shape at the end of the batch, $T_{\min}$, $T_{\max}$, $R_{\min}$, $R_{\max}$ are the operating constraints determined by the bounds and minimum and maximum rate changes of the temperature profiles, respectively, C is the concentration, $C_{\max}$ is the maximum concentration at the end of batch required to achieve a desired crystallization yield, and $\alpha_{\min}$ and $\alpha_{\max}$ are the minimum and maximum bounds on the aspect ration for the desired crystal shape. For most crystallization processes there are significant variations in the metastable zone width (MSZW), which is incorporated in the optimization by considering uncertainties in the nucleation parameters ($\theta_b = [k_b, b]$). The variations in the nucleation kinetics are usually significantly larger than the uncertainties in the growth hence this will be considered in this study. In the MHE scheme a fixed iteration approach was used and the convergence of the two nucleation parameters is shown on Figure 3. It can be seen that the MHE converges practically after 3-4 iterations, after which the robust profile is also constant. Figure 4 shows the metastable zone width in the phase diagram delimited by the solubility and nucleation curves. Because of the uncertainties in the nucleation kinetics there is a nucleation region, with a width given by the 99% confidence intervals obtained via Monte Carlo simulations. The operating profiles resulted by solving (24)-(25) for the nominal case ($w = 0$) and one robust case ($w = 0.3$) are shown in the phase diagram in Figure 4. The target CSD was monomodal with a shape biased toward larger particles to improve filtration. The robust operating profile which resulted after 4 iterations, corresponds to a trajectory, which is further away from the nucleation zone throughout the entire batch. The operating profiles are implemented using a supersaturation controller. Figure 5 shows the time-domain representation of the operating curves corresponding to Figure 4. Since the robust profile operates at lower superasaturation the cooling is slower than in the nominal case resulting in longer batch time for similar yield. The robust operating policy also indicates slower cooling and even a slight increase in the temperature during the initial part of the batch when the nuclei are generated. This is in correlation with the often used industrial practice, according to which slow cooling and moderate increase in temperature after the onset of nucleation can result in improved consistency in the final CSD. Monte Carlo simulations were performed by randomly sampling (100 samples) the uncertain parameter space θ_b and applying the nominal and robust temperature profiles. Figure 6 demonstrates that the robust operating curve leads to significantly reduced variability in the product quality compared to the nominal operating policy. The incorporation of the constraint with respect to the aspect ratio allows controlling the crystal habit (aspect ratio) and shape of the CSD at the same time. Figure 7 represents the variations during repeated batches in the product quality expressed as the aspect ratio and maximum Feret diameter. The Monte Carlo simulations show that the robust iterative learning control approach with the profile shown in Figures 4 and 5, provides a significantly lower variability in the crystal size and aspect ratio due to parameter uncertainty.

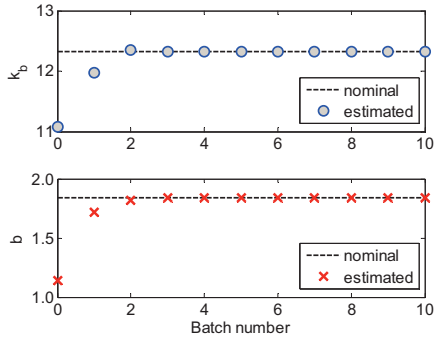


Fig. 3. Evolution of the parameters estimated by the MHE over the batches.

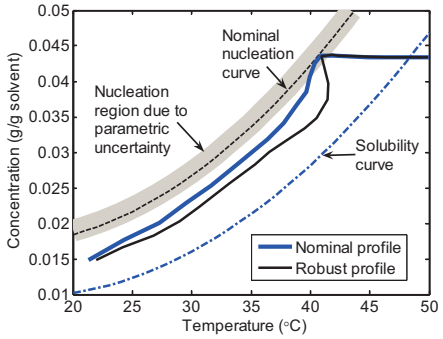


Fig. 4. Phase diagram with nominal ($w = 0$) and robust ($w = 0.3$) operating curves for monomodal target CSD.

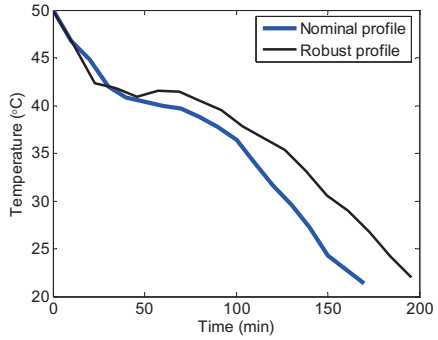


Fig. 5. Time-domain representation of the nominal and robust operating profiles corresponding to Figure 3.

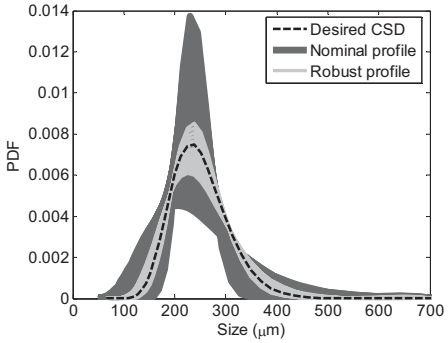


Fig. 6. Monomodal target and product CSDs resulting from Monte Carlo simulations with the uncertain nucleation parameters using the nominal and robust operating profiles.

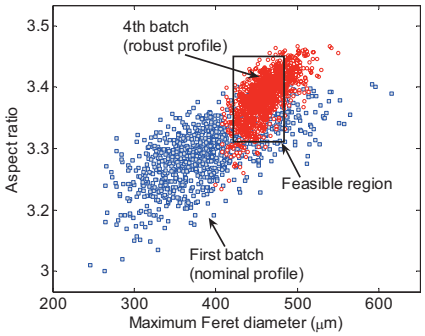


Fig. 7. Monte Carlo simulations showing the performance of the robust iterative learning control after 4 batches.

In the case of the robust ILC most of the batches provide product within the feasible region whereas the original nominal trajectory leads to a very high proportion of off-spec product, which could lead to decreased efficiency of the downstream processes or problems in formulation.

6. CONCLUSIONS

The paper presents a novel robust batch-to-batch control approach for the design of crystalline products by shaping the crystal size distribution and habit. A distributional optimization approach is used to design a robust concentration versus temperature profile, which is used as a setpoint for a lower level supersaturation controller. A two dimensional population balance model is solved using the method of characteristics, to capture the dynamic evolution of the aspect ratio which is incorporated into the robust optimization problem to control the shape of the size distribution and habit at the same time. Simulation results demonstrate the benefits of the proposed approach, which can decrease variability in size and shape of the product in a few batches.

ACKNOWLEDGEMENTS

Financial support by EPSRC (EP/E022294/1) and the Royal Academy of Engineering, U.K. is gratefully acknowledged.

REFERENCES

Biegler, L., J. Rawlings (1991). Optimization Approaches to Nonlinear Model Predictive Control, *Chemical Process Control*, South Padre Island, TX, 1991.

Braatz, R.D. (2002). Advanced control of crystallization processes, Annual Reviews in Control 26, 87-99.

Diehl, M., H.G. Bock, J.P. Schlöder, R. Findeisen, Z.K. Nagy, F. Allgöwer (2002). Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations, *J. of Process Control* 12, 577-585.

Franke, R., E. Arnold, H. Linke. HQP: A solver for nonlin-early constrained optimization. http://hqp.sourceforge.net.

Fujiwara, M., Z.K. Nagy, J.W. Chew, R.D. Braatz (2005). First-principles and direct design approaches for the control of pharmaceutical crystallization, *J. of Process Control* 15, 493-504.

Lee, K., J.H. Lee, M. Fujiwara, D.L. Ma, R.D. Braatz (2002). Run-to-run control of multidimensional crystal size distribution in a batch crystallizer, *Proc. of the American Control Conf.*, IEEE Press, Piscataway, NJ, 1013-1018.

LeVeque, R. (1992). Numerical Methods for Conservation Laws, Birkhauser.

Nagy, Z.K., B. Mahn, R. Franke, F. Allgower (2007). Efficient output feedback nonlinear model predictive control for temperature control of industrial batch reactors, Control Engineering Practice, 15, 839-859.

Nagy, Z.K., R.D. Braatz (2003). Robust nonlinear model predictive control of batch processes, *AIChE J.* 49, 1776-1786.

Nagy, Z.K., R.D. Braatz (2004). Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis, *J. of Proc. Control*, 14, 411-422.

Rawlings, J., S. Miller, W. Witkowski (1993). Model identification and control of solution crystallization processes A review, *Ind. Eng. Chem. Res.* 32, 1275-1296.

Rusli, E., J.H. Lee, R.D. Braatz (2006). Optimal distributional control of crystal size and shape, *Proc. of the Fifth World Congress on Particle Technology*, Orlando, FL, paper 240f.

Modeling and Control System Design of a Crystallizer Train for Para-xylene Production

Souichi Amano * **Genichi Emoto** * **Hiroya Seki** **

* *Production Technology Department, Mitsubishi Chemical Corporation, Kurashiki, 712-8054 Japan*
** *Chemical Resources Laboratory, Tokyo Institute of Technology, Yokohama, 226-8503 Japan*
(Corresponding author, e-mail: hseki@ pse.res.titech.ac.jp)

Abstract: A dynamic process model of an industrial crystallizer train for para-xylene production, which consists of five scraped surface crystallizers, two hydrocyclone separators, and two centrifugal separators, is developed for control system design. The model is identified by using real plant data. Optimal operating policies, which consider feed maximization and load distribution among the crystallizers, are derived, and multiloop controller is configured to realize the operating policy.

Keywords: Crystallization, industrial application, plantwide control

1. INTRODUCTION

Crystallization is one of the most popular unit operations for separation and purification used in the chemical industry. Despite their importance, process and control system design for separation processes based on crystallization technology has received much less attention compared with distillation column processes [Mendez et al. (2005)].

Most of the studies on control system design for crystallization processes focus on operations of a single crystallizer as an isolated unit [Rawlings et al. (1993), Hasebe et al. (2000)], although crystallizers never exist in isolation and simultaneous consideration of subprocessing units such as filtration and drying, etc. should be equally important [Chang et al. (1998)]. Some studies handle operations of multiple crystallizers [Garside (1985), Liu et al. (1991)], but study on process and control system design for crystallization processes from the plantwide perspective is quite limited [Wibowo et al. (2001), Ward et al. (2007)].

In this paper, modeling and control system design of an industrial crystallizer train, which comprises five crystallizers, two centrifugal separators and two cyclone separators, are discussed. The process concerns the product recovery section in a para-xylene production plant. The process underwent several revamps during the course of a long history of commercial operation, and the process became rather complicated, so that quantitative analysis based on a mathematical model would help provide us with improved operations.

First, a dynamic process model is developed, which describes crystallization kinetics, mass balance, and heat balance for the whole plant. The model is then identified by using the actual plant data. Based on the developed model, optimal operating policies are derived through optimization calculations. In setting up the optimization problem, practical considerations such as constraint handling issues are discussed in detail. Finally, a multi-loop control system is configured which realizes the optimal operation.

2. PROCESS DESCRIPTION

2.1 Para-xylene production process

Para-xylene is an aromatic hydrocarbon used primarily to make intermediates for manufacturing polyester. It is the main feedstock for purified terephthalic acid and dimethyl terephthalate, which in turn are used to produce polyethylene terephthalate (PET) for use in fibres for textiles, bottles for soft drinks and water.

Figure 1 shows a typical commercial production process of para-xylene, where fresh feed that contains mixed xylene (mixture of ortho-xylene, meta-xylene, and para-xylene) and ethylebenzene is sent from the upstream plant and pure para-xylene (normally $> 99.5\%$) is recovered from the feedstock by fractionation and crystallization. Crystallization is one of the conventional methods for the recovery of pure para-xylene; currently adsorption may be the most popular. Filtrate from the para-xylene recovery section is sent to the reaction section, where ortho-xylene, meta-

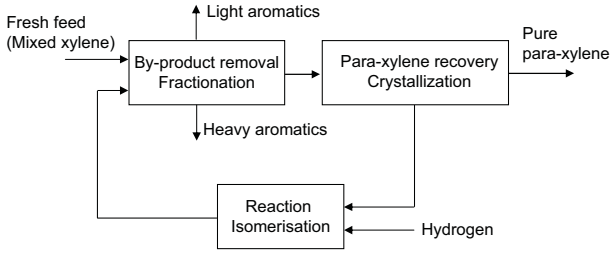


Fig. 1. Typical para-xylene production process

xylene are converted into para-xylene through isomerization reaction.

2.2 Para-Xylene recovery based on crystallization technology

Since many physical properties of the individual xylene isomers are similiar, high purity separation of each individual xylene isomer is difficult. Crystallization is one of the methods for that purpose, resorting to the differences in the melting points among the xylene isomers.

Figure 2 shows the crystallizer train under study, which consists of two separate cyrstallization stages. The first stage uses several (in this example, three) scraped surface crystallizers to bring the temperature of the xylene mixture close to the para-xylene/meta-xylene eutectic point. The first-stage crystallizers are followed by a solid-liquid separation process using a centrifuge (screen-bowl type in this example). The cold xylene filtrate from the first stage cools the feedstream (not shown in the figure) and is sent to the isomerization section. To achieve the maximum production rate from a certain feedstock, the first-stage crystallization temperature should be decreased as low as possible, down to the eutectic point of para-xylene and meta-xylene.

The solid cake para-xylene crystals and the adherent mother liquor from the first stage are melted in the melt tank, and pumped to the second-stage crystallizers. The second stage is made up of the main crystallizer, and the auxiliary crystallizer located on the recycle stream. The slurry from the main crystallizer is sent to the final centrifugal separator, and the filter cake is melted to form the final para-xylene product. A major portion of the mother liqour from the main crystallizer is returned to the first-stage after a part of para-xylene in the mother liquor is recovered by the auxiliary crystallizer.

Due to the presence of the recycle streams at several locations, which have been added during the course of a long history of commercial operation, the process becomes highly interacting, so that careful analysis on the basis of a mathematical model would be necessary in designing control system.

3. MODELING

3.1 Crystallizer

The crystallizers are assumed to be mixed-suspension mixed-product removal (MSMPR) systems. In addition, the following assumptions are made for model development:

- Only growth and nucleation are considered as crystallization kinetics; breakage and agglomeration are ignored.
- Para-xylene crystal growth is fast enough so that the liquid phase para-xylene is always saturated (the assumption of the high growth rate limit). Nucleation occurs at the crystallizer wall.

These assumptions are adopted from the study by Patience et al. (2001), who studied experimentally the crystallization kinetics of para-xylene in a scraped surface crystallizer.

Denoting the crystal size distribution (CSD) in the crystallizer as $f(x,t)$, its $i-$th moment μ_i is defined as

$$\mu_i = \int_0^\infty f(x,t)x^i dx.$$

By using the method of moment, the population balance equation can be written as:

$$\frac{d\mu_0}{dt} = B + \mu_0^{in} - \mu_0^{out},$$
$$\frac{d\mu_i}{dt} = iG\mu_{i-1} + \mu_i^{in} - \mu_i^{out} \qquad (i \geq 1),$$

where B and G are the nucleation rate and growth rate of para-xylene crystal respectively, μ_i^{in} and μ_i^{out} are the moment flows in and out of the crystallizers which can be calculated from the MSMPR assumption. The empirical expression is used for the nucleation rate B:

$$B = k_b \Delta C^b,$$

where ΔC is defined as the supersaturation created by the temperature difference between the magma and the crystallizer wall:

$$\Delta C = \frac{C^*(T) - C^*(T_J)}{C^*(T_J)}.$$

Here, T is the temperature of the magma, T_J is the temperature of the crystallizer wall, which is assumed to be equal to the jacket temperature, and $C^*(T)$ is the temperature dependent solubility of para-xylene.

The mass balance of the liquid phase para-xylene is written as

$$\frac{dm_{PX}}{dt} = F^{in}C^{in} - F^{out}C^*(T) - 3\rho k_v G\mu_2,$$

where m_{PX} is the liquid hold up of para-xylene in the crystallizer, F^{in} and F^{out} are the inlet and outlet liquid flow rates respectively, C^{in} is the para-xylene concentration of the inlet flow, ρ is the density of para-xylene crystals, k_v is the shape factor. Note that the liquid concentration of para-xylene in the crystallizer is assumed to be saturated.

The heat balance is written as

$$\frac{dH}{dt} = H^{in} - H^{out} + 3\rho k_v G\mu_2 \Delta H_c - UA(T - T_J),$$

where H is the overall enthalpy of the crystallizer, H^{in} and H^{out} are the enthalpy in and out of the crystallizer respectively, ΔH_c is the heat of crystallization, and UA is the overall heat transfer coefficient. Because of the fouling of the crystallizer wall, the heat transfer coefficient is treated as slowly time-varying.

The assumption of the high growth rate limit, that is, the growth rate of para-xylene crystals is so large that the liquid phase para-xylene concentration is always saturated, renders the model equations a DAE system; the growth rate is not explicitly given in the above equations. But the model equation can be easily converted into the ODE by the procedure shown by Patience et al. (2001).

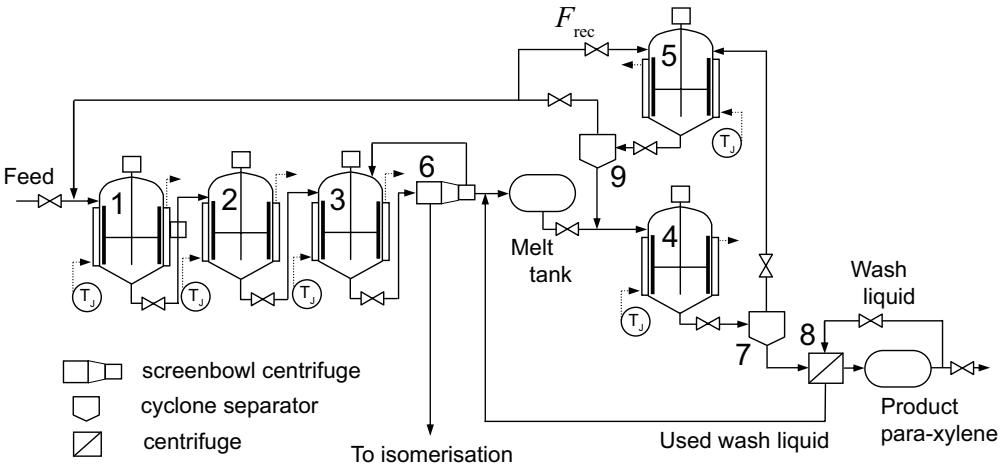


Fig. 2. Process flow of the para-xylene recovery section. The numbers ($1 \sim 9$) denote the equipment numbers which are used as subscript to distiguish the equipments.

3.2 Cyclone separator

The hydrocyclone separates the inlet slurry flow into two streams: the overflow and underflow streams. With the help of centrifugal force, the solid particles contained in the inlet stream are concentrated in the underflow. Ideally the overflow stream contains no solid particle, but it is practically assumed that some of the crystals whose size is smaller than $\bar{d}$ escape into the overflow stream. An ideal separation is assumed, where the crystals over the size $\bar{d}$ will not be included in the overflow. Crystals under the size $\bar{d}$ will be included both in the underflow and overflow, and they are distributed according to the liquid flow rates of these streams.

To obtain the amount of crystals smaller than $\bar{d}$, the crystal size distribution has to be recovered from its associated moment information. However, it is known that infinite number of the moments are needed to reconstruct the CSD [McGraw et al. (1998)].

To avoid this problem, the logarithmic normal distribution is assumed for the CSD. From the values of the moments $\mu_i's$, the mean crystal size m and variance σ^2 can be recovered from the relation:

$$\log(\mu_n/\mu_0) = \frac{n^2}{2}\sigma^2 + nm.$$

In this study, m and σ are determined through the least squares fit by using the moments up to the 4-th order.

Then the amount of crystals below the size $\bar{d}$ can be calculated as

$$\int_0^{\bar{d}} x^n f(x)dx = \exp(\frac{n^2\sigma^2 + 2nm}{2}) \cdot\frac{1}{2}\left(\mathrm{erf}(\frac{\log\bar{d} - m - n\sigma^2}{\sqrt{2}\sigma}) + 1\right),$$

where the error function is defined as

$$\mathrm{erf}(x) = \int_0^x \frac{2}{\pi}\exp(-\lambda^2)d\lambda.$$

No holdup is assumed for the cyclone separators. Then, the balance equations for the hydrocyclone are readily derived.

3.3 Centrifugal separator

At the centrifugal separators, it is assumed that the para-xylene crystals of the size smaller than $\hat{d}$ pass through the screen, accompanying the mother liquor. The amount of such crystals is calculated in the same way as in the hydrocyclone separator model.

Constant void fraction is assumed for the filter cake ($\varepsilon = 0.4$), and the average degree of saturation S_{av} (the percentage of the void in the cake filled with mother liquor) is assumed to be a function of the average crystal size $d_{23} = \mu_3/\mu_2$:

$$S_{av} = S_{av}(\hat{S}, d_{23}),$$

where $\hat{S}$ is a parameter to define the empirical expression.

When the cake is washed (as in the second stage centrifuge), part of the mother liquor in the cake is replaced by the wash liquor. The percentage of the mother liquor replaced by the wash liquid is expressed by the empirical expression, which is a function of S_{av} and the ratio of the amounts of the wash liquid and the mother liqour. The amount of the remaining mother liquor in the cake largely accounts for the product purity.

Then, the balance equations for the centrifuge are readily derived; no holdup is assumed.

The screen-bowl type centrifugal separator at the outlet of the 1st stage is modeled as a combined system comprising a cyclone and a centrifuge; the bowl part is modeled as the cyclone. The parameter α is introduced as the ratio between the bowl filtrate (overflow stream) and the liquid inlet, which will be used for model identification in the next section.

3.4 Overall process model

By combining the models for the crystallizers, the hydrocyclone separators, the centrifugal separators and other

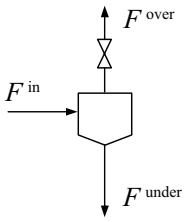


Fig. 3. Hydrocyclone separator modeling

storage tanks, a nonlinear dynamic process model in the following form is derived:

$$\begin{aligned} \dot{x} &= f(x,u,p) \\ y &= h(x,u,p), \end{aligned} \tag{1}$$

where x is the state variable, u is the measured independent variables, p is the unknown parameters and unmeasured independent variables, and y is the measured variables. Here, the variables are defined for the purpose of model identification.

3.5 Model identification

The unknown parameters are obtained through least squares fit of the model calculation with the plant data:

$$\min_{x,p} \quad (\tilde{y}-y)^T(\tilde{y}-y) \tag{2}$$

subject to

$$0 = f(x,u,p), \tag{3}$$

where $\tilde{y}$ is the plant data. Eq. (3) assumes that model identification is done for steady states.

Several data sets for $(\tilde{y},u)$, which have been obtained by heavily filtering 1 hour average data from the real plant, are used for the least squares fit. The available measurements $\tilde{y}$ consist of the holdups of the five crystallizers, the holdups of the melt tank and product tank, the temperatures of the crystallizers, the production rate, the recycle flow rate to the isomerization reaction, para-xylene concentrations of the melt tank and product tank.

As a result of the sensitivity analysis of the minimization problem (2), the identifiable parameter set $p \in \Re^6$ has been selected as

$$p = \left(\alpha_6 \ \bar{d}_6 \ \hat{d}_6 \ \hat{S}_6 \ \hat{S}_8 \ \bar{d}_9\right)^T,$$

where the subscript are defined as equipment number in Fig. 2. The overall heat transfer coefficient UA_i of each crystallizer ($i = 1 \sim 5$) is considered as time-varying and is also used for model identification as the fitting parameters. It should be noted that the parameters concerning crystallization kinetics such as b and k_b are not identifiable from the available measurements, so that their values are adopted from the paper by Patience et al. (2001).

Figure 4 shows one of the fitting results: the parity plot of the fitting result for the production rate. The operation condition used for model identification covers ±20% of the nominal production rate.

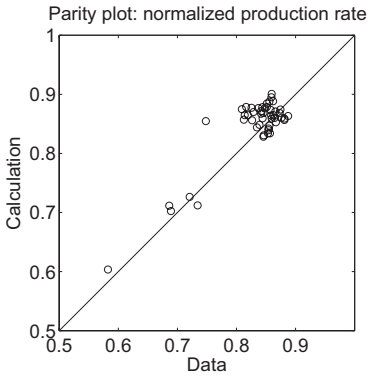


Fig. 4. Parity plot of production rate measurements

4. CONTROL SYSTEM DESIGN

4.1 Definition of manipulated and controlled variables

By using the identified process model, basic regulatory control system is designed; the seven inventory control loops are closed with each effluent flow, and the temperature control loop of each crystallizer is closed by manipulating its corresponding jacket temperature.

The manipulated variables for further control system design are defined as $u_C \in \Re^9$, and they consist of the temperature setpoints of the crystallizers, the ratios of the flow rates of the overflow and the underflow of the cyclones, the wash liquid flow rate, and the recycle flow rate from the overflow of the cyclone to the auxiliary crystallizer:

$$u_C = \left(T_1 \ T_2 \ T_3 \ T_4 \ T_5 \ \alpha_7 \ \alpha_9 \ F_w \ F_{rec}\right)^T,$$

where the subscripts are defined as the equipment number in Fig. 2 and α_i ($i = 7, 9$) is the flow rate ratio in the cyclone separators, and F_w is the wash flow rate, and F_{rec} is the recycle flow rate.

The controlled variables are defined as y_C, for which constraints may be considered, and the process model for control system design is described as

$$\begin{aligned} \dot{x} &= f_C(x,u_C) \\ y_C &= h_C(x,u_C). \end{aligned} \tag{4}$$

4.2 Steady state optimal operation policy

In deriving optimal operating policies, the following constraints are considered.

- *Lower limit for the para-xylene purity x_{prod}* This is a product specification. The purity is determined by the amount of accompanying mother liquor, which is affected by the average crystal size (the larger, the better) and the intensity of the wash at the centrifuge.
- *Lower limits for the jacket temperatures at the 1st stage crystallizers $T_{J,1} \sim T_{J,3}$* The yield of the para-xylene recovery section is determined by how low the 1st stage crystallizer temperature can be reduced. The refrigerator capacity determines the lower limits of the jacket temperatures.
- *Upper limits for the temperature difference between the jacket and crystallizer $\Delta T_1 \sim \Delta T_5$* One of the

major operational concerns is the fouling of the crystallizer wall, which is caused by the crystal deposition on the wall surface and exacerbated by too high a super-saturation at the crystallizer wall. Para-xylene crystal deposition on the wall results in poor heat transfer and limits the production rate.

- *Upper limits for the slurry concentrations in the second stage crystallizers* Cs_4, Cs_5 The slurry concentrations in the second stage crystallizers tend to be high and they are limited by the torque limit of the agitator. If the slurry concentration is too high, mixing in the crystallizer would become imperfect.
- *Upper and lower flow rate limits for the cyclone separators* Operation of the cyclone in an abnormal flow rate regime results in inappropriate classification of crystals.
- *Upper limits for the slurry concentrations in the underflow of the cyclone separators* Cs_7, Cs_9 The slurry concentration in the underflow of the cyclone separator tends to be large. Too high a slurry concentration results in clogging of the pipe.
- *Upper limit for the para-xylene concentration in the melt tank* If this concentration is too high, some of the solid para-xylene with low purity from the 1st stage do not dissolve in the melt tank.

The following two modes of operations are considered for developing optimal operating policies.

Feed maximization Feed maximization is realized by solving the following optimization problem:

$$\begin{aligned} \max_{u_C} \quad & F_p \\ \text{subject to} \quad & \\ & 0 = f_C(x, u_C) \\ & y_C^{LL} \le y_C \le y_C^{UL} \\ & u_C^{LL} \le u_C \le u_C^{UL}, \end{aligned} \tag{5,6}$$

where $(\cdot)^{LL}$ and $(\cdot)^{UL}$ are the lower limits and upper limits respectively; F_p is the production rate that is defined as the effluent of the product tank minus the wash liquid.

Prospective active constraints are found to be the lower limits of the jacket temperatures of the 1st stage crystallizers, the upper limits for the temperature differences and slurry concentrations of the second stage crystallizers, and the slurry concentration in the underflow of the cyclone on the recycle stream. It has been found that around 2% increase in the production rate could be possible compared with the conventional operation.

Load distribution Load for the crystallizer is expressed as the temperature difference (ΔT) between the crystallizer and the jacket. For a prescribed production rate $\bar{F}_p$, flexible operation by distributing the loads between the two crystallizers at the second stage would be advantageous; when fouling of the crystallizer wall of one of the crystallizers is severe, which situation may be observed by decrease in the heat transfer coefficient, the load for that crystallizer is lowered while the load for the other crystallizer is increased to keep the production rate. For such cases, the following optimization problem can be conceived:

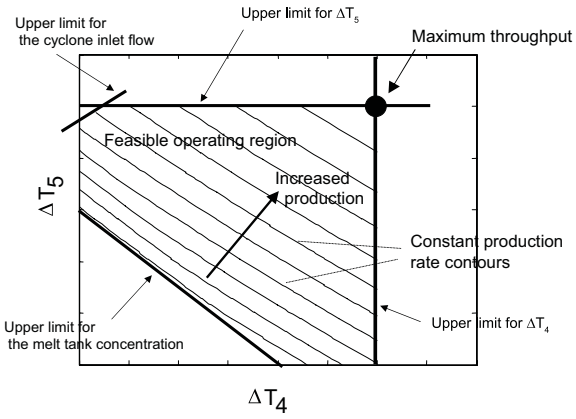


Fig. 5. Optimization landscape described as a contour plot of production rate

$$\begin{aligned} \min_{u_C} \quad & w\Delta T_4 + (1-w)\Delta T_5 \\ \text{subject to} \quad & \\ & 0 = f_C(x, u_C) \\ & F_p = \bar{F}_p \\ & y_C^{LL} \le y_C \le y_C^{UL} \\ & u_C^{LL} \le u_C \le u_C^{UL}, \end{aligned} \tag{7,8}$$

where w ($0 \le w \le 1$) is the weight used for distributing the load.

Figure 5 shows the optimization landscape obtained by solving the minimization problem (7) for various values of production rate $\bar{F}_p$ and weight w, in which the load distribution for a fixed production rate is described as a contour plot in the $\Delta T_4 - \Delta T_5$ plane. The maximum throughput is realized when the upper limit constraints for both of the temperature differences become active, whereas the throughput is decreased for a moderate load where there is room for the temperature difference. The lower limits for the jacket temperatures of the 1st stage crystallizers and the upper limits for the slurry concentrations in the 2nd stage crystallizers and the underflow of the cyclone are always active.

4.3 Optimizing control

To realize the optimal operating policies derived in the previous subsection, a 6×6 multi-loop control is configured. As the manipulated and controlled variables, the following variables are selected:

Manipulated: T_4, T_5, α_7, α_9, F_w, F_{rec}
Controlled: Cs_4, ΔT_4, Cs_5, ΔT_5, Cs_9, x_{prod},

where Cs_4, Cs_5 and Cs_9 are the slurry concentrations of the 2nd stage crystallizers and the underflow of the cyclone separator on the recycle stream, x_{prod} is the para-xylene concentration in the product stream.

Constant setpoints are given to Cs_4, Cs_5, Cs_9, and x_{prod}, because constraints for these variables are known to be always active with the optimal operations, while the setpoints to ΔT_4 and ΔT_5 are varied according to the load distribution policy. Care should be taken in giving setpoints to ΔT_4 and ΔT_5, because other constraints for

Table 1. RGA analysis for the multi-loop controller design

	T_4	T_5	α_7	α_9	F_w	F_{rec}
Cs_4	0.0093	0.33	0.81	-0.19	0.041	0
ΔT_4	0.85	-0.33	0.18	-0.093	0.39	0
Cs_5	-0.057	0.055	0	0.0028	0	1.0
ΔT_5	-0.30	0.73	-0.0024	0.57	0.0074	-0.0027
Cs_9	-0.0034	0.33	0.0031	0.67	0.0022	0.0019
x_{prod}	0.5	-0.12	0.01	0.038	0.56	0.0

such variables as the cyclone inlet flow rate and melt tank concentration may become active.

A pairing of these variables in the multi-loop control system is determined through the relative gain array (RGA) analysis [Bristol (1966)] shown in Table 1.

5. SIMULATION STUDY

Figure 6 shows a simulation result of the designed control system when the setpoints of the temperature differences ΔT_4 and ΔT_5 are changed (all the numerical values are eliminated from the plot to keep any proprietary information confidential). For the first half of the simulation, the temperature difference ΔT_5 is changed stepwise, while the temperature difference of the other crystallizer ΔT_4 is held constant. This operation increases the production rate. For the second half, ΔT_4 is decreased stepwise while ΔT_5 is held constant. As a result, the load of the crystallizer 4 is reduced while the load of the crystallizer 5 is increased, compared with the initial state of the simulation.

Toward the end of the simulation when the load on the crystallizer 4 is reduced, the cyclone inlet flow almost hits the upper limit, which is anticipated from the analysis shown in Fig. 5.

6. CONCLUSION

A process model of an industrial crystallizer train for para-xylene recovery has been developed and a multi-loop control system has been configured.

Since the process is highly interactive due to the existence of the recycle streams, and the active constraints are subject to change depending upon operating conditions as shown in Fig. 5, application of multivariable model predictive control with constraint handling capability may be justified, if override control logic is regarded tedious.

One of the major concerns in the crystallizer operations is fouling of the crystallizer wall due to large supersaturation, which leads to decreased heat transfer and production rate. A monitoring and control system which is capable of identifying the deteriorating heat transfer coefficient to automatically adjust ΔT would be helpful.

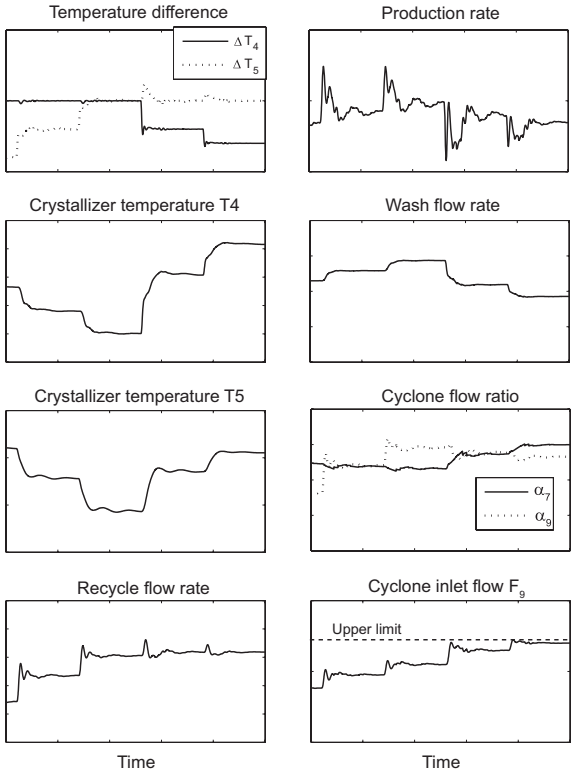


Fig. 6. Simulation result

REFERENCES

E.H. Bristol. On a new measure of interactions for multivariable process control. *IEEE Trans.*, volume AC-11, pages 133-134, 1966.

W.C. Chang and K.M. Ng. Synthesis of processing system around a crystallizer. *AIChE Journal*, volume 44, pages 2240–2251, 1998.

J. Garside. Industrial crystallization from solution. *Chem. Eng. Sci.*, volume 40, pages 3–26, 1985.

S. Hasebe, T. Hamamura, K. Naito, K. Sotowa, M. Kano, I. Hashimoto, H. Betsuyaku, and H. Takeda. Optimal operation of a continuous DTB crystallizer. *J. Process Control*, volume 10, pages 441–448, 2000.

C.H. Liu, D.H. Zhang, C.G. Sun, and Z.Q. Shen. The modelling and simulation of a multistage crystallizer. *The Chemical Engineering Journal*, volume 46, pages 9–14, 1991.

R. McGraw, S. Nemesure, and S.E. Schwartz. Properties and evolution of aerosols with size distributions having identical moments. *J. Aerosol Sci.*, volume 29, pages 761-772, 1998.

C.A. Mendez, J. Myers, S. Roberts, J. Logsdon, A. Vaia, and I.E. Grossmann. MINLP model for synthesis of paraxylene separation processess based on crystallization technology. *Proc. ESCAPE 15*, pages 829–834, Barcelona, Spain, 2005.

D.B. Patience, J.B. Rawlings, and H.A. Mohameed. Crystallization of para-xylene in scraped-surface crystallizers. *AIChE Journal*, volume 47, pages 2441–2451, 2001.

J.B. Rawlings, S.M. Miller, and W.R. Witkowski. Model identification and control of solution crystallization processes: a review. *Ind. Eng. Chem. Res.*, volume 32, pages 1275–1296, 1993.

J.D. Ward, C.C. Yu, and M.F. Doherty. Plantwide operation of processes with crystallization. *AIChE Journal*, volume 53, pages 2885–2896, 2007.

C. Wibowo, W.C. Chang, and K.M. Ng. Design of integrated crystallization systems. *AIChE Journal*, volume 47, pages 2474–2492, 2001.

Evaluation of the Effect of the Solubility Model on Antisolvent Crystallization Optimization

D. J. Widenski*, A. Abbas, J. A. Romagnoli***

**Chemical Engineering Department, Louisiana State University, Baton Rouge, LA 70803 USA (e-mail:dwiden1, jose@lsu.edu)*
***School of Chemical and Biomolecular Engineering, University of Sydney, Sydney, NSW 2006 Australia (e-mail: alia@usyd.edu.au)*

Abstract: The use of predictive solubility models can be of great use for crystallization modeling, and can decrease the amount of experimental data needed to create a robust crystallization model. In this paper, predictive solubility models such as MOSCED, UNIFAC, NRTL-SAC, and the Jouyban-Acree model are compared against an empirical model for predicted solubility accuracy. The best models are subsequently compared against the empirical model for the antisolvent crystallization of acetaminophen in acetone using water. Two different optimization objective functions are executed for each solubility model to generate corresponding optimal profiles. The effect of these optimal profiles on the predicted crystal properties is evaluated.

Keywords: Optimization, Solubility Model, Crystallization, Antisolvent, NRTL-SAC, Jouyban-Acree, Population Balance

1. INTRODUCTION

Crystallization is an important unit operation for the production of pharmaceuticals, fertilizers, and fine chemicals. Optimal crystallization operation often requires a crystallization model. This crystallization model, typically based on population balances (Hulburt and Katz, 1964; Ramkrishna, 1985; Randolph and Larson, 1988), requires a companion solubility model. Empirical solubility models have been extensively used in crystallization modelling (Zhou et al., 2006; Nowee et al., 2008; Lindenberg et al., 2009). It is of interest to understand how other predictive solubility models such as the MOSCED, NRTL-SAC, UNIFAC, and Jouyban-Acree models can be incorporated into crystallization models and how their accuracy of predicting the solubility profiles influences both the crystallization model prediction and optimal profile calculation. The outcome of combining predictive solubility modelling with the crystallization model is expected to reduce the need for solubility experimental data and consequently streamline the optimization of the crystallization process.

The solubility prediction is an important aspect of any crystallization model because its prediction is the basis for any crystallization phenomena. Crystallization is caused by supersaturation which is defined as the difference between the solution concentration and the equilibrium concentration (absolute supersaturation), or the ratio of the solution and equilibrium concentrations (relative supersaturation).

This paper investigates the effect of different solubility models on the optimization of antisolvent crystallization. Both the effect of the model on the predicted optimal profile and on the result of these models' optimal profiles implemented into a validated crystallization model will be evaluated. Specifically, we examine the effect on the supersaturation, mean size, and volume percent crystal size distribution (CSD) profiles. Although there has been extensive work done in the area of crystallization control and optimization (Braatz, 2002; Zhou et al., 2006; Nowee et al., 2008; Sheikhzadeh et al., 2008), as far as we are aware there is no study that has investigated the use of predictive solubility models in developing optimal antisolvent feed profiles.

2. SOLUBILITY MODELS

2.1 MOSCED & UNIFAC Models

The MOSCED model (Lazzaroni et al., 2005), generates infinite dilution activity coefficients. In order to obtain a non-infinite dilution activity coefficient, another activity coefficient model is required. The Van Laar, Wilson, and NRTL models were each combined with the MOSCED model to evaluate which would give the best prediction to known experimental data. The next solubility model considered is the UNIFAC model (Anderson and Prausnitz, 1978). The UNIFAC model predicts activity coefficients based on group contributions. The MOSCED and UNIFAC models predicted equilibrium profiles for acetaminophen in acetone and water are shown in Figure 1.

The MOSCED models all give very poor solubility predictions. They all greatly underestimate the solubility. The NRTL and Wilson models give better estimates to the shape of the solubility curve than the Van Laar model does. The UNIFAC model is the worst of the models both greatly overestimating the solubility and weakly representing the shape of the curve.

2.2 NRTL-SAC, Jouyban-Acree, and Empirical Models

The next solubility model considered is the NRTL-SAC model (Chen et al., 2004, 2006). The NRTL-SAC model is a NRTL activity coefficient model that has been modified using segment theory (Chen et al., 2004, 2006). The last predictive solubility model considered is the Jouyban-Acree model (Jouyban et al., 2006).

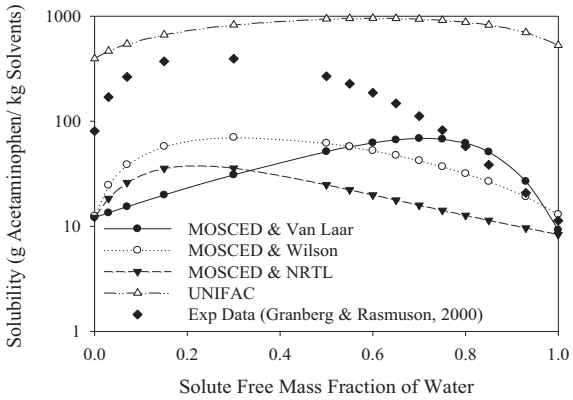


Figure 1: MOSCED & UNIFAC Solubility Predictions.

The Jouyban-Acree model is a semi-empirical model developed to predict the solubility of pharmaceuticals in organic solutions. This model requires the solubilities of both pure components in a binary solute-solvent system, and predicts the solubility of a solute in a solvent mixture. The last solubility model considered is an empirical model generated from data from Granberg and Rasmuson (2000) by Zhou et al. (2006).

$$C^* = -5.01902 \text{ x } 10^{-12} w^6 + 1.69767 \text{ x } 10^{-9} w^5 - 2.46765 \text{ x } 10^{-7} w^4 + 2.19262 \text{ x } 10^{-5} w^3 - 1.27018 \text{ x } 10^{-3} w^2 + 3.42614 \text{ x } 10^{-2} w + 7.96086 \text{ x } 10^{-2} \quad (1)$$

Where C^* is the equilibrium concentration (kg acetaminophen/kg solvents), and w is the solute free mass percent of water. The NRTL-SAC, Jouyban-Acree, and empirical model predicted solubilities are plotted in Figure 2. The NRTL-SAC and Jouyban-Acree solubility models both predict the equilibrium solubility much better than the MOSCED or UNIFAC models did. The empirical model fits the data very well and will be considered as the standard solubility model for benchmarking. Since the UNIFAC and MOSCED models gave such poor solubility predictions, only the NRTL-SAC and Jouyban-Acree models will be compared against the empirical model in the optimization sensitivity study in the subsequent sections.

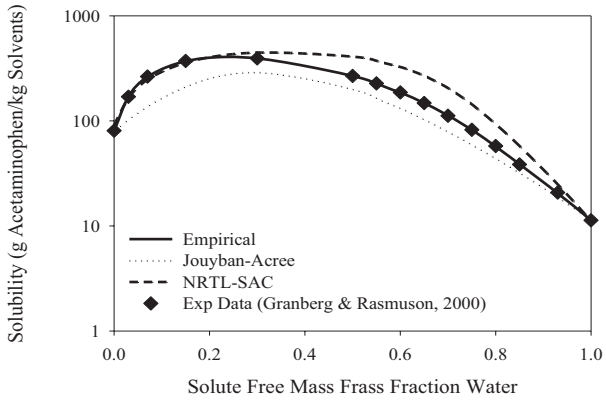


Figure 2: NRTL-SAC and Jouyban-Acree Solubility Predictions.

3. CRYSTALLIZATION MODEL

In order to evaluate the effect of the solubility model on the predicted optimal trajectories, a crystallization model is required.

3.1 Population Balance

The population balance for a crystallization system having size-independent crystal growth and without attrition or agglomeration is defined in (2).

$$\frac{\partial n(L,t)}{\partial t} + \frac{n(L,t)}{V}\frac{dV}{dt} + G\frac{\partial n(L,t)}{\partial L} - B = 0 \quad (2)$$

Where $n(L,t)$ is the crystal density (# of particles/m^3), V is the volume (m^3), G is the growth rate (m/s), and B is the nucleation rate (# of particles/ s m^3). The population balance was solved by discretization using backward finite differences. The discretization consisted of 250 geometrically spaced intervals from 0.5-1000 microns.

3.2 Crystallization Kinetics

The antisolvent crystallization kinetics for acetaminophen in acetone with water as the antisolvent were taken from Zhou et al (2006). The authors developed their own kinetic rates (3-6), from previous crystallization data performed by Granberg et al. (1999, 2001).

$$N = 8.56080 \text{ x } 10^8 \text{ x } \exp\left\{-1.22850 \text{ x } 10^{-3} \frac{\ln^3\left(\frac{\rho_c}{C^*\rho_s}\right)}{\ln^2\left(\frac{C}{C^*}\right)}\right\} \quad (3)$$

$$G = k_g(C - C^*)^g \quad (4)$$

$$k_g = 4.01067 \text{ x } 10^{-8} w^2 - 1.76198 \text{ x } 10^{-6} w + 5.78135 \text{ x } 10^{-5} \quad (5)$$

$$g = -4.22536 \text{ x } 10^{-3} w + 1.77428 \quad (6)$$

Where N is the nucleation rate (no. of particles/m^3), ρ_c is the crystal density of acetaminophen (kg/m^3), C^* is the equilibrium concentration defined previously, C is the solution concentration (kg acetaminophen/kg solvents), ρ_s is the density of the solution (kg/m^3), G is the crystal growth rate (m/s), and w is the solute free mass percent of antisolvent (water) in the solution. Also, the growth kinetics is only valid for solute free water mass percents greater than 30%.

4. OPTIMIZATION

The first optimization objective (O-1) was to create a final volume mean crystal size (D_{43}) of 200 microns, and jointly minimize the total amount of nucleation by minimizing the zeroth moment. The optimization constraints were to end with a solute free antisolvent mass percent of water of 88%, and the mass flow rate of water could range between 0 and 400 g/min. The duration of the experiment was fixed at 4200 s. The control interval was discretized into 10 fixed 360 s intervals where the antisolvent flow rate could be adjusted in a piecewise constant manner. The final 600 s had a fixed antisolvent flow rate of zero. This was done to ensure that all remaining supersaturation is consumed at the end of the run. The optimizations were implemented using the gPROMS package (Process System Enterprise, UK) using the gOPT entity. The objective function used is defined in (7) subject to initial conditions in (8).

$$OBJ = \begin{matrix}\max D_{43} \\ \min \mu_0\end{matrix} \begin{cases} D_{43} \leq 200\ \mu\text{m} \\ w_f = 88\% \\ 0 \leq \dfrac{dW}{dt} \leq 400\ g/\min \text{Water} \end{cases} \quad (7)$$

$$T = 16\ ^\circ\text{C} \quad w_i = 40\% \quad n_i(L,0) = 0 \quad C_i = C_i^* \quad (8)$$

This optimization was carried out using the crystallization model in Section 3 separately with each of the empirical, Jouyban-Acree, and NRTL-SAC solubility models. The MOSCED and UNIFAC models were not considered because when those models where incorporated into the crystallization model they did not predict any crystallization phenomena.

4.1 Optimal Antisolvent Feed Profiles for O-1.

Each solubility model resulted in an optimal profile (Figure 3).. The empirical and Jouyban-Acree models generated similar optimal profiles (denoted Profile A.1 and Profile B.1 respectively) with a small initial flow rate at the beginning of the experiment, moderate flow rate in the middle, and higher flow rate at the end. In contrast, the NRTL-SAC model calculates an optimal profile (denoted Profile C.1) that has a moderate initial flow rate followed by a high flow rate in the middle, and no flow at the end.

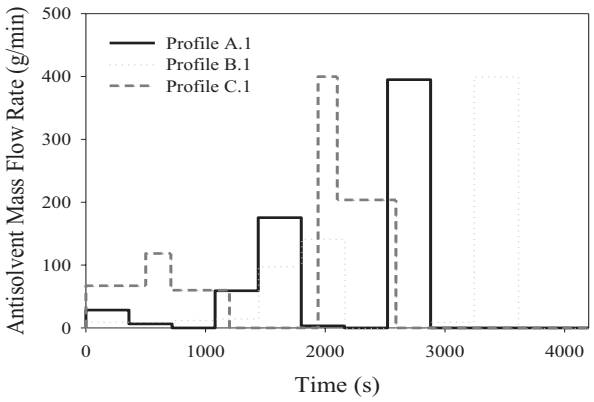


Figure 3: Optimal antisolvent feed profiles for O-1.

The second optimization objective (O-2) was to create a larger final volume mean size (D_{43}) of 400 microns while again minimizing the total amount of nucleation by minimizing the zeroth moment. The objective function formulation for O-2 was the same as for O-1 with the exception that D_{43} now cannot be greater than 400 microns.

4.2 Optimal Antisolvent Feed Profiles for O-2.

Each solubility model resulted in a new optimal profile for O-2 (Figure 4). The empirical and Jouyban-Acree models again generated similar optimal profiles (denoted Profile A.2 and Profile B.2 respectively) with a small initial flow rate at the beginning of the experiment, a high flow rate in the middle for A.2, and a high flow rate at the end for B.2. In contrast, the NRTL-SAC model calculates an optimal profile (denoted Profile C.2) that has a moderate initial flow rate followed by a low flow rate in the middle, and a moderate flow rate at the end.

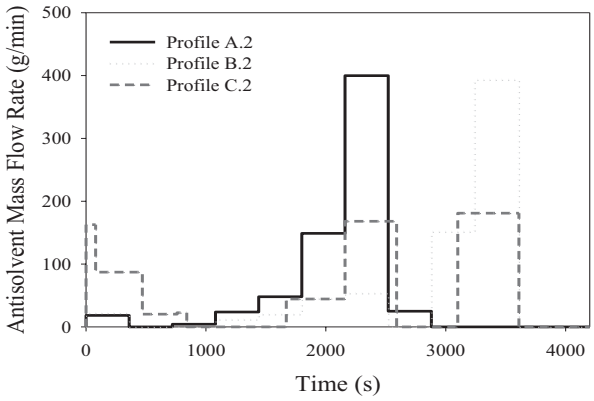


Figure 4: Optimal antisolvent feed profiles for O-2.

5. OPTIMIZATION SENSITIVITY ANALYSIS

The crystallization model was executed for each generated optimal feed profile (A.1-C.2) using the empirical solubility model. The empirical model is used as the benchmark since it

showed very close agreement to experimental solubility data. This should predict what these optimal profiles would actually produce in a real crystallizer. Results are shown in the next sections.

5.2 Optimal Profiles for O-1 Evaluation

When the optimal profiles are implemented into the empirical solubility model there are several observed differences in the simulated supersaturation profiles (Figure 5) under Profiles A.1, B.1, and C.1. The NRTL-SAC optimal profile (C.1) causes the supersaturation to peak earlier than the other two models, while the supersaturation caused by the Jouyban-Acree profile (B.1) is shown to be similar in shape to the empirical profile (A.1) but with a delay. Next, the effect on the volume mean size growth is shown in Figure 6.

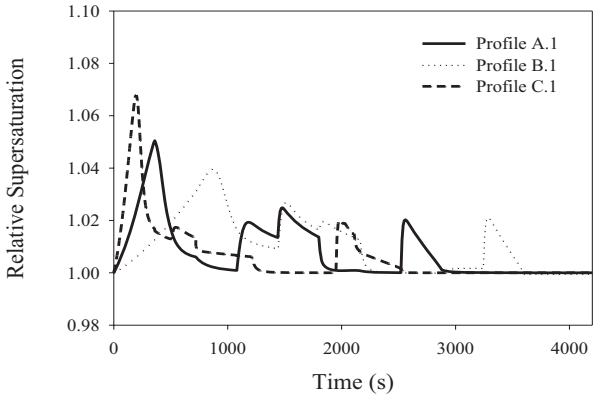


Figure 5: Relative supersaturation profiles for each optimal antisolvent feed profile for O-1.

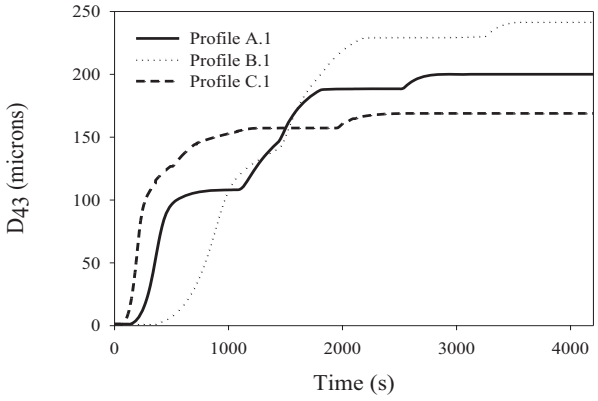


Figure 6: Volume mean size for each optimal antisolvent feed profile for O-1.

Each volume mean size profile can be explained by its corresponding supersaturation curve. Whenever the supersaturation increases there is a corresponding increase in the volume mean size. The NRTL-SAC optimal profile's (C.1) supersaturation profile only has one large early supersaturation peak which causes the first primary increase in crystal size, and a second peak which causes a small increase in crystal size. The empirical optimal profile's (A.1) generated supersaturation profile has four peaks which cause four increases in crystal size. Likewise the Jouyban-Acree optimal profile (B.1) causes the supersaturation profile to have four increases in D_{43}. Using the empirical solubility model, the empirical optimal profile (A.1) satisfies its objective of 200 microns, the Jouyban-Acree optimal profile (B.1) is higher at 242 microns, and the NRTL-SAC optimal profile (C.1) is lower at 169 microns. Both predictive models optimal profiles did not meet the optimization objective but are within 20% of the desired value.

Figure 7 shows the volume percent CSD for each optimal profile. All three optimal profiles give similar distributions with the NRTL-SAC optimal profile (C.1) distribution having a lower mean size than the others, and the Jouyban-Acree optimal profile (B.1) distribution having a larger mean size. All three optimal profiles generated distributions with similar width.

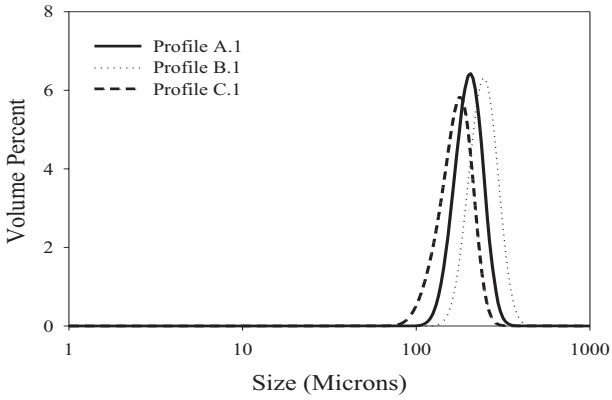


Figure 7: Volume percent CSD for each optimal antisolvent feed profile for O-1.

For this objective function (O-1) only the empirical model's optimal profile (A.1) was able to satisfy the objective to create a volume mean size of 200 microns, but the predictive models' profiles (B.1 & C.1) were able to be within 20% of the desired value. Also, all three profiles were successfully able to suppress nucleation to produce unimodal profiles.

5.3 Optimal Profiles for O-2 Evaluation

The next objective function considered is the 400 volume mean size objective function (O-2). As seen in Figure 8 the generated supersaturation profiles follow the same trend as for the first objective function (O-1). The NRTL-SAC optimal profile (C.2) generates a supersaturation profile that is nearly identical to the supersaturation profile that C.1 generated for O-1. The empirical optimal profile (A.2) generates a supersaturation amount that is above 1.02 from 500 to 2500 s. The Jouyban-Acree optimal profile (B.2) generates a supersaturation peak that is similar to (A.2) but not as high of a supersaturation amount.

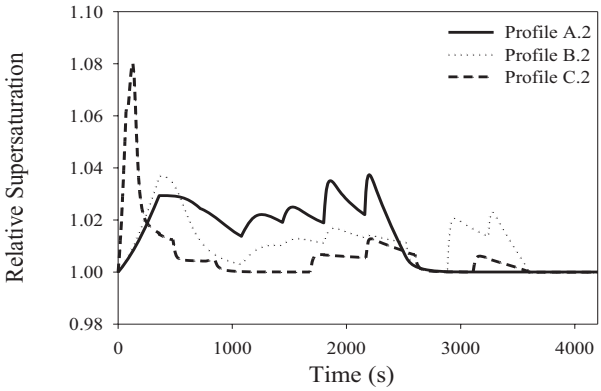


Figure 8: Relative supersaturation profiles for each optimal antisolvent feed profile for O-2.

The Jouyban-Acree (B.2) and NRTL-SAC (C.2) optimal profiles both generated a much smaller mean size because they did not generate the required supersaturation. The Jouyban-Acree optimal profile (B.2) generated a volume mean size of 271 microns and the NRTL-SAC optimal profile (C.2) generated a volume mean size of 162 microns. Both predictive solubility models' optimal profiles do not satisfy O-2 as well as they satisfied O-1.

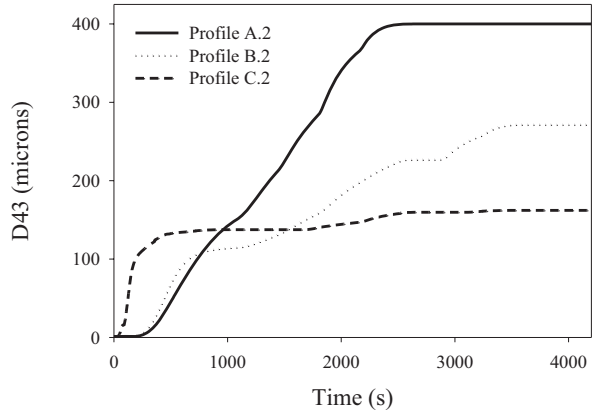


Figure 9: Volume mean size for each optimal antisolvent feed profile for O-2.

Figure 10 shows the volume percent CSD for each optimal profile for O-2. For this case there is a larger difference between the three profiles. Clearly, the Jouyban-Acree (B.2) and NRTL-SAC (C.2) optimal profiles did not satisfy the optimization objective. Also, the distribution width had more variation between the three profiles. The empirical profile (A.2) had the lowest distribution width, followed by the Jouyban-Acree (B.2) and the NRTL-SAC (C.2) model had the largest distribution width.

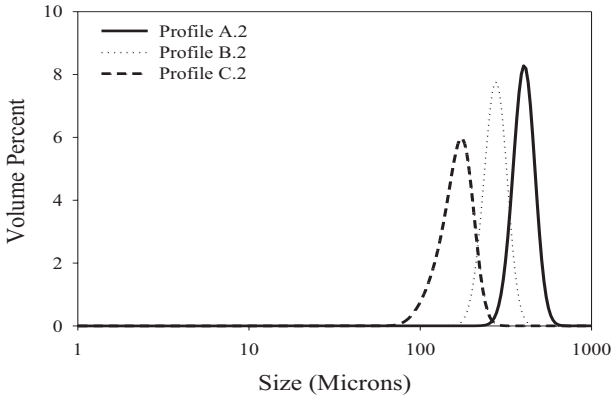


Figure 10: Volume percent CSD for each optimal antisolvent feed profile for O-2.

Just as for the first case (O-1) only the empirical model's optimal profile (A.2) was able to satisfy the objective (O-2) to create a volume mean size of 400 microns. Both predictive model profiles (B.2 & C.2) produced a much smaller mean size. The Jouyban-Acree profile (B.2) produced particles 32% smaller, and the NRTL-SAC profile (C.2) produced particles 60% smaller. Even though they did not produce the proper volume mean size, all three profiles were successfully able to suppress nucleation to produce unimodal profiles.

5.5 Optimization Evaluation

The reason why the optimal flow rates are similar for both the empirical and Jouyban-Acree model is that the slopes of both solubility curves are very similar. Since the slope of the solubility curve is what dictates the supersaturation profile, it would be expected to give similar supersaturation profiles. The NRTL-SAC model has a different slope in its solubility profile which causes the larger deviation in these reported results. In order for a predictive solubility model to produce predictive optimal profiles it must be accurate both quantitatively and qualitatively.

Table 1: Final volume mean crystal size derived from each optimal profile.

	Final Volume Mean Size & Percent Error			
Optimal Feed Profile	O-1 (200)	Prediction Percent Error	O-2 (400)	Prediction Percent Error
Empirical	200	0	400	0
Jouyban-Acree	242	21%	271	-32%
NRTL-SAC	169	-16%	162	-60%

Only the optimal profiles (A.1, A.2) generated from the empirical solubility model were able to satisfy both optimization objectives. When other optimal profiles were used the final volume mean size was as much as 60% under predicted and 21% over predicted when implemented into the

empirical solubility model. The deviation from the objective criteria increased as the volume mean size increased.

6. CONCLUSION

The effect of several solubility models were evaluated on the predicted optimal antisolvent feed profiles. The solubility model did have an effect on the optimal profile, and generated a unique optimal antisolvent feed profile. The use of the predictive solubility models' optimal profiles in the empirical solubility model did not satisfy the original objective function. The use of an incorrect solubility model will create a sub-optimal antisolvent feed profile that will not satisfy its intended crystallization optimization objective in a real system. This underpins the significance of the solubility profile in crystallisation optimizations work.

REFERENCES

Anderson T.F., Prautsnitz J.M. (1978), Application of the UNIQUAC Equation to Calculation of Multicomponent Phase Equilibria 2: Liquid-Liquid Equlibria, *Industrial & Engineering Chemistry Process Design and Development*, 17-4, 561-567

Braatz R.D. (2002), Advanced Control of Crystallization Processes, *Annual Reviews in Control*, 26, 87-99.

Chen C.C., Crafts P.A. (2006), Correlation and Prediction of Drug Molecule Solubility in Mixed Solvent Systems with the Nonrandom Two-Liquid Segment Activity Coefficient (NRTL-SAC) Model, *Industrial & Engineering Chemistry Research*, 45, 4816-4824.

Chen C.C., Song S. (2004), Solubility Modelling with a Nonrandom Two-Liquid Segment Activity Coefficient Model, *Industrial & Engineering Chemistry Research*, 43, 8354-8362.

Granberg R. A., Bloch D. G., Rasmuson A. C. (1999), Crystallization of Paracetamol in Acetone-Water Mixtures, *Journal of Crystal Growth, 198/199*, 1287-1293.

Granberg R. A., Ducreux C., Gracin S., Rasmuson A. C. (2001), Primary Nucleation of Paracetamol in Acetone-Water Mixtures, *Chemical Engineering Science*, *56*, 2305-2313.

Granberg R.A., Rasmuson A.C. (2000), Solubility of Paracetamol in Binary and Ternary Mixtures of Water, *Journal of Chemical & Engineering Data*, 45, 478-483.

Hulbert H.M., Katz S. (1964), Some Problems in Particle Technology, *Chemical Engineering Science,* 19, 555.

Jouyban A., Chan H.K., Chew N.Y.K., Khoubnasabjafari M., Acree, Jr. W.E. (2006), Solubility Prediction of Paracetamol in Binary and Ternary Solvent Mixtures Using Jouyban-Acree Model, *Chemical and Pharmaceutical Bulletin,* 54-4, 428-431.

Lazzaroni M., Bush D., Eckert C. (2005), Revision of MOSCED Parameters and Extension to Solid Solubility Calculations, *Industrial & Engineering Chemistry Research*, 44, 4075-4083.

Lindenberg C., Krattli M., Cornel J., Mazzotti M., Brozio J. (2009), Design and Optimization of a Combined Cooling/Antisolvent Process, *Crystal Growth & Design*, 9:2, 1124-1136.

Nowee S.M., Abbas A., Romagnoli J.A. (2008), Model-Based Optimal Strategies for Controlling Particle Size in Antisolvent Crystallization Operations, *Crystal Growth & Design,* 8:8, 2698-2706.

Ramkrishna D. (1985), The Status of Population Balances, Reviews in Chemical Engineering, 3:1, 49-95.

Randolph A.D., Larson M.A. (1988), *Theory of Particulate Processes (2nd ed.), Academic Press, San Diego,* CA.

Sheikhzadeh M., Trivkovic M., Rohani S. (2008), Real-time optimal control of an anti-solvent isothermal semi-batch crystallization process, *Chemical Engineering Science*, 63, 829-839.

Zhou G.X., Fujiwara M., Woo X.Y., Rusli E., Tung H.H., Starbuck C., Davidson O., Ge Z., Braatz R.D. (2006), Direct Design of Pharmaceutical Antisolvent Crystallization through Concentration Control, *Crystal Growth and Design*, 6-4, 892-898.

Numerical Studies of Wavelet-based Method as an Alternative Solution for Population Balance Problems in a Batch Crystalliser

Johan Utomo, Tonghua Zhang, Nicoleta Balliu and Moses O. Tadé*

Department of Chemical Engineering, Curtin University of Technology,
GPO Box U 1987, Perth, WA 6845, Australia
**corresponding author:(Tel:61-8-9266 4998; e-mail: m.o.tade@curtin.edu.au).*

Abstract: This paper deals with the numerical simulation studies of the nucleation and crystal growth process in a batch crystalliser. The population density functions may extend over orders of several magnitudes and the size distribution can be very sharp, thus accurate numerical solution of the population density functions can be challenging. The main interest for finding a more representative population balance solutions has motivated many researchers to develop many methods for the last four decades. In this paper, four methods have been implemented including wavelet-based method. The results obtained in three case studies, have demonstrated the wavelet scheme to be an alternative in providing accurate, fast and robust solutions.

Keywords: Crystallisation, Numerical scheme, Population Balance, Wavelet method

1. INTRODUCTION

Hulburt and Katz (1964) introduced a statistical mechanical formulation, known as population balance, for modelling the crystallisation process. The population balance equation (PBE) can be defined as a mathematical description characterising particles undergoing the mechanisms of birth, growth, death and leaving a certain particle phase space. In crystallisation, those mechanisms can be categorized as nucleation, growth, agglomeration and breakage. The population density functions may extend over orders of several magnitudes and the size distribution can be very sharp, thus the accurate numerical simulation of the population density functions can be challenging and has motivated several researchers in this area for decades to develop specialised algorithms for solving PBE, for example Ramkrishna (1981), Hounslow et al.(1988), Litster et al. (1995), Kumar and Ramkrishna (1996), Nicmanis and Hounslow (1996), and Mahoney and Ramkrishna (2002). All of these methods can be classified into four categories, such as method of weight residuals/orthogonal collocation, finite elements methods, finite difference schemes/discrete population balances methods, and other methods. There are major drawbacks from those methods such as computationally expensive, lack of stability and accuracy of the solution and in applicability of the solved models to be implemented in control based models. Extensive comparative discussion of those methods can be found in the literature (Kostoglou et al. 1994; Ramkrishna 2000; Vanni 2000). Therefore, the need for accurate, fast, robust and low order solution is essential for design, control and optimisation purposes.

This paper reports the application of wavelet based methods as an alternative, for solving population balance problems in a batch crystalliser condition. Previous study (Utomo et al. 2006) has been extended by comparing with finite difference based methods, such as upwind finite difference (U-FD), biased upwind finite difference (BU-FD), and method of weighted residuals, such as, orthogonal collocation with finite element method (OCFE). Different types of population balance are illustrated in the three cases discussed in this paper. They are having a high non-linearity, a steep-front profile, and stiff characteristic. The solutions are benchmarked with respects to their size (spatial grid points used), accuracy (mean and average error) and the computation time (t-CPU).

2. NUMERICAL SCHEMES

2.1 Previous Methods

Finite difference (FD) methods have been commonly used for the solution of all types of partial differential equations (ODEs) systems. FD method approximates the continuous function, $f(x)$, with Taylor expansion series (Hangos and Cameron, 2001). They can be a first order or second order approximations. In our case, FD method was used to approximate the first partial derivative of population density over its size and converts the PDE into a set of ODEs.

In this paper, the upwind finite difference and biased upwind finite difference schemes were applied to effectively handle the instability and to avoid the spurious oscillation as generated by a centred FD scheme. The five-point (fourth-order accuracy) upwind and biased upwind on uniform grids were implemented as described in (Wouwer et al. 2005). Orthogonal collocation technique was developed and applied in various cases of boundary value problems. The trial functions are chosen as sets of orthogonal polynomials and

the collocation points are the roots of these polynomials. The use of orthogonal polynomials is to reduce the error as the polynomial order increases (Gupta 1995; Hangos et al. 2001). OCFE is the combination method of dividing the regions into a number of elements and by applying orthogonal collocation techniques for each element can improve the solution where the profile is very steep. In the region, where there is a sharp transition, numbers of small elements can be applied while the remainder utilizes larger size elements. Selection of the elements size is therefore essential.

2.2 Wavelet-based methods

In 1992, Daubechies in her famous text-book, "Ten lectures on wavelets" (Daubechies 1992), predicted that a wavelet based software package to solve partial differential equations will be available in the market. The prediction has not been met as today there is no software available except for the Wavelet Toolbox in MATLAB®, which cannot be used for solving any partial differential equation (PDE).

The earliest wavelet application in chemical engineering was Wavelet Galerkin (WG). It was due to the work of Chen et al. (1996). Wavelet method with Galerkin scheme was utilised to solve the breakage population balance in a batch crystalliser. One of the challenges of WG is the expansion coefficients in WG was not specified in the physical space, while most of the PDEs can be directly solved in the physical space rather than converting and transforming its solution back to physical space. The second method was Wavelet Optimised Finite Difference (WOFD), developed by Jameson (1998). To date, this method has not been employed in the chemical engineering field. The third method was Wavelet Orthogonal Collocation (WOC). Its first application in chemical engineering was due to Liu and Cameron (2001). This method was successfully applied to solve the population balance and steep front concentration profiles in adsorption. However, WOC has not been applied for solving the complex cases which may involve the non-linearity and full dimensional variables. To sum up, the comparison of the three methods discussed are given below, which may initiate further development of a new wavelet-based numerical scheme for solving PDEs.

Table 1. Comparative components of three wavelet-based methods, WG: Wavelet Galerkin, WOFD: Wavelet Optimised Finite Difference and WOC: Wavelet Orthogonal Collocation

Comparative components	***WG***	***WOFD***	***WOC***
Basis calculation	Wavelet	**Physical**	**Physical**
BC treatment	Difficult	**Easy**	Moderate
Non-linearity handling	Difficult	**Moderate**	**Moderate**
Adaptive scheme	No	**Yes**	**Yes**
Computation capacity	Fixed	Fixed	**Reduced**

2.3 Daubechies orthonormal wavelets

Wavelet can be used as a basis function to represent a certain function. In the wavelet function, two-basis functions can be found, the scaling function and the wavelet function. The scaling function coefficient illustrates a local average of the solution can be calculated from the collocation points. The function (coarse illustration) and the wavelet function coefficient describes detailed information of the function (refinements) that cannot be found from the average coefficient. Compared to Fourier expansion, wavelet approximation gives smaller error and is highly localized at discontinuity regions (Nielsen 1998). Compared to the traditional trigonometric basis functions which have infinite support, wavelets have compact support, therefore wavelets are able to approximate a function by the placement of the right wavelets at appropriate locations. From Daubechies's work (1988), scaling function (ϕ) and wavelet function (ψ) can be described by a set of L (an even integer) coefficients (p_k : k = 0,1,…, L-1) through the two-scale relationship:

$$\phi(x) = \sum_{k=0}^{L-1} p_k \phi(2x-k) \tag{1}$$

and the wavelet function

$$\psi(x) = \sum_{k=2-L}^{1} (-1)^k \, p_{1-k} \phi(2x-k) \tag{2}$$

The support for the scaling function is in the interval 0 to (L-1), whilst for the wavelet function is in the interval (1-L/2) to (L/2). The coefficients p_k are called the wavelet filter coefficients.

Denote $L^2(R)$ as the space of square integrable functions on the real line. Let V_j be the subspace as the L^2-closure of the linear combination of:

$$\phi_{jk}(x) = 2^{j/2} \phi\left(2^j x - k\right) \tag{3}$$

for $k \in Z = \{...,-1,0,1...\}$. A function $f(x) \in V_j$ can be represented by the wavelet series:

$$f(x) = \sum_{k \in Z} f_{jk} \phi_{jk}(x) \tag{4}$$

The multi-resolution properties of wavelets give another advantage to represent functions in differential equations which can be solved numerically (Motard et al. 1994). Detailed information about Daubechies orthonormal wavelets can be found in Daubechies (1988).

2.4 Wavelet Orthogonal Collocation(WOC)

This method was introduced by Betoluzza and Naldi (1996) for solving partial differential equations. In 2001, it was developed and applied for solving population balance problems by Liu and Cameron (2001). The interpolation functions are generated by autocorrelation of the compactly supported Daubechies scaling functions $\phi(x)$. Then the function θ called autocorrelation function verifies the interpolation property due to the orthonormality.

$$\theta(0) = \int \phi(x)\phi(x)dx = 1 \tag{5}$$

and

$$\theta(n) = \int \phi(x)\,\phi(x-n)\,dx = 0,\ n \neq 0 \tag{6}$$

The approximate solution of our problem will be a function u_j in the term of its dyadic points to obtain the wavelet expression:

$$u_j(x) = \sum u_j\left(2^{-j}n\right)\theta\left(2^j x - n\right) \tag{7}$$

Detailed information can be found in Liu and Cameron (2001) and Bertoluza and Naldi (1996).

3. CASE STUDIES

Three case studies of population balances, which have sharp transition phenomena in their particle size distribution in the batch crystallizer, were tested in this paper. Even though the case studies considered here are simple, the analytical solutions are available for comparison purposes.

3.1 Case I: Population balance with nucleation and size-independent growth

The population balance for nucleation mechanism and size independent growth is described by the partial differential equation:

$$\frac{\partial n(x,t)}{\partial t} + G\frac{\partial n(x,t)}{\partial x} = B_0 \qquad G = 1;\ B_0 = \exp(-x) \tag{8}$$

where n is the number of particle (population density), x is the particle size, G is the growth rate, and B_0 is the nucleation rate.

With initial and boundary conditions such as:

$$n(t,0) = 0; \quad n(0,x) = 0 \tag{9}$$

The analytical solution for this case is:

$$n(t,x) = 1 - \exp(-x) \qquad x - t < 0$$
$$n(t,x) = \exp(-x)\left[\exp(-t) - 1\right] \qquad x - t > 0 \tag{10}$$

The dimensionless particle size L is defined as follow:

$$L = x/x,\max = [0:1]$$

Where: $x = [x,\min : x,\max] = [0:2]$

3.2 Case II: Population balance with size-independent growth only and initially seeded

One dimensional population balance for size dependent growth mechanism is described by the partial differential equation below:

$$\frac{\partial n(x,t)}{\partial t} + G\frac{\partial n(x,t)}{\partial x} = 0 \qquad G = 1 \tag{11}$$

With initial and boundary conditions such as:

$$n(t,0) = 0; \quad n(0,x) = \exp(-100\,(x-1)^2 \times 1/6.6) \tag{12}$$

The analytical solution for the second case is :

$$n(t,x) = \exp(-100(x - G.t - 1)^2 \times 1/6.6) \tag{13}$$

The dimensionless particle size L is defined as follow:

$$L = x/x,\max = [0:1]$$

Where: $x = [x,\min : x,\max] = [0:5]$

3.3 Case III: Seeded batch crystalliser with nucleation and growth

$$\frac{\partial n(L,t)}{\partial t} + G\frac{\partial n(L,t)}{\partial L} = B \tag{14a}$$

$$B = k_b\, M_T^j\, \Delta c^b \tag{14b}$$
$$G = k_g\, \Delta c^g \tag{14c}$$

where G is the growth rate, and B is the nucleation rate, k_g and k_b are the growth and nucleation constant, M_T is the suspension density, Δc is the supersaturation, and superscript g and b are the exponential constants for growth and nucleation rate respectively.

With initial and boundary conditions such as:

$$n(0,L) = n_0\,\delta(L - L_0)$$
$$n(t,0) = B/G \tag{15}$$

The supersaturation balance can be written as

$$\frac{d\Delta c}{dt} = \frac{3W_{S0}\,L_S^2\,G}{L_{S0}^3\,S} + \frac{3\rho\,k_v\,A_N\,G}{k_a} \tag{16}$$

where W_{S0} is the mass of seed crystals, $L_{S0}$is average size of seed crystals, L_S is the average size of growing crystals, S is the amount of solvent used, ρ is the crystal density, k_v is the volume shape factor, A_N is the total surface area of crystals, and k_a is the area shape factor.

The particle size L is defined as follow:

$$L = [L,\min : L,\max] = [400:1400]\,\mu m$$

4. RESULTS AND DISCUSSION

All the simulation results presented have been executed on an Intel(R) Core(TM) 2 CPU, with 2.00 GHz and 2.00 Gigabytes of RAM. A MATLAB® version 7.4.0.287 (R2007a) was used as the computation software to simulate the models. The built-in integrator of "ode15s" was utilised for integrating set of ordinary differential equations. The relative and absolute error of the integrator was specified at value of 10^{-3}.

4.1 Case I

The first case describes a simple population balance system which presents sharp front size distribution profiles. The PBE has a nucleation as function of size and a constant growth rate

and the analytical solution was available from (Chang et al. 1984). Four numerical schemes were applied in this case, their performance were tested in order to see the suitability, accuracy and stability of tested methods in handling the non-linearity and the sharp profile characteristic. The orthogonal collocation with finite elements (OCFE) scheme with 2 elements, which comprise of 31 grid points was applied. The other two methods were based on finite difference scheme. The upwind finite difference (U-FD) of 2 points and the biased upwind finite difference (BU-FD) of 5 points were employed. Both of them were discretised in space to generate 129 grid points. Lastly, wavelet orthogonal collocation (WOC) was used, and their performance were benchmarked with the analytical solution.

The average error (AE) was defined as the square root of sum of square error divided by the number of grid points. As it was calculated on an individual time basis, AE does not depend on the number of equations (grid points). The value of AE can illustrate the total absolute error of the grids for a certain time. A small value of AE may also illustrate a stable solution. While AE is the global error figure, the maximum error (ME) could show a local error or an overshoot from the reference value. ME was defined as the maximum value of the square root of sum of square error at a certain time.

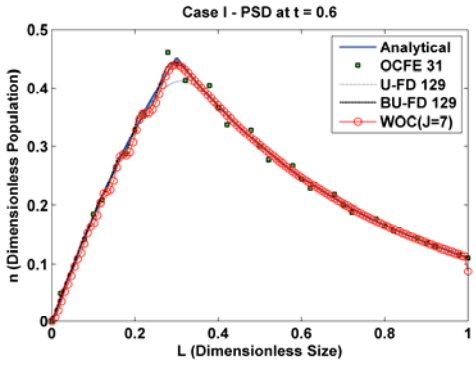


Fig. 1. Particle size distribution, case I at 0.6 seconds, simulated by four methods and compared with analytical solution.

Figure 1 shows a comparative particle size distribution (PSD) at 0.6 s. The particle distribution was initially zero and the nucleation start to generate nuclei and at the same, the born nuclei grow at a constant rate of 1.0. OCFE 31 profile produced a slightly overestimated particle density than the analytical solution. While, all the other numerical schemes, including WOC, with the same resolution gave comparable results. The simplest two points upwind scheme gave an under predicted population at the peak point, while the simulation results from biased upwind and wavelet are equivalent in term of minimising the error at the peak point. From the AE point of view, U-FD was the least while BU-FD was the first accurate then followed by WOC. When the wavelet level (J) was increased from 7 to 9, the ME values were about the same but an increased accuracy was achieved by 62%, consequently at the same time the computation time (t-CPU) was 440% higher.

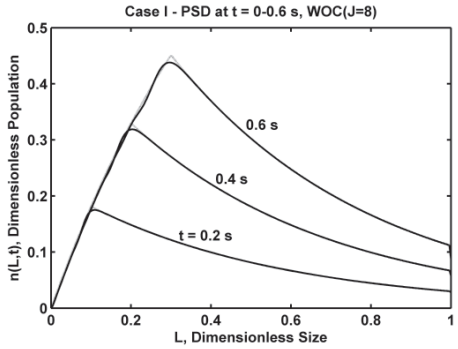


Fig. 2. Particle size distribution of WOC (J=8) at 0.2, 0.4 and 0.6 second, the black line: WOC solution and grey line: analytical solution.

In this simulation, all the computation time was short (less than 5 seconds) because all the methods have the same structure of a matrix form. Matrix to matrix calculation was superior than loop calculation (using *for* loop) for its computation time and its adaptability to a more complex case. Figure 2 shows the solutions produced by WOC (J=8) at various time from t = 0.2, 0.4 and 0.6 seconds. At this point, it can be concluded that OCFE methods can be with reasonable accuracy level and WOC method can be employed as an equivalent alternative solution for handling the case of sharp fronts profile caused by non-linear nucleation function. A question that arose is whether these methods are able to track a very sharp profile as shown in the next case.

Table 2. Numerical performance results for case I, N: grid points, AE: average error, ME: maximum error, t-CPU: computation time

Case I @ 0.6 s	N	AE	ME	t-CPU
OCFE (2)	31	0.4383	0.7861	< 1s
U-FD (2)	129	0.0082	0.0391	< 1s
BU-FD (5)	129	0.0031	0.0120	< 1s
WOC (J=7)	129	0.0064	0.0243	< 1s
WOC (J=8)	257	0.0033	0.0242	1.4 s
WOC (J=9)	513	0.0024	0.0239	4.4 s

4.2 Case II

In a seeded batch crystalliser, where the nucleation can be minimised, particle size distribution will be controlled only by the initial condition of seeding and a crystal growth. In this case, the crystal growth was assumed again constant, and the seed condition was artificially made to present a very sharp front of particle size distribution. The previous study done by Utomo, et al. (2006) reported that OCFE method cannot be applied as the unstable solutions were obtained. Moreover, the upwind finite difference scheme gave delayed solutions. Therefore, only WOC method as an equivalent method was tested in this case. The effect of wavelet resolution (J) and vanishing moments (M) were observed to closely study its performance.

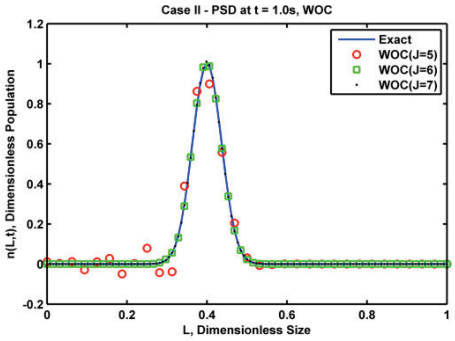


Fig. 3. Particle size distribution of WOC at t = 1s, using various J =5,6,7.

Figure 3 shows the WOC solutions with various wavelet resolution (J=5, 6, and 7) as compared to the exact analytical solution. It is clear that except for J=5, WOC presented a good validation result with high accuracy. WOC for J=5, however, not only gave a high ME but also oscillation and a negative value problem. The final solution of J=5 was stable, but early oscillation recorded the highest average error in the solution. The same problem was mentioned by (Muhr et al. 1996), can be easily rectified by decreasing the spacing or utilising adaptive gridding.

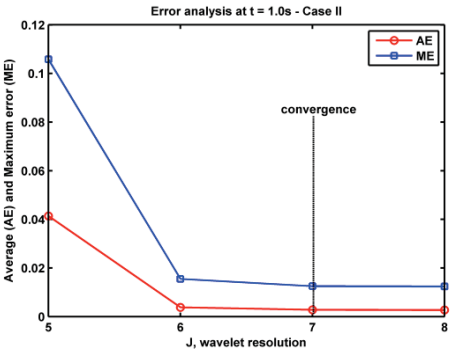


Fig. 4. The error analysis at t = 1s for various J, where AE: average error and ME: maximum error.

The convergence issue is important for numerical simulation. The WOC method for the tested resolution could easily give good convergence at a resolution as low as 6. Both AE and ME could be used for error analysis to show the convergence at a certain time. As shown in Figure 4, the convergence was actually reached as the J was increased from 6 to 7 and further increased in resolution would not give any improvement in the accuracy. The effect of vanishing moments (M) could only be observed, when the lowest resolution was applied. M = 0 and 1 were the most optimal values in giving the lowest level of error. At this stage, it can be summarised that WOC could have been also employed as an alternative for a very sharp front's profile. The selection of resolution is more sensitive to the computational performance rather than the sequential choice of vanishing moments. Selection of M becomes sensitive only when the symptom of instability was observed. To demonstrate WOC capabilities as an alternative method, the more complex population balance in a batch crystalliser case study was performed in the next section.

Table 3. Numerical performance results for case II; M: vanishing moments, N: grid points, AE: average error, ME: maximum error, t-CPU: computation time

Case II @ 1.0 s	M	N	AE	ME	t-CPU
WOC (J=5)	0	33	0.0369	0.0933	< 1.0 s
	1		0.0368	0.0933	
	2		0.0406	0.0931	
	3		0.0414	0.1058	
WOC (J=6)	0-3	65	0.0038	0.0155	< 1.0 s
WOC (J=7)	0-3	129	0.0028	0.0125	< 1.0 s
WOC (J=8)	0-3	257	0.0027	0.0124	1.3 s

4.3 Case III

The last case presents a seeded batch crystalliser with capacity of 25.5 kg solvent, running in an isothermal condition for batch time of 6000 seconds. The nucleation and growth kinetics are described in (14b) and (14c). The initial condition was the seed condition at the average size of 500 μm and the boundary condition is outlined in (15). To solve the system, a population balance equation (14a) coupled with the mass (supersaturation) balance for the solute and solid phase as in (16), thus the dynamic of crystal size distribution (CSD) can be computed. All the parameters used in this case were adopted from (Tavare et al. 1986).

For illustrative purposes, Figure 5, shows the profile of supersaturation and crystal growth rate during 6000 seconds batch operation. The initial condition of 0.015 kg/kg solvent would give a corresponding crystal growth of 1.68×10^{-10} m/s. The WOC method with J = 7 and M = 1 was employed and a reasonable result was obtained in Figure 6 as the experimental results was not available.

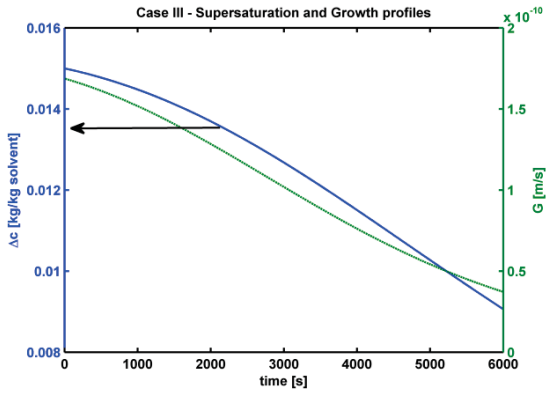


Fig. 5. Supersaturation (left) and crystal growth rate (right) profiles for case III up to 6000 seconds.

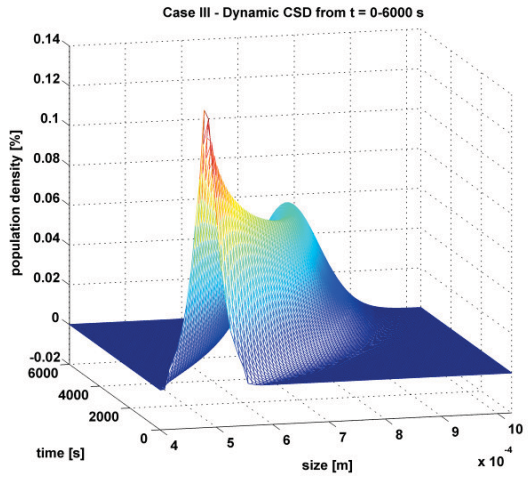


Fig. 6. Dynamic crystal size distribution for case III, for 400-1000 μm in size and 0-6000 seconds.

5. CONCLUSIONS

In this paper, three case studies which present a very sharp size distribution profile have demonstrated the potential of wavelet-based numerical scheme as an alternative in providing accurate, fast and robust solutions. Further research on a new wavelet numerical scheme and wavelet application in chemical engineering field is essentially required and promising. From the computational efficiency result shown, with the WOC algorithm, the model is suitable to be employed in online control system, however, from control engineers' perspective, low-order models are needed.

REFERENCES

Bertoluzza, S. and G. Naldi (1996). A Wavelet Collocation Method for the Numerical Solution of Partial Differential Equations. *Applied and Computational Harmonic Analysis*, 3(1), 1-9.

Chang, R.-Y. and M.-L. Wang (1984). Shifted Legendre function approximation of differential equations; application to crystallization processes. *Computers & Chemical Engineering*, 8(2), 117-125.

Chen, M.-Q., C. Hwang and Y.-P. Shih (1996). A wavelet-Galerkin method for solving population balance equations. *Computers & Chemical Engineering*, 20(2), 131-145.

Daubechies, I. (1988). Orthonormal bases compactly supported wavelets. *Communications Pure and Applied Mathematics*(41), 909-996.

Daubechies, I. (1992). Ten Lectures on Wavelets. *Philadelphia: Society for Industrial and Applied Mathematics*.

Gupta, S. K. (1995). *Numerical Methods for Engineers*. New Age International, New Delhi.

Hangos, K. and I. T. Cameron (2001). *Process Modelling and Model Analysis*. Academic Press, London.

Hounslow, M. J., R. L. Ryall and V. R. Marshall (1988). A discretized population balance for nucleation, growth and aggregation. *AIChE Journal*, 34, 1821-1832.

Hulburt, H. M. and S. Katz (1964). Some problems in particle technology : A statistical mechanical formulation. *Chemical Engineering Science*, 19(8), 555-574.

Jameson, L. (1998). A Wavelet-Optimized, Very High Order Adaptive Grid And Order Numerical Method. *Siam J. Sci. Comput.*, 19(6), 1980-2013.

Kostoglou, M. and A. J. Karabelas (1994). Evaluation of Zero Order Methods for Simulating Particle Coagulation. *Journal of Colloid and Interface Science*, 163(2), 420-431.

Kumar, S. and D. Ramkrishna (1996). On the solution of population balance equations by discretization--I. A fixed pivot technique. *Chemical Engineering Science*, 51(8), 1311-1332.

Litster, J. D., D. J. Smit and M. J. Hounslow (1995). Adjustable Discretized Population Balance for Growth and Aggregation. *AIChE Journal*, 41(3), 591-603.

Liu, Y. and I. T. Cameron (2001). A new wavelet-based method for the solution of the population balance equation. *Chemical Engineering Science*, 56(18), 5283-5294.

Mahoney, A. W. and D. Ramkrishna (2002). Efficient solution of population balance equations with discontinuities by finite elements. *Chemical Engineering Science*, 57(7), 1107-1119.

Motard, R. L. and B. Joseph (1994). *Wavelet applications in chemical engineering*. Kluwer Academic Publisher, London.

Muhr, H., R. David, J. Villermaux and P. H. Jezequel (1996). Crystallization and precipitation engineering-VI. Solving population balance in the case of the precipitation of silver bromide crystals with high primary nucleation rates by using the first order upwind differentiation. *Chemical Engineering Science*, 51(2), 309-319.

Nicmanis, M. and M. J. Hounslow (1996). A finite element analysis of the steady state population balance equation for particulate systems: Aggregation and growth. *Computers & Chemical Engineering*, 20(Supplement 1), S261-S266.

Nielsen, O. M. (1998). Wavelets in Scientific Computing. Ph.D. Thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby.

Ramkrishna, D. (1981). Analysis of population balance--IV : The precise connection between Monte Carlo simulation and population balances. *Chemical Engineering Science*, 36(7), 1203-1209.

Ramkrishna, D. (2000). *Population balance : theory and applications to particulate systems in engineering*. Academic Press, London.

Tavare, N. S. and J. Garside (1986). Simultaneous Estimation of Crystal Nucleation and Growth Kinetics From Batch Experiments. *Chemical Engineering Research and Design*, 64(a), 109-118.

Utomo, J., N. E. Balliu and M. O. Tade (2006). Challenges of Modelling A Population Balance Using Wavelet. F. J. Doyle III, J. O. Trieweiler and A. R. Secchi (ed.),*Preprints Volume II*,643-648, IFAC, Gramado, Brazil.

Vanni, M. (2000). Approximate Population Balance Equations for Aggregation-Breakage Processes. *Journal of Colloid and Interface Science*, 221(2), 143-160.

Wouwer, A. V., P. Saucez, W. E. Schiesser and S. Thompson (2005). A MATLAB implementation of upwind finite differences and adaptive grids in the method of lines. *Journal of Computational and Applied Mathematics*, 183(2), 245-258.

Model-predictive Control Algorithms

Oral Session

Economic Dynamic Real-Time Optimization and Nonlinear Model-Predictive Control on Infinite Horizons

Lynn Würth * **James B. Rawlings** ** **Wolfgang Marquardt** ***

* *AVT - Process Systems Engineering, RWTH Aachen University, D-52056 Aachen, Germany (email: lynn.wuerth@avt.rwth-aachen.de).*
** *Dept. Chem. & Biol. Eng. - Univ. of Wisconsin. Madison (WI), USA (email: rawlings@engr.wisc.edu)*
*** *AVT - Process Systems Engineering, RWTH Aachen University, D-52056 Aachen, Germany (email: wolfgang.marquardt@avt.rwth-aachen.de).*

Abstract: This paper investigates the formulation of nonlinear model-predictive control problems with economic objectives on an infinite horizon.
The proposed formulation guarantees nominal stability for closed-loop operation. Furthermore, a novel solution method of the infinite horizon method through a transformation of the independent time variable is employed. The closed-loop optimization with infinite horizon is compared to a finite-horizon formulation. A small case study is presented for illustration purposes.

Keywords: Dynamic Real-Time Optimization, Economic Objective, Nonlinear Model-Predictive Control, Infinite Horizon, Stability.

1. INTRODUCTION

The interest in economic Dynamic Real-Time Optimization (DRTO) or Nonlinear-Model Predictive Control (NMPC) with economic objectives has increased (Backx et al. (2000), Helbig et al. (2000), Engell (2007), Rawlings and Amrit (2008), Zavala (2008)), as the development of efficient methods for solving these types of optimization problems has significantly progressed in recent years. Compared to the traditional formulation of NMPC problems with quadratic cost criteria minimizing the deviation from a fixed steady-state set-point, the economic dynamic optimization problem exploits all the dynamic degrees of freedom available to maximize the profit of the plant on a given time horizon. Furthermore, the profit is maximized at a sampling rate of high frequency, whereas the traditional steady-state optimization is performed at a slow rate and only when the process is in a steady-state. Thus, disturbances with a favorable impact on the profit can also be exploited efficiently instead of compensating them by minimizing a steady-state offset.
However, theoretical studies of DRTO and NMPC problems with economic objectives are still lacking. Huesman et al. (2008) have pointed out that certain formulations with linear economic objective functions lead to multiple solutions. They concluded that some degrees of freedom are left for optimization. These degrees of freedom can be exploited by introducing a second optimization problem to improve the operability.
Rawlings and Amrit (2008) have shown that in order to optimize process economics it is sometimes advantageous not to reach the steady-state quickly. Secondly, they pointed out that the formulation with an economic objective results in characteristic trajectories behaving like a turnpike. The trajectory is attracted to a constant path and finally moves away from the constant path at the end of the horizon. The turnpike is a characteristic property of the economic optimization problem with finite-horizon, which has been introduced and studied in the economics literature before (Carlson et al., 1991).
In the economics literature an infinite horizon formulation is often employed. Infinite horizons have also been considered in the MPC literature as stabilizing although impractical for online application. A review of methods providing stability but circumventing infinite horizons by adding a terminal constraint, a terminal cost function, or by employing a terminal constraint set with a local stabilizing controller have been reviewed by Mayne et al. (2000).

An infinite-horizon formulation for (nonlinear) economic dynamic optimization on a receding horizon is explored in this paper. There are several advantages related to an infinite horizon formulation:

(1) The somewhat arbitrary choice of the final time of the optimization horizon is avoided. A natural formulation of the optimization problem is achieved for continuous processes if the final time is not specified.
(2) The infinite-time horizon formulation leads to closed-loop operation with guaranteed stability. This property was also exploited in the literature of linear MPC, where the infinite-horizon formulation can

be transformed into a finite horizon formulation by adding a terminal term to the cost functional (Muske and Rawlings, 1993).

In the following, the formulation of the economic optimization problem will be investigated and a comparison between finite and infinite horizon problems will be carried out. A new approach is presented for solving the infinite-horizon problem numerically. We use a transformation of the infinite-time onto a finite-time horizon and combine this transformation with an adaptive discretization. As already noted in the literature investigating stability in MPC, very long horizons provide stable control but lead to high computational costs. Many solution methods employed in MPC use a discretization of the control variables with a uniform spacing of grid points. An alternative adaptive discretization (Schlegel et al., 2005) introduces grid points mainly within the transient parts of the profile and reduces the computational load. Note that this method can also be applied to solve the infinite-horizon problem in the MPC regulator case to track a given set-point.

2. ECONOMIC DRTO PROBLEM

2.1 Finite-Horizon Formulation

The moving horizon formulation of the DRTO problem is similar to the formulation used in nonlinear model-predictive control, although an economic objective is chosen to provide an economically optimal operation at all times (Helbig et al., 2000). The moving horizon problem is defined as follows:

$$\min_{\boldsymbol{u}^j(t)} \Phi(\boldsymbol{x}, \boldsymbol{u}, t_0, t_f) \tag{1}$$

$$\text{s.t.} \quad \dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}^j(t)), \tag{2}$$

$$\boldsymbol{y}(t) = \boldsymbol{g}(\boldsymbol{x}(t), \boldsymbol{u}^j(t)), \tag{3}$$

$$\boldsymbol{x}(t^j) = \hat{\boldsymbol{x}}^j, \tag{4}$$

$$\boldsymbol{0} \geq \boldsymbol{h}(\boldsymbol{x}(t), \boldsymbol{y}(t), \boldsymbol{u}^j(t)), \tag{5}$$

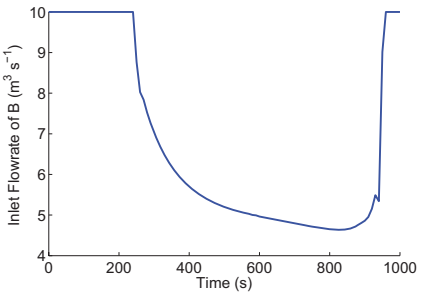
$$\boldsymbol{0} \geq \boldsymbol{e}(\boldsymbol{x}(t_f^j)), \tag{6}$$

$$t \in [t^j, t_f^j], \tag{7}$$

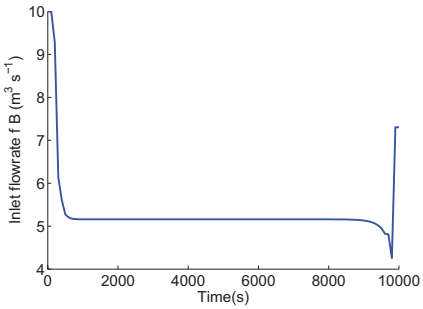
$$t^j := t^{j-1} + \Delta t, \tag{8}$$

$$j = 0, 1, ...J. \tag{9}$$

$\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$ are state variables with initial conditions $\hat{\boldsymbol{x}}^j$; $\boldsymbol{y}(t) \in \mathbb{R}^{n_y}$ are algebraic output variables. The dynamic process model (2) is given by $\boldsymbol{f}(\cdot)$. The time-dependent control variables $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$ are degrees of freedom for the optimization problem. The optimization problem is solved on the time horizon $[t^j, t_f^j]$ at each sampling instant t^j; the control is implemented on the current sampling interval (assuming negligible computational time), and the optimization horizon is then shifted by the sampling interval Δt. Equations (5) and (6) describe the path constraints $\boldsymbol{h}(\cdot)$ on the input and state variables and the endpoint constraints $\boldsymbol{e}(\cdot)$ on the state variables. Process operation is determined by economic decision criteria, which enter into the definition of the objective function $\Phi(\cdot)$. Exemplarily the profit function can be defined as



(a) Final time of 1000s



(b) Final time of 10000s

Fig. 1. Finite horizon

$$\Phi = -\int_0^T (c_{prod}\dot{n}_{prod} - c_{reac}\dot{n}_{reac} - \dot{q})dt, \tag{10}$$

where c_{prod} and c_{reac} are the costs of the products and reactants, and $\dot{n}_{prod}$ and $\dot{n}_{reac}$ are the flowrates of the products and reactants. The term $\dot{q}$ includes utility costs, depreciation, and other expenses.

If we choose a finite horizon in a dynamic optimization problem with economic objective, we often observe a so-called turnpike effect when looking at the solution profile (Rawlings and Amrit (2008), Carlson et al. (1991)). The turnpike effect means that the trajectory spends most time at a balanced equilibrium path, which is independent of the initial condition and the final time. Under certain conditions, this turnpike reduces to a singleton. In that case, the characteristic behavior is that the trajectory is attracted by a stationary path, at a certain time t_m and stays on this constant path, until it reaches a point t_n close to the end of the horizon and moves away from the path at the end of the horizon. This gives rise to trajectories as shown in Figure 1. The trajectories represent the inlet flowrate of the reactant of the CSTR presented in Section 5; they were computed for different final times. In these figures the solution path strongly depends on the choice of the final time. In the first figure, the final time is not long enough to let the process reach the turnpike. However, if a long horizon is chosen, the process gets on a constant path $\bar{k}$ (Carlson et al., 1991). The optimal steady-state path represents an attractor for the finite horizon path.

In practice, often a long horizon length is chosen to achieve closed-loop stability (Mayne et al., 2000). For stable plants, the final time T is often chosen to be large compared to the settling time of the plant. The somewhat arbitrary choice of the final time suggests that more

research is required to reconsider the current formulations and to look for adequate formulations of economic dynamic optimization problems providing stability in closed loop.

2.2 Infinite Horizon Formulation

For a given optimal trajectory, according to Bellmann's optimality principle, the trajectory starting from any point on the optimal trajectory is optimal for the corresponding problem initiated at that point on the trajectory. This implies that the formulation of the optimization problem on an infinite horizon is providing stability in closed loop.

If the profit function is optimized on a long or an infinite horizon, the time value of money should be accounted for. This is accomplished through the parameter ρ discounting the future profit to the present value. If we deal with discrete payment periods, discounting the future amounts to today's value can be included in the calculation of the net present value. The objective function maximizing the net present value of cash flows C_k in N discrete time periods and with discounting rate ρ is formulated as

$$\Phi = \sum_{k=0}^{N} \frac{C_k}{1+\rho^k}, \tag{11}$$

where k is the index of the time period. In this work we investigate an objective function with continuous discounting on an infinite horizon:

$$\Phi = -\int_0^{\infty} e^{-\rho t}(c_{prod}\dot{n}_{prod} - c_{reac}\dot{n}_{reac} - \dot{q})dt. \tag{12}$$

The exponential formulation is usually employed for optimal control problems in continuous time and is equivalent to the discrete-time formulation in equation (11). Both formulations use, however, the same discount factor ρ, which can be chosen as the annual market interest rate.

Note that researchers in macroeconomics have included infinite horizons in their problem formulations very early to model e.g. economic growth (Barro and Sala-i-Martin, 1995). This is due to the fact that there is no natural finite time in these types of problems and the consequences of investment are very long-lived. Since this theory is not well-known in the systems and control community, an example of an economic growth model by Cass (1966) is shown here for illustration of the economic growth problem formulation:

$$y(t) = f(k(t)) \tag{13}$$
$$c(t) = y(t) - z(t) \tag{14}$$
$$\dot{k}(t) = z(t) - \mu k(t) \tag{15}$$
$$k(0) = k_0 \tag{16}$$

$k(t)$ is the stock of capital accumulated at time t. The production function $f(\cdot)$ associates an output y with the capital stock k. The output $y(t)$ can be either consumed at a rate $c(t)$ or invested at a rate $z(t)$. The capital stock depreciates at a constant rate μ. Cass (1966) considered the welfare functional

$$W = \int_0^T e^{-\rho t} U(c(t))dt, \tag{17}$$

where $U(c)$ is the concave utility function depending on the level of consumption $c(t)$. The maximization of W is the standard optimal control problem and the final time T is often assumed to be infinite in economic growth theory. If the discount factor ρ is strictly positive, the objective function on an infinite horizon is bounded. However, if the discount factor is zero, the objective function becomes unbounded. Methods for reducing the unbounded objective on an infinite horizon to finite rewards have been presented by Carlson et al. (1991). We consider in this paper the case where the discount factor is positive, as it seems to be a reasonable assumption to include the time value of money on a long or infinite-time horizon. The discount factor is chosen in the range of the interest rate of the market.
Very few numerical solution methods exist in the literature of mathematical economics to solve the infinite-time horizon problem. Often, the indirect methods of optimal control are used to derive the First-Order Necessary Conditions of Optimality and a two-point boundary-value problem has to be solved. In the following, we will apply a numerical solution method with adaptive grid refinement to solve nonlinear infinite-horizon problems.

3. INFINITE HORIZON SOLUTION APPROACH

Solving the infinite-horizon problem is not straightforward and several attempts approximating the infinite horizon with finite horizons or reformulating the infinite-time problem exist in the literature.

3.1 Time Transformation

A common approach replaces the infinite-time horizon by a finite-time horizon, thereby introducing a truncation error. A second possibility is to use a variable transformation to transform the infinite-horizon into a finite-horizon problem. We choose the second approach in this work, since it does not introduce truncation errors and allows to obtain a solution of higher accuracy of the optimization solution on the infinite horizon.
The infinite horizon with $t \in [0, \infty)$ can be transformed into a finite horizon with $\tau \in [0, 1]$ using the variable transformation

$$\tau = t/(1+t) \tag{18}$$

of the independent variable t. A similar variable transformation was presented by Kunkel and Hagen (2000) to obtain solutions of the infinite-horizon optimal control problem. The variable transformation of eq. (2) yields the transformed system:

$$\frac{d\boldsymbol{x}}{d\tau} = \frac{\boldsymbol{f}(\boldsymbol{x}(\tau), \boldsymbol{u}(\tau))}{(1-\tau)^2}. \tag{19}$$

We can see that a singularity is introduced at $\tau = 1$, which corresponds to $t = \infty$.
In order to study the singularity at $\tau = 1$, we will restrict the analysis to the scalar system

$$\frac{dx}{d\tau} = \frac{f(x(\tau))}{(1-\tau)^2}. \tag{20}$$

The eigenvalue problem corresponding to eq. (19) is

$$\frac{dx}{d\tau} = \frac{\lambda x}{(1-\tau)^2}. \tag{21}$$

The general solution of this differential equation is

$$x(\tau) = ce^{\frac{\lambda}{1-\tau}}. \tag{22}$$

If we assume that the system is within the region of attraction of a stable steady-state, we obtain the limit

$$\lim_{\tau \to 1} c e^{\frac{\lambda}{1-\tau}} = 0, \tag{23}$$

if $\lambda < 0$ and the system will reach steady-state as $\tau \to 1$.

However, if $\lambda > 0$, the system is unstable, and

$$\lim_{\tau \to 1} c e^{\frac{\lambda}{1-\tau}} = \infty. \tag{24}$$

In that case we have to handle the singularity at $\tau = 1$ by imposing a boundary condition at final time. The steady-state solution can be imposed as terminal boundary condition. If we are dealing with an unstable system, a numerical solution approach different from the single-shooting method presented in the following section must be employed. The boundary value problem problem could be solved using multiple-shooting or orthogonal collocation methods. In order to generalize the analysis of the singularity above, an extension to the multivariable case should be performed including an analysis of the stable and unstable modes of the system.

3.2 Adaptive Discretization

The continuous control variables in optimization problem (1) are discretized after the time transformation with eq. (18) using piecewise-constant or piece-wise linear approximations. The discretization and the formulation of the NLP for the piecewise constant approximation reads as

$$\mathbf{u}_i(\tau_k) = \mathbf{c}_{\mathbf{u}_{i,k}}, k = 1, ..., N, i = 1, ..., n_u, \tag{25}$$

where N is the number of discretization intervals and n_u the number of control variables. Choosing the discretized controls $\mathbf{z}_i := [\mathbf{c}_{\mathbf{u}_{i,k}}]$, as the n_z optimization variables, the dynamic optimization problem can be transcribed into the NLP

$$\min_{\boldsymbol{z}} f(\boldsymbol{z}) := \Phi(\boldsymbol{z}) \tag{26}$$

$$\text{s.t. } \boldsymbol{g}(\boldsymbol{z}) \geq 0. \tag{27}$$

The nonlinear program can be solved by employing a standard SQP algorithm. Since the optimization algorithm requires repetitive function evaluations and gradients, the objective function Φ, constraints $\mathbf{g}(\cdot)$ and their gradients are evaluated by a simultaneous integration of the process model and the sensitivity equation system.
The optimization problem is solved using the dynamic optimization software DyOS (2002), which adopts an adaptive control vector parameterization (Schlegel et al., 2005). The adaptation is done using a refinement method based on a wavelet analysis of the control profile. If the horizon of the optimization problem is infinite or has a long finite time, an adaptive discretization is essential to deal with the computational load associated with these long or infinite horizons. The adaptive discretization allows to introduce the grid points selectively in transient regions of the control profile, and therefore avoids overloading the optimization problem with many optimization parameters. Furthermore, a highly accurate solution profile can be obtained.

3.3 Closed-loop Implementation

The solution approach outlined in the previous section is implemented in an algorithm for closed-loop dynamic optimization. The optimization is performed on the transformed timehorizon $[0, 1]$, and the results are converted back to the original timehorizon.

for $j = 1, N$ (number of sample intervals) **do**

(1) **Solve** the optimization problem (26), where the control variables $\boldsymbol{u}^j(\tau)$ have been discretized on $[0, 1]$, to obtain the solution $\boldsymbol{z}^j$.
(2) **Transform** the control variables $\boldsymbol{u}^j(\tau)$, $\tau \in [0, 1]$ back to the original representation $\boldsymbol{u}^j(t)$, $t \in [0, \infty)$.
(3) **Implement** the control variables $\boldsymbol{u}^j(t)$ for one sampling interval $[t_j, t_{j+1}]$. Get measurements and compute state estimates.
(4) **Horizon shift:** Reduce the time horizon by one sampling interval and use the shifted solution as initial guess for the next optimization problem.
(5) **Transform** the shifted variables $\boldsymbol{u}^j(t)$ to the finite horizon representation $\boldsymbol{u}^j(\tau)$, $\tau \in [0, 1]$ using eq. (18).

end for

4. INFINITE-HORIZON NMPC

Apart from the economic dynamic optimization problem, the presented solution method for infinite-horizon formulations can also be used for the NMPC regulator problem, if we deal with a pure tracking problem. In that case the economic objective function is replaced by the quadratic objective function

$$\Phi = \int_0^\infty (\Delta \boldsymbol{u}^T \boldsymbol{Q} \Delta \boldsymbol{u} + (\boldsymbol{y}^{set} - \boldsymbol{y})^T \boldsymbol{W} (\boldsymbol{y}^{set} - \boldsymbol{y})) dt, \tag{28}$$

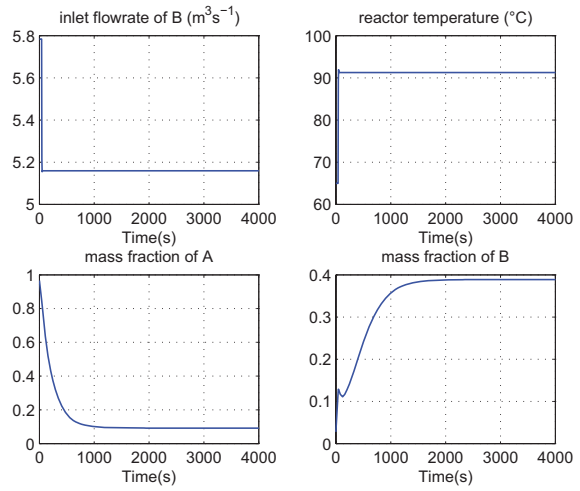
minimizing the deviation from fixed set-points $\boldsymbol{y}^{set}$ and the control moves $\Delta \boldsymbol{u}$. As a finite horizon is traditionally used in NMPC, some methods guarantee closed-loop stability by introducing e.g. a terminal constraint. However, introducing a terminal constraint can lead to feasibility problems. By employing the infinite-horizon solution method, the disadvantages of the finite-horizon methods to guarantee closed-loop stability can be avoided.

5. CASE STUDY

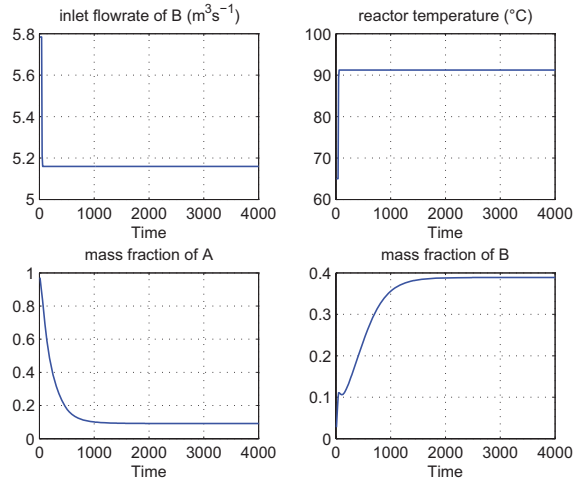
The approach is applied to the benchmark case of the Williams-Otto continuous stirred tank reactor, as introduced by Forbes (1994). The reactions taking place in the reactor are

$$A + B \xrightarrow{k_1} C,\ C + B \xrightarrow{k_2} P + E,\ P + C \xrightarrow{k_3} G.$$

Reactant A is already present in the reactor, whereas reactant B is fed continuously to the reactor. During the exothermic reactions the desired products P and E as well as the side-product G are formed. At the end of the batch, the conversion to the main products P and E should be maximized. During the batch, constraints on the inlet flow rate of reactant B ($F_{B_{in}}$) and the reactor temperature (T_r) must be fulfilled. The manipulated control variables of this process are $F_{B_{in}}$ and T_r. The CSTR is assumed to be well-mixed and the dynamics of the cooling system are neglected. The reactor system is open-loop stable and therefore the singularity issues at $\tau = 1$ do not arise in this case study.

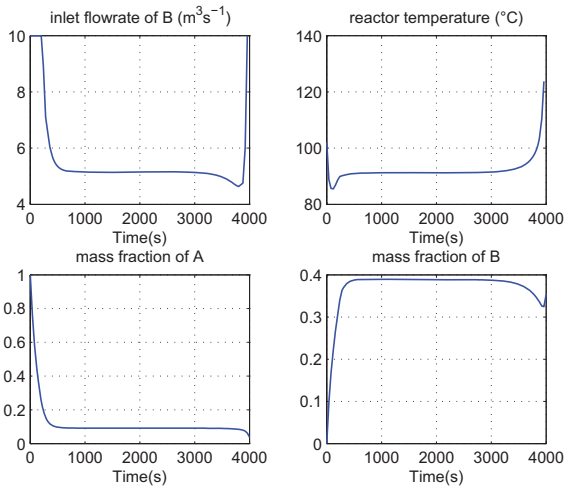


(a) Infinite horizon, open-loop

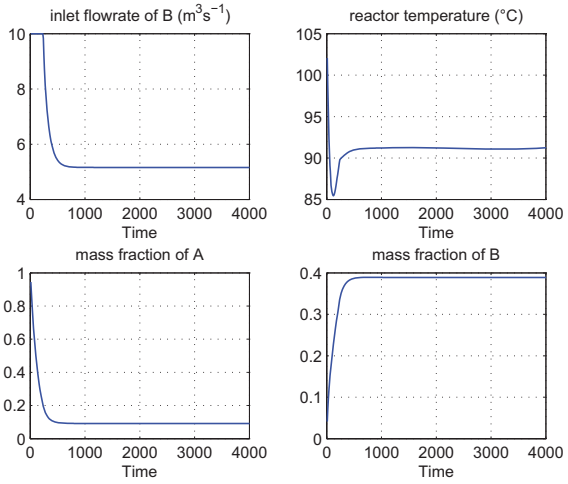


(b) Infinite horizon, closed-loop

Fig. 2. Infinite horizon results



(a) Finite horizon, open-loop



(b) Finite horizon, closed-loop

Fig. 3. Finite horizon results

5.1 Economic DRTO Problem

The economic objective is to maximize the profit, which consists of the revenue obtained from the products minus the costs of the reactants over the infinite time horizon. Furthermore, an annual discount rate of 5% is chosen to account for the time value of money. The optimization problem is formulated as follows:

$$\max_{F_{B_{in}}(t),T_r(t)} \Phi = \int_0^\infty e^{-\rho t}(c_p\dot{n}_p + c_e\dot{n}_e - c_a\dot{n}_a - c_b\dot{n}_b)dt \tag{29}$$

s.t. *process model*, and

$$0\frac{\text{kg}}{\text{sec}} \leq F_{B_{in}}(t) \leq 5.784\frac{\text{kg}}{\text{sec}}, \tag{30}$$

$$0\ ^\circ\text{C} \leq T_r(t) \leq 150\ ^\circ\text{C}. \tag{31}$$

5.2 Closed-Loop Results

The economic performance of the operation of the CSTR is optimized for both infinite and finite horizon formulations on a receding horizon. The results for the infinite-horizon case are shown in Figure 2. The open-loop optimization problem is solved using the time transformation for $\tau \in [0, 1]$ and the results are simulated on a long horizon in the original time variable. The closed-loop results are obtained using the algorithm with adaptive grid refinement as sketched in Section 3.3. Figure 2 shows that the nominal trajectories in open-loop and closed-loop are almost identical except minor deviations due to the online adaptation scheme. This result confirms that the solution of the optimal problem on an infinite horizon in closed-loop is providing a nominally stable control with an economic objective. Furthermore, the computational effort has been low as only few degrees of freedom are required in the adaptive grid refinement approach. It is interesting to observe that the control variables reach their steady-state values very fast, but the state variables require more time to reach the steady-state values.

The results for the finite-horizon formulation with a final time of 4000 s are shown in Figure 3. In this case, the open-loop trajectory shows the turnpike behavior. Since

only the first control interval is implemented in closed-loop, the end of the trajectory in open-loop is actually never implemented. This leads to a discrepancy between the open-loop and the closed-loop behavior, as Bellmann's optimality principle is not fulfilled. Nevertheless, the figures show that the system reaches the steady-state quite fast.

Comparing the closed-loop results obtained with the infinite- and the finite-time horizon formulation, the figures show that the transient part of the trajectories at the beginning is quite different, but that the steady-states obtained after a certain time are identical. The temperature profile is different in the infinite-horizon (Figure 2.b) and the finite-horizon cases (3.b), as the temperature increases from 65 °C to 90 °C in (2.b) and the temperature decreases from 102 °C to 91 °C in (3.b). It was observed that this difference only occurs for relatively short horizons. If the final time chosen for the finite-horizon case increases, the trajectory of the finite-horizon case approximates the infinite-horizon case more closely. As expected, the longer the horizon becomes, the closer the solution of the finite-horizon problem will be to the infinite-time solution. This is due to the fact that the transient parts of the trajectories have less impact on the profit function for increasing length of the time horizon. As the same steady-state is reached by both finite- and infinite-time horizon formulations, the same profit is also obtained at steady-state.

These results show that also the finite-horizon formulation with economic objective can provide closed-loop stability, as the system gets on the turnpike (which corresponds to a stationary path in this case) and stays there for most of the time. Hence, by choosing a long time horizon it is possible to achieve closed-loop nominal stability, because the trajectory is attracted to the stationary path. However, as shown in Figure (1.a), if the finite time horizon is not long enough the resulting trajectories will not reach the optimal stationary path.

6. CONCLUDING REMARKS

The closed-loop solution of nonlinear DRTO or NMPC problems was studied. Nominal stability in closed-loop economic optimization is achieved via an infinite-horizon formulation. The comparison of the infinite- to the finite-horizon formulation shows that the formulation of DRTO problems on finite horizons can also provide closed-loop stability, if the optimization time horizon is chosen long enough such that the trajectory is attracted to a constant path.

Secondly, a new numerical approach was introduced for solving infinite-horizon problems addressed in NMPC and DRTO. The method achieves high computational accuracy because the infinite horizon is transformed into a finite horizon through a simple variable transformation. The advantage is that the truncation error occurring by choosing an arbitrary final time is avoided. Furthermore, the computational load is still low because of adaptive grid refinement resulting in low number of degrees of freedom for optimization.

In the future, it is of interest to further investigate the properties of the finite-horizon economic optimization problem required to achieve closed-loop stability. On the other hand, the infinite-horizon formulation with adaptive grid refinement is a promising approach to guarantee closed-loop stability. The solution method for infinite-horizon problems will be further developed and extended to open-loop unstable systems.

REFERENCES

Backx, T., Bosgra, O., and Marquardt, W. (2000). Integration of model-predictive control and optimization of processes. In *IFAC Symposium "Advanced Control of Chemical Processes", ADCHEM 2000*, 249–260. Pisa, Italy.

Barro, R. and Sala-i-Martin, X. (1995). *Economic Growth*. McGraw-Hill Advanced Series in Economics.

Carlson, D.A., Haurie, A.B., and Leizarowitz, A. (1991). *Infinite Horizon Optimal Control*. Springer Verlag.

Cass, D. (1966). Optimal growth in an aggregative model of capital accumulation, a turnpike theorem. *Econometrica*, 34, 833–850.

DyOS (2002). *DyOS User Manual*. Process Systems Engineering, Aachener Verfahrenstechnik, RWTH Aachen University.

Engell, S. (2007). Feedback control for optimal process operation. *Journal of Process Control*, 17, 203–219.

Forbes, J. (1994). *Model Structure and Adjustable Parameter Selection for Operations Optimizations*. Ph.D. thesis, McMaster University, Hamilton, Canada.

Helbig, A., Abel, O., and Marquardt, W. (2000). Structural concepts for optimization based control of transient processes. In F. Allgöwer and A. Zheng (eds.), *Nonlinear Model Predictive Control*, volume 26 of *Progress in Systems and Control Theory*, 295–311. Birkhäuser, Basel.

Huesman, A.E.M., Bosgra, O.H., and den Hof, P.M.J.V. (2008). Integrating MPC and RTO in the process industry by economic dynamic lexicographic optimization; an open-loop exploration. In *Proceedings of the 2008 AICHE Annual Meeting*. Philadelphia, USA.

Kunkel, P. and Hagen, O.V.D. (2000). Numerical solution of infinite-horizon optimal control problems. *Computational Economics*, 16(3).

Mayne, D.Q., Rawlings, J.B., RAO, C.V., and Scokaert, P.O.M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789–814.

Muske, K.R. and Rawlings, J.B. (1993). Model predictive control with linear models. *A.I.Ch.E. Journal*, 39(2), 262–287.

Rawlings, J.B. and Amrit, R. (2008). Optimizing process economic performance using model predictive control. In *Assessment and Future Directions of Nonlinear Model Predictive Control*. Pavia, Italy.

Schlegel, M., Stockmann, K., Binder, T., and Marquardt, W. (2005). Dynamic optimization using adaptive control vector parameterization. *Comp. & Chem. Eng.*, 29(8), 1731–1751.

Zavala, V. (2008). *Computational Strategies for the Optimal Operation of Large-Scale Chemical Processes*. Ph.D. thesis, Carnegie Mellon University.

Soft Constraints for Robust MPC of Uncertain Systems

Guru Prasath [*,**] **John Bagterp Jørgensen** [**,1]

[*] *FLSmidth Private Limited. Automation Division*
Chennai, India (e-mail: mgp-in@flsaindia.co.in)

[**] *Informatics and Mathematical Modelling*
Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
(e-mail: jbj@imm.dtu.dk).

Abstract: In this paper we develop a robust constrained predictive controller for linear systems. The controller is equipped with soft output constraints that are used in a novel way to have robustness against model plant mismatch. By simulation we compare the performance of the new robust constrained predictive controller to a nominal predictive controller. In the nominal case, the performance of the robust predictive controller is comparable to the performance of the nominal predictive controller. In the case of plant model mismatch, the robust predictive controller performs significantly better than the nominal predictive controller.

Keywords: Linear Model Predictive Control, Robust Predictive Control, Soft Constraints

1. INTRODUCTION

Model predictive control has become a standard technology in high level control of chemical processes. However, little advice is available regarding tuning methodologies of such controllers in the face of the inevitable plant model mismatch. The closed-loop performance of nominal linear model predictive control can be quite poor when the models are uncertain. Consequently, some years after commissioning, many high-level control systems are turned off due to bad closed-loop performance. This is often due to changes in the plant dynamics caused by wear and tear combined with lack of the necessary human resources at the plant to re-tune and maintain the MPC. Model predictive controllers with robust performance against model-plant mismatch is therefore crucial in long-term maintenance and success of MPC system. Using soft output constraints in a novel way, we demonstrate by simulation that the poor performance of predictive control in the case of plant model mismatch can be improved significantly. Therefore, we suggest use of the soft constraints to tune and improve the performance of linear model predictive control.

Specifically, we investigate the effect of uncertain models on the performance of a regularized l_2 model predictive controller with input constraints, input-rate constraints and soft output constraints (Maciejowski, 2002; Goodwin et al., 2005; Qin and Badgwell, 2003). Previously, the soft output constraints have been used to replace hard output constraint and guarantee feasibility (Scokaert and Rawlings, 1999). We use the soft output constraints to create a dead zone around the set point and demonstrate by simulation that the performance of such an MPC does not degrade much in the nominal case but improves significantly in the case of plant model mismatch. This technique is similar but not identical to the funnels used by Honeywell in RMPC (Qin and Badgwell, 2003; Havlena and Lu, 2005; Havlena and Findejs, 2005). Compared to classical process control, our use of the soft constraints has some similarities to PID control with dead zones (Shinskey, 1988).

We use a finite impulse response (FIR) model for prediction of the process outputs. In contrast to state space parameterizations, the FIR model is in a form that can easily be applied in robust predictive control, i.e. predictive control based on robust linear programming or second-order cone programming (Hansson, 2000; Vandenberghe et al., 2002; Boyd and Vandenberghe, 2004). To facilitate comparative performance studies of l_2 and robust MPC, a FIR based l_2-MPC benchmark has been established (Prasath and Jørgensen, 2008). The soft output constraint included in the MPC acts as a dead zone to the controller to reduce its sensitivity to noise and uncertainty when the process output is close to its target. This use of soft constraints for robustness is new, simple, and gives good performance. Bemporad and Morari (1999) provide an excellent survey of methodologies for robust model predictive control.

This paper is organized as follows. We derive the predictive controller consisting of a regulator and an estimator with soft output constraints in Section 2. Section 3 illustrates by simulation the performance of MPC with and without soft constraints for both deterministic and stochastic processes. Conclusions are given in Section 4.

2. FIR MODEL BASED MPC

Model predictive control systems consist of an estimator and a regulator as illustrated in Figure 1. The inputs to

[1] Corresponding author.

the MPC are the target values, r, for the process outputs, z, and the measured process outputs, y. The output from the MPC is the manipulated variables, u.

2.1 Plant and Sensors

The plant is assumed to be a linear state space system

$$\boldsymbol{x}_{k+1} = A\boldsymbol{x}_k + Bu_k + B_d d_k + G\boldsymbol{w}_k \tag{1a}$$
$$\boldsymbol{z}_k = C\boldsymbol{x}_k \tag{1b}$$

with x being the states, u being the manipulated variables (MVs), d being unmeasured disturbances, and w being stochastic process noise. z denotes the controlled variables (CVs). The measured outputs, y, are the controlled outputs, z, corrupted by measurement noise, v. Consequently

$$\boldsymbol{y}_k = \boldsymbol{z}_k + \boldsymbol{v}_k \tag{1c}$$

The initial state, the process noise, and the measurement noise are assumed to be normally distributed stochastic vectors

$$\boldsymbol{x}_0 \sim N(\bar{x}_0, P_0) \tag{2a}$$
$$\boldsymbol{w}_k \sim N_{iid}(0, Q) \tag{2b}$$
$$\boldsymbol{v}_k \sim N_{iid}(0, R) \tag{2c}$$

The measured output, y, is the signal available for feedback and used by the estimator. u is the signal generated by the control system and implemented on the plant.

2.2 Regulator

Stable processes can be represented by the finite impulse response (FIR) model

$$z_k = b_k + \sum_{i=1}^{n} H_i u_{k-i} \tag{3}$$

in which $\{H_i\}_{i=1}^n$ are the impulse response coefficients (Markov parameters). b_k is a bias term generated by the estimator. b_k accounts for discrepancies between the predicted output and the actual output. In this paper, the output predictions used by the regulator are based on the FIR model (3). Consequently, using the FIR model (3), the regularized l_2 output tracking problem with input and soft output constraints may be formulated as

$$\min_{\{z,u,\eta\}} \phi = \frac{1}{2}\sum_{k=0}^{N-1} \|z_{k+1} - r_{k+1}\|^2_{Q_z} + \|\Delta u_k\|^2_{S_u} + \sum_{k=1}^{N} \frac{1}{2}\|\eta_k\|^2_{S_\eta} + s'_{\eta_k}\eta_k \tag{4a}$$

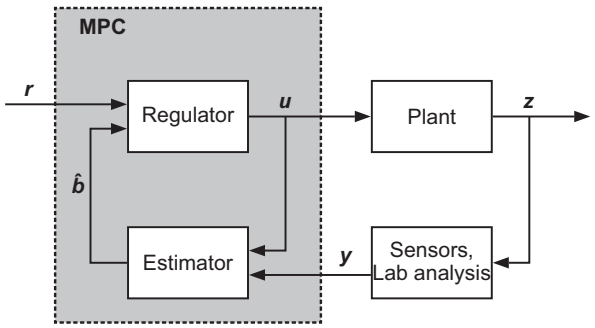


Fig. 1. Generic model predictive control system.

subject to the constraints

$$z_k = b_k + \sum_{i=1}^{n} H_i u_{k-i} \qquad k = 1, \ldots, N \tag{4b}$$
$$u_{\min} \le u_k \le u_{\max} \qquad k = 0, \ldots, N-1 \tag{4c}$$
$$\Delta u_{\min} \le \Delta u_k \le \Delta u_{\max} \qquad k = 0, \ldots, N-1 \tag{4d}$$
$$z_k \le z_{\max,k} + \eta_k \qquad k = 1, \ldots, N \tag{4e}$$
$$z_k \ge z_{\min,k} - \eta_k \qquad k = 1, \ldots, N \tag{4f}$$
$$\eta_k \ge 0 \qquad k = 1, \ldots, N \tag{4g}$$

in which $\Delta u_k = u_k - u_{k-1}$. In this formulation, the control and the prediction horizon are identical. If desired, a prediction horizon longer than the control horizon could be included in the formulation. However, we prefer instead to select the control horizon sufficiently long such that any boundary effects at the end of the horizon has no influence on the solution in the beginning of the horizon. (4) can be converted to a constrained linear-quadratic optimal control problem. Efficient algorithms exists for the solution of such problems with long prediction horizons, N. In this paper we adopt another approach and formulate a dense quadratic program in standard form that is equivalent with (4).

Define the vectors Z, R, U and η as

$$Z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_N \end{bmatrix} \quad R = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_N \end{bmatrix} \quad U = \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{N-1} \end{bmatrix} \quad \eta = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_N \end{bmatrix} \tag{5}$$

Then the predictions by the impulse response model (4) may be expressed as

$$Z = c + \Gamma U \tag{6}$$

For the case $N = 6$ and $n = 3$, Γ is assembled as

$$\Gamma = \begin{bmatrix} H_1 & 0 & 0 & 0 & 0 & 0 \\ H_2 & H_1 & 0 & 0 & 0 & 0 \\ H_3 & H_2 & H_1 & 0 & 0 & 0 \\ 0 & H_3 & H_2 & H_1 & 0 & 0 \\ 0 & 0 & H_3 & H_2 & H_1 & 0 \\ 0 & 0 & 0 & H_3 & H_2 & H_1 \end{bmatrix}$$

and c is

$$c = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \begin{bmatrix} b_1 + (H_2 u_{-1} + H_3 u_{-2}) \\ b_2 + (H_3 u_{-1}) \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix}$$

Similarly, for the case $N = 6$, define the matrices Λ and I_0 by

$$\Lambda = \begin{bmatrix} I & 0 & 0 & 0 & 0 & 0 \\ -I & I & 0 & 0 & 0 & 0 \\ 0 & -I & I & 0 & 0 & 0 \\ 0 & 0 & -I & I & 0 & 0 \\ 0 & 0 & 0 & -I & I & 0 \\ 0 & 0 & 0 & 0 & -I & I \end{bmatrix} \quad I_0 = \begin{bmatrix} I \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Define $s_\eta = \begin{bmatrix} s'_{\eta_1} & s'_{\eta_2} & \ldots & s'_{\eta_N} \end{bmatrix}'$ and

$$\mathcal{Q}_z = \begin{bmatrix} Q_z & & & \\ & Q_z & & \\ & & \ddots & \\ & & & Q_z \end{bmatrix} \quad \mathcal{S}_i = \begin{bmatrix} S_i & & & \\ & S_i & & \\ & & \ddots & \\ & & & S_i \end{bmatrix}$$

with $i = \{u, \eta\}$. Then the objective function (4) may be expressed as

$$\begin{aligned}
\phi &= \frac{1}{2}\sum_{k=0}^{N-1} \|z_{k+1} - r_{k+1}\|^2_{Q_z} + \|\Delta u_k\|^2_{S_u} \\
&\qquad + \frac{1}{2}\|\eta_{k+1}\|^2_{S_\eta} + s'_{\eta_{k+1}}\eta_{k+1} \\
&= \frac{1}{2}\|Z - R\|^2_{\mathcal{Q}_z} + \frac{1}{2}\|\Lambda U - I_0 u_{-1}\|^2_{\mathcal{S}_u} \\
&\qquad + \frac{1}{2}\|\eta\|^2_{\mathcal{S}_\eta} + s'_\eta\eta \\
&= \frac{1}{2}\|c + \Gamma U - R\|^2_{\mathcal{Q}_z} + \frac{1}{2}\|\Lambda U - I_0 u_{-1}\|^2_{\mathcal{S}_u} \\
&\quad + \frac{1}{2}\|\eta\|^2_{\mathcal{S}_\eta} + s'_\eta\eta \qquad (7) \\
&= \frac{1}{2}U'\left(\Gamma'\mathcal{Q}_z\Gamma + \Lambda'\mathcal{S}_u\Lambda\right)U \\
&\quad + \left(\Gamma'\mathcal{Q}_z(c - R) - \Lambda'\mathcal{S}_u I_0 u_{-1}\right)'U \\
&\quad + \left(\frac{1}{2}\|c - R\|^2_{\mathcal{Q}_z} + \frac{1}{2}\|I_0 u_{-1}\|^2_{\mathcal{S}_u}\right) \\
&\quad + \frac{1}{2}\eta'\mathcal{S}_\eta\eta + s'_\eta\eta \\
&= \frac{1}{2}U'HU + g'U + \rho + \frac{1}{2}\eta'\mathcal{S}_\eta\eta + s'_\eta\eta \\
&= \frac{1}{2}x'\bar{H}x + \bar{g}'x + \rho
\end{aligned}$$

with

$$H = \Gamma'\mathcal{Q}_z\Gamma + \Lambda'\mathcal{S}_u\Lambda \qquad (8a)$$

$$g = \Gamma'\mathcal{Q}_z(c - R) - \Lambda'\mathcal{S}_u I_0 u_{-1} \qquad (8b)$$

$$\rho = \frac{1}{2}\|c - R\|^2_{\mathcal{Q}_z} + \frac{1}{2}\|u_{-1}\|^2_{S_u} \qquad (8c)$$

$$x = \begin{bmatrix} U \\ \eta \end{bmatrix} \quad \bar{H} = \begin{bmatrix} H & 0 \\ 0 & \mathcal{S}_\eta \end{bmatrix} \quad \bar{g} = \begin{bmatrix} g \\ s_\eta \end{bmatrix} \qquad (8d)$$

Consequently, we may solve the FIR based MPC regulator problem (4) by solution of the following convex quadratic program

$$\min_x \quad \psi = \frac{1}{2}x'\bar{H}x + \bar{g}'x \qquad (9a)$$

$$s.t. \quad x_{\min} \le x \le x_{\max} \qquad (9b)$$

$$b_l \le \bar{A}x \le b_u \qquad (9c)$$

in which

$$x_{\min} = \begin{bmatrix} U_{\min} \\ 0 \end{bmatrix} \quad x_{\max} = \begin{bmatrix} U_{\max} \\ \infty \end{bmatrix} \qquad (10a)$$

$$b_l = \begin{bmatrix} \Delta U_{\min} \\ -\infty \\ Z_{\min} - c \end{bmatrix} A = \begin{bmatrix} \Lambda & 0 \\ \Gamma & -I \\ \Gamma & I \end{bmatrix} b_u = \begin{bmatrix} \Delta U_{\max} \\ Z_{\max} - c \\ \infty \end{bmatrix} \qquad (10b)$$

In a model predictive controller only the first vector, u_0^*, of $U^* = \left[(u_0^*)'\ (u_1^*)'\ \dots\ (u_{N-1}^*)'\right]'$, is implemented on the process. At the next sample time the open-loop optimization is repeated with new information due to a new measurement.

2.3 Soft Constraint Principle

Figure 2 illustrates the stage cost function for l_2 model predictive control (nominal MPC) and l_2 model predictive control with a dead zone (soft MPC). The stage cost

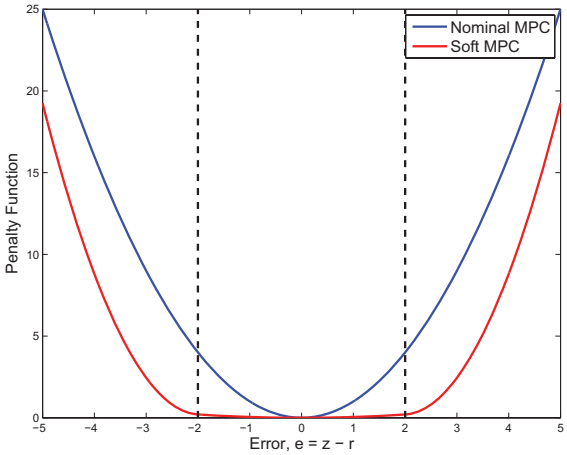


Fig. 2. The set point deviation penalty function for nominal MPC and soft MPC.

function, or penalty function, is plotted as function of the set-point error, $e = z - r$. The penalty function of the nominal MPC is a quadratic function. The penalty function of the soft MPC is constructed such that it is zero or almost zero within the dead-zone between the soft limits and growths quadratically when the set-point error exceeds the soft limits. The small penalty within the soft limits ensures that the controller produces a steady state offset free response. By having the penalty small within the soft constraints, the controller does not react much to small errors. In this way we avoid that the controller introduces significant real disturbances to the process because it reacts to say measurement noise or plant-model mismatch. Outside the soft limits, it is assumed that the deviation from target is due to a real process disturbance, and the soft MPC may be designed to react in the same way as the nominal MPC.

2.4 Simple Estimator

To have offset free steady state control when unknown step responses occur, we must have integrators in the feedback loop. This may be achieved using a FIR model in difference variables. Assume that the relation between the inputs and outputs may be represented as

$$\Delta y_k = \Delta z_k = e_k + \sum_{i=1}^{n} H_i \Delta u_{k-i} \qquad (11)$$

in which Δ is the backward difference operator, i.e. $\Delta y_k = y_k - y_{k-1}$, $\Delta z_k = z_k - z_{k-1}$, and $\Delta u_k = u_k - u_{k-1}$. This representation is identical with the FIR model (3)

$$y_k = z_k = \hat{b}_k + \sum_{i=1}^{n} H_i u_{k-i} \qquad (12)$$

if $\hat{b}_k$ is computed by

$$e_k = \Delta y_k - \sum_{i=1}^{n} H_i \Delta u_{k-i} \qquad (13a)$$

$$\hat{b}_k = \hat{b}_{k-1} + e_k \qquad (13b)$$

Note that in the regulator optimization problem $b_1 = b_2 = \dots = b_N = \hat{b}_k$ at each time instant. This is based on the assumption that the disturbances enter the process

as constant output disturbances. Of course this may not be how the disturbances enter the process in practice, and significant performance deterioration may result as a consequence of this representation.

3. SIMULATIONS

In this Section we consider plants of the form

$$\boldsymbol{Z}(s) = G(s)U(s) + G_d(s)\left(D(s) + \boldsymbol{W}(s)\right) \tag{14a}$$

$$\boldsymbol{y}(t_k) = \boldsymbol{z}(t_k) + \boldsymbol{v}(t_k) \tag{14b}$$

with the transfer functions

$$G(s) = \frac{K(\beta s+1)}{(\tau_1 s+1)(\tau_2 s+1)}\, e^{-\tau s} \tag{15a}$$

$$G_d(s) = \frac{K_d(\beta_d s+1)}{(\tau_{d1} s+1)(\tau_{d2} s+1)}\, e^{-\tau_d s} \tag{15b}$$

The disturbance model, $G_d(s)$, is kept fixed at its nominal value, while the transfer function, $G(s)$, from $U(s)$ varies around its nominal value, $G_0(s)$. This is used to illustrate the consequence of model uncertainty on the MPC closed-loop performance. The nominal system is $K = K_d = 1$, $\tau_1 = \tau_2 = \tau_{d1} = \tau_{d2} = 5$, $\beta = \beta_d = 2$, and $\tau = \tau_d = 5$. The system is converted to a discrete time state space model (1) using a sample time of $T_s = 1$ and a zero-order-hold assumption on the inputs.

The predictive controller is based on the impulse response coefficients of the following system

$$Z(s) = \hat{G}(s)U(s) \tag{16}$$

in which $\hat{G}(s)$ is identical to the nominal plant $G_0(s)$.

The simple estimator described in Section 2.4 is used for bias estimation. The input limits are $u_{\min} = -1$, $u_{\max} = 1$, $\Delta u_{\min} = -0.2$, and $\Delta u_{\max} = 0.2$. The horizon of the impulse response model is $n = 40$ and the control horizon is $N = 120$. The MPC is tuned with $Q_z = 1$ and $S = 10^{-3}$.

The unknown deterministic process disturbance, $D(s)$, the stochastic process disturbances, $\boldsymbol{W}(s)$ or $\boldsymbol{w}_k$, and the measurement noise, $\boldsymbol{v}(t_k) = \boldsymbol{v}_k$, used in the simulations are illustrated in Figure 3. The stochastic process disturbances is $\boldsymbol{w}_k \sim N(0, 0.01)$, and the stochastic measurement noise is $\boldsymbol{v}_k \sim N(0, 0.01)$.

3.1 Nominal Stochastic System

We consider the case when the model used by the controller is identical to the deterministic part of the plant model. However, the plant has in addition to the deterministic part stochastic process disturbances and stochastic measurement noise as illustrated in Figure 3.

Consider the case with no determistic disturbance, i.e. $D(s) = 0$. The performances of the nominal MPC and the soft MPC applied to this system are compared in Figure 4. The output variances produced by the two controllers are almost identical, while the input variance of the soft MPC is much smaller than the input variance of the nominal MPC. Due to the low penalties within the soft limits, the soft MPC does not react to measurement noise and do not need to compensate such previous erroneous measurement noise induced input moves.

Figure 5 illustrates the performance of the nominal MPC and the soft MPC when the model is identical to the

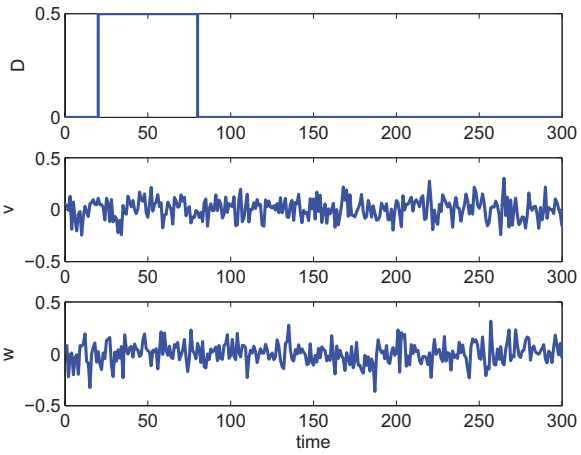


Fig. 3. External signals used in the closed loop simulations. $D(s)$ or d_k is the unknown deterministic disturbance, $\boldsymbol{v}_k$ is the stochastic measurement noise, and $\boldsymbol{w}_k$ is the stochastic process noise,

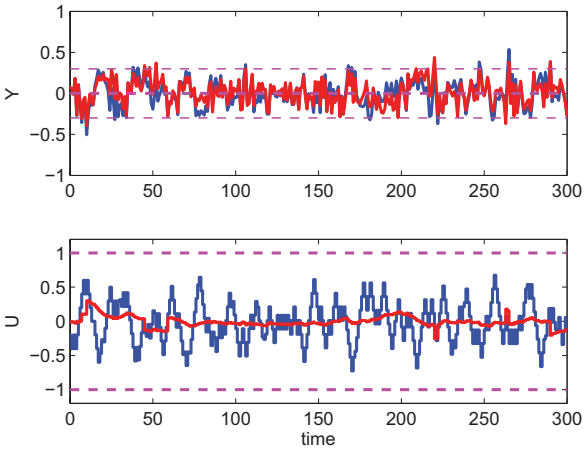


Fig. 4. Comparison of normal and soft MPC with nominal models applied to a stochastic system with no deterministic disturbance (Nominal MPC = blue, Soft MPC = red).

plant model and the external signals illustrated in Figure 3 are applied to the model (14). Also in this case, the controlled variable, Y (or Z), of the two controllers are similar while the manipulated variable, U, of the soft MPC has significantly less variance than the manipulated variable, U, of the nominal MPC.

3.2 Uncertain Determistic System

We consider a deterministic system without stochastic process noise nor stochastic measurement noise. However, the model used by the controllers is different from the plant model. The process is perturbed by an unknown deterministic disturbance, $D(s)$, as illustrated in figure 3.

We compare the performance of the nominal MPC and the soft MPC for model-plant mismatches defined by the time delay, τ, the gain K, the time constant τ_1, and the zero β.

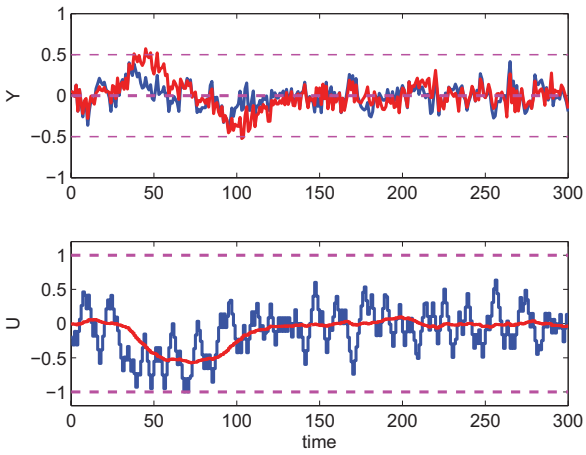


Fig. 5. Comparison of normal and soft MPC with nominal models applied to a stochastic system with an unknown deterministic disturbance (Nominal MPC = blue, Soft MPC = red). The external signals are shown in Figure 3.

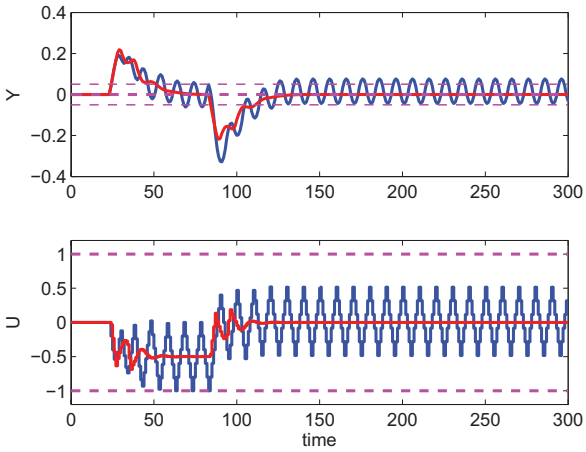


Fig. 6. Closed-loop MPC performance with time delay uncertainty. The plant delay is $\tau = 3$ and the model delay is $\hat{\tau} = 5$ (Nominal MPC = blue, Soft MPC = red).

Figure 6 and Figure 7 illustrate closed-loop performances achieved by the nominal and soft MPC when there is time delay plant-model mismatch. The soft MPC has smaller input variation than the nominal MPC, and the soft MPC provies better control than the nominal MPC in terms of set point deviations.

Figures 8-10 illustrate the performances of the nominal MPC and the soft MPC in the case of model-plant mismatch in the gain, the time constant, and the zero, respectively. In all cases, the soft MPC has significantly less input variation than the nominal MPC. Furthermore, the outputs are significantly better controlled by the soft MPC than by the nominal MPC.

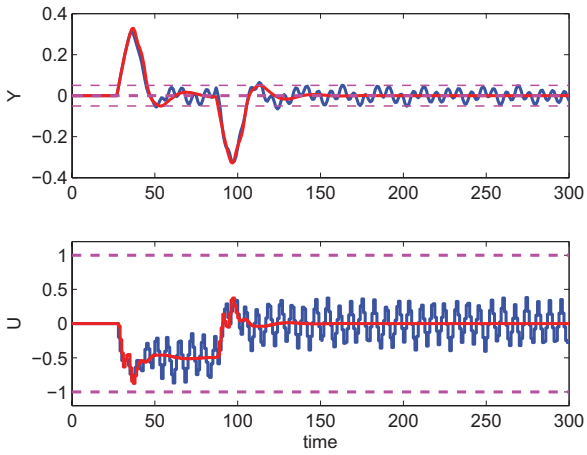


Fig. 7. Closed-loop MPC performance with time delay uncertainty. The plant delay is $\tau = 7$ and the model delay is $\hat{\tau} = 5$ (Nominal MPC = blue, Soft MPC = red).

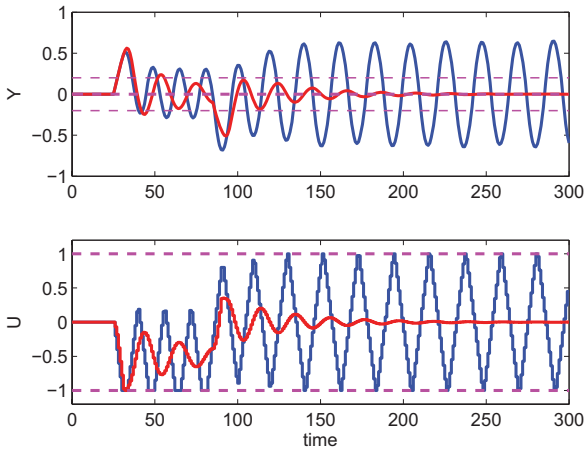


Fig. 8. Closed-loop MPC performance with gain uncertainty. The plant gain is $K = 2$ and the model gain is $\hat{K} = 1$ (Nominal MPC = blue, Soft MPC = red).

3.3 Uncertain Stochastic System

Figure 11 illustrates the closed loop performance of a nominal MPC and a soft MPC applied to the system (14) with the external signals in Figure 3 and a plant-model mismatch in the gain. The plant gain is $K = 2$ and the model gain is $\hat{K} = 1$. By inspection, it is obvious that the performance of the soft MPC is significantly better than the performance of the nominal MPC. The superior performance is achieved by having a small set point deviation penalty within the soft constraints such that the controller does not react aggressively when close to the set point. In this way it avoids perturbing the system due to stochastic measurement noise and plant-model mismatch.

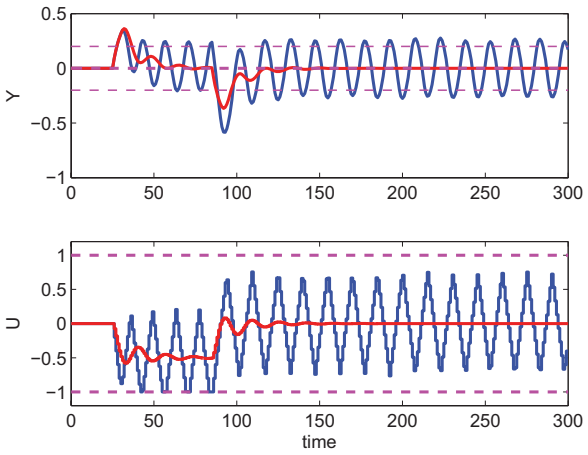


Fig. 9. Closed-loop MPC performance with time constant uncertainty. The plant time constant is $\tau_1 = 2$ and the model time constant is $\hat{\tau}_1 = 5$ (Nominal MPC = blue, Soft MPC = red).

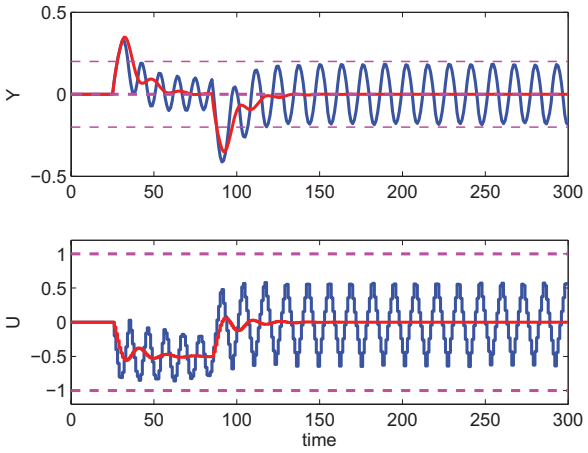


Fig. 10. Closed-loop MPC performance with zero uncertainty. The plant zero is $\beta = 4.5$ and the model zero is $\hat{\beta} = 2$ (Nominal MPC = blue, Soft MPC = red).

4. CONCLUSION

We have developed a l_2 regularized predictive controller with soft constraints and demonstrated efficient application of this controller to uncertain stochastic systems. We call this controller soft MPC. It is illustrated and verified by simulations that this soft MPC provides significantly better closed loop performance than nominal MPC. The soft MPC also provides much better performance degradation in the face of plant-model mismatch than nominal MPC. These features are expected to contribute to better closed loop performance, easier maintenance, easier tuning, and longer lifetime of model predictive controllers for chemical processes.

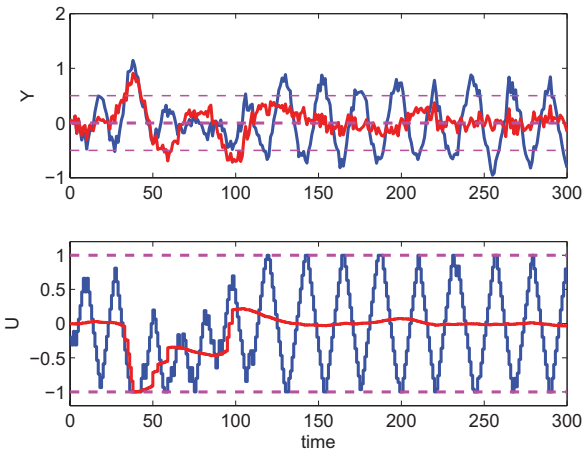


Fig. 11. Closed-loop MPC performance for a system with the external signals in Figure 3 and gain uncertainty. The plant gain is $K = 2$ and the model gain is $\hat{K} = 1$. Nominal MPC = blue, Soft MPC = red.

REFERENCES

Bemporad, A. and Morari, M. (1999). Robust model predictive control: A survey. In A. Garulli, A. Tesi, and A. Vicino (eds.), *Robustness in Identification and Control*, volume 245 of *Lecture Notes in Control and Information Sciences*, 207–226. Springer-Verlag.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press, Cambridge, UK.

Goodwin, G.C., Seron, M.M., and De Dona, J.A. (2005). *Constrained Control and Estimation. An Optimization Approach*. Springer, New York.

Hansson, A. (2000). A primal-dual interior-point method for robust optimal control of linear discrete-time systems. *IEEE Transactions on Automatic Control*, 45, 1639–1655.

Havlena, V. and Lu, J. (2005). A distributed automation framework for plant-wide control, optimisation, scheduling and planning. In *IFAC World Congress 2005*. IFAC, Prague, Czech Republic.

Havlena, V. and Findejs, J. (2005). Application of model predictive control to advanced combustion control. *Control Engineering Practice*, 13, 671–680.

Maciejowski, J.M. (2002). *Predictive Control with Constraints*. Prentice Hall, Harlow, England.

Prasath, G. and Jørgensen, J.B. (2008). Model predictive control based on finite impulse response models. In *ACC 2008*. American Control Conference 2008, Seattle, Washington.

Qin, S.J. and Badgwell, T.A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11, 733–764.

Scokaert, P.O.M. and Rawlings, J.B. (1999). Feasibility issues in linear model predictive control. *AIChE Journal*, 45, 1649–1659.

Shinskey, F.G. (1988). *Process Control Systems. Application, Design, and Tuning*. McGraw-Hill, New York.

Vandenberghe, L., Boyd, S., and Nouralishadi, M. (2002). Robust linear programming and optimal control. In *15th IFAC World Congress 2002*. IFAC, Barcelona, Spain.

Dynamic Operability for the Calculation of Transient Output Constraints for Non-Square Linear Model Predictive Controllers

Fernando V. Lima and Christos Georgakis*

Department of Chemical and Biological Engineering &
Systems Research Institute, Tufts University, Medford,
*MA, 02155, USA (*e-mail:christos.georgakis@tufts.edu)*

Abstract: This paper introduces a dynamic operability-based approach for the determination of feasible output constraints during transient operation. This approach is based on previously published steady-state operability developments and the concept of output funnels. In this study, high-dimensional non-square systems with more outputs than inputs are of particular interest. Such systems are challenging because it is impossible to control all the outputs at specific set-points when there are fewer degrees of freedom available than the controlled variables. Thus, interval, instead of set-point, control is needed for at least some of the output variables. In order to motivate the new concepts, two non-square case studies are addressed, one illustrative and one industrial - obtained from the control system of a Steam Methane Reformer process. The calculated constraints are validated by running DMCplus™ (AspenTech) closed-loop simulations for the extreme values of the disturbances. These constraints are intended for use online in model-based controllers (e.g., Model Predictive Controllers) to ensure that each of the outputs will remain inside a feasibility envelope during transient operation.

Keywords: Output Variables, Constraints, Model-based Control, Operability, Dynamic Systems.

1. INTRODUCTION

Model Predictive Control (MPC) is a long standing multivariable constrained control methodology that utilizes an explicit process model to predict the future behavior of a chemical plant. At each control interval, the MPC algorithm attempts to optimize the future plant behavior by computing a sequence of future manipulated variable adjustments. The first of the optimal sequence of calculated input moves is implemented into the plant and the entire calculation is repeated at subsequent control intervals using updated process measurements. MPC has been extensively studied in academia and widely accepted in the chemical industry for its ability to handle complex multivariable and highly interactive process control problems (Qin and Badgwell, 2003). MPC-type controllers in industrial practice aim to control non-square systems in which there are more controlled outputs than manipulated inputs. In such systems it is impossible to control all the outputs at specific set-points because there are fewer degrees of freedom available than the controlled variables.

Based on the input constraints, generally specified *a priori* due to the physical limitations of the process, an important design task is to determine the output ranges within which one wants to control the process. The improper selection of these constraints can make the controller infeasible when a disturbance moves the process far away from its usual operating region. Past practice requires that output constraints are enforced whenever feasible and softened whenever they become infeasible (Rawlings, 2000). The steady-state operability methodology originally introduced for square systems (Vinson and Georgakis, 2000) and extended for non-square systems (Lima and Georgakis, 2006; Lima and Georgakis, 2008a) provides a method for selecting such output constraints systematically, so that they are as tight as possible but also do not render the controller infeasible. Specifically for non-square systems, the interval operability framework was introduced (Lima and Georgakis, 2006) to assess the input-output open-loop operability of multivariable non-square systems at the steady-state, a necessary condition for the overall process operability. The application of this framework to high-dimensional square and non-square systems is discussed in another set of publications (Lima and Georgakis, 2008b; Lima, Georgakis, Smith and Schnelle, 2008), where a Linear Programming (LP) based approach is introduced to calculate the tightest feasible set of steady-state output constraints when interval operability is necessary.

This paper extends this interval operability framework to enable the determination of feasible output constraints during transient for high-dimensional non-square systems. Although the previously developed steady-state operability approaches are necessary to quantify the overall operability of a process and to determine the steady-state output constraints for MPC, the development of a dynamic operability methodology for non-square systems will have great impact on MPC controller design. Specifically, dynamic operability analysis can be used to systematically calculate the amount of constraint relaxation necessary in order to prevent the occurrence of transient infeasibilities, when disturbances affect the process (see Dimitriadis and Pistikopoulos (1995) for dynamic flexibility analysis). This extension is accomplished here by designing a funnel for each of the output variables, which provides output

constraints that guarantee feasible process operation in closed-loop. Previously, output funnels have been used to define MPC controllers' output trajectories in commercial packages (Qin and Badgwell, 2003). Specifically, Honeywell's RMPCT (Robust Multivariable Predictive Control Technology) controller defines a funnel for the outputs or Controlled Variables (CV) constraints. When a predicted CV trajectory leaves its funnel, the controller algorithm penalizes this trajectory to bring the CV back within its range (Qin and Badgwell, 2003; Maciejowski, 2002). Here such funnels are used to design output constraints during transient operation. This design is especially important for underdamped systems in general, where overshoots may occur during process operation, and overdamped or critically damped systems when disturbance dynamics are faster than input dynamics. For the opposite case when input dynamics are faster than disturbance ones, the output constraints calculated using one of the steady-state operability methodologies are also applicable during transient.

2. PROCESS OPERABILITY

Before introducing the dynamic operability approach, it is necessary to briefly define the sets of variables used for steady-state interval operability calculations (Lima and Georgakis, 2008a). The Available Input Set (**AIS**) is the set of values that the process input, or manipulated, variables (**u**) can take, based on the constraints of the process. For an $n \times m \times q$ (*n outputs*, *m inputs* and *q disturbances*) linear system:

$$\mathbf{AIS} = \left\{ \mathbf{u} \mid u_i^{\min} \le u_i \le u_i^{\max};\ 1 \le i \le m \right\} \tag{1}$$

The Desired Output Set (**DOS**) is given by the ranges of the outputs (**y**) that are desired to be achieved and is represented by:

$$\mathbf{DOS} = \left\{ \mathbf{y} \mid y_i^{\min} \le y_i \le y_i^{\max};\ 1 \le i \le n \right\} \tag{2}$$

The Expected Disturbance Set (**EDS**) represents the expected steady-state values of the disturbances (**d**):

$$\mathbf{EDS} = \left\{ \mathbf{d} \mid d_i^{\min} \le d_i \le d_i^{\max};\ 1 \le i \le q \right\} \tag{3}$$

Based on the steady-state linear model of the process, expressed by the process gain matrix (**G**) and the disturbance gain matrix ($\mathbf{G}_d$), the Achievable Output Set for a specific disturbance vector, **AOS(d)**, is defined by the ranges of the outputs that can be achieved using the inputs inside the **AIS**:

$$\mathbf{AOS}(\mathbf{d}) = \left\{ \mathbf{y} \mid \mathbf{y} = \mathbf{G}\mathbf{u} + \mathbf{G}_d\mathbf{d};\ \mathbf{u} \in \mathbf{AIS},\ \mathbf{d} \text{ is fixed} \right\} \tag{4}$$

The Achievable Output Interval Set (**AOIS**) is defined as the tightest feasible set of output constraints that can be achieved, with the available range of the manipulated variables and when the disturbances remain within their expected values (see references (Lima and Georgakis, 2008a; Lima and Georgakis, 2008b; Lima, Georgakis, Smith and Schnelle, 2008) for the algorithms developed for the calculation of this important set). Using these defined sets and some of the previously published interval operability concepts and calculations, a dynamic operability approach, based on the design of output funnels, for the determination of output constraints during transients is introduced next through a 2-D illustrative example. This is followed by the analysis of the Steam Methane Reformer (SMR) process example, which is 9-D and underdamped.

3. ILLUSTRATIVE EXAMPLE

In order to introduce the dynamic operability approach, consider a 2-D example from Lima and Georgakis (2006) with 2 outputs, 1 input and 1 disturbance (2 x 1 x 1). This example has the following steady-state gain model and constraining sets (see information on process dynamics below):

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1.41 \\ 0.66 \end{pmatrix} u_1 + \begin{pmatrix} -0.6 \\ 0.4 \end{pmatrix} d_1 \tag{5}$$

$$\mathbf{AIS} = \{ u_1 \mid -1 \le u_1 \le 1 \}$$

$$\mathbf{EDS} = \{ d_1 \mid -1 \le d_1 \le 1 \}$$

$$\mathbf{DOS} = \left\{ \mathbf{y} \in \Re^2 \mid \|y\|_\infty \le 1 \right\}$$

Two funnels, one for each output, with specific amplitudes and decay characteristics will be designed for the two output variables of this system. Each of these funnels is designed from the moment that a disturbance is inserted into the system and provides an envelope where the control problem is always feasible if the output constraints remain inside of this envelope. This envelope starts at the funnel amplitude value (defined below), decays at a specific rate and ends at a designed steady-state constraint calculated using one of the interval operability approaches cited above. Cases where the disturbance variable takes its extreme values are of particular interest because they represent the worst cases, which if satisfied, ensure feasible operation for all the other cases.

For the system above, the dynamics of each of the input-output and disturbance-output pairs are plotted in Figs. 1, 2, 3, and 4 for pairs (y_1, u_1), (y_1, d_1), (y_2, d_1), and (y_2, u_1), respectively. In these figures, these dynamics are represented by step response coefficients, which would be obtained in practice by plant testing.

The funnel amplitude associated with output i (a_i) is defined as follows:

$$a_i = k_{d,j-i} s_j \tag{6}$$

where $k_{d,j-i}$ corresponds to the value of the steady-state disturbance gain associated with the disturbance-output pair j-i and s_j is the step disturbance value. For this example, s_j will be assumed at the extreme values of the disturbance within the **EDS**, i.e. d_1 is moved from 0 to ± 1. If $d_1 = 1$, then $a_1 = -0.6$ and $a_2 = 0.4$. When $d_1 = -1$, $a_1 = 0.6$ and $a_2 = -0.4$. It is assumed that using steady-state disturbance gains to calculate the starting point of the funnel decay, as opposed to the maximum absolute value of the dynamic gains, will be enough to provide an envelope that contains the entire closed-loop response. This is based on the assumption that the inputs are able to compensate for the presence of overshoots, caused by these dynamic gains, in most practical cases during closed-loop operation, especially if a model-based controller, such as MPC, is implemented. The decay for each output funnel (λ_i) is determined by the slowest dynamics among all

the input-output and disturbance-output pairs for the corresponding output. Each output is analyzed separately because the disturbance dynamics might be slower for one of the outputs, while the input dynamics may be slower for the other. These dynamics are estimated from the step response curves using two approaches, depending on the characteristics of the analyzed curve:

1) Exponential fit (typically for oscillatory responses): an exponential is fitted to two selected points of the step response curve. These points are selected such that most of the curve is below (or above, depending on the sign of the dynamic gains) the fitted exponential. The dynamics of the analyzed pair are estimated by the following exponential decay:

$$y_{\exp} = a_{\exp} \exp\left(-\lambda_{\exp} t\right) + y_{\infty} \tag{7}$$

where y_{∞} corresponds to the steady-state gain of the analyzed step response curve. Using the two selected points and eq. (7), a system of 2 equations and 2 unknowns can be solved for the two parameters of the exponential, $a_{\exp}$ and $\lambda_{\exp}$. For the example above, this approach is used for pairs (y_1, u_1), (y_1, d_1), and (y_2, d_1), whose exponential fits are shown in Figs. 1, 2, and 3, respectively, along with the fitted points selected for each case. For such pairs, the following exponential fits are obtained:

$$\begin{aligned}
&\left(y_1,\, u_1\right) = 0.47\exp\left(-1.89\times10^{-2} t\right) + 1.41 \Rightarrow \lambda_{\exp} = 1.89\times10^{-2} \\
&\left(y_1,\, d_1\right) = -1.70\exp\left(-2.12\times10^{-2} t\right) - 0.60 \Rightarrow \lambda_{\exp} = 2.12\times10^{-2} \\
&\left(y_2,\, d_1\right) = 0.18\exp\left(-4.22\times10^{-2} t\right) + 0.40 \Rightarrow \lambda_{\exp} = 4.22\times10^{-2}
\end{aligned} \tag{8}$$

2) First-order models estimated using ARX (Auto-Regressive model with eXogenous inputs, subroutine ARX in Matlab) (Ljung, 1999): the following first-order ARX model with a zero-order holder in the z domain is fitted to the step response coefficients of a specific pair:

$$y(z) = \frac{bz^{-1}}{1+az^{-1}} u(z) \text{ with } \lambda_{\text{arx}} = -\ln\left(p_z\right) \tag{9}$$

where λ_{arx} represents the model dynamics and is calculated by taking the negative natural log of the transfer function pole in the z domain, $p_z = -a$. This approach is used whenever a pair dynamics can be well approximated by a first-order model. This approach is applied here to pair (y_2, u_1), which is shown in Fig. 4, and the following model and λ_{arx} are obtained:

$$y_2(z) = \frac{0.0283}{z-0.9601} u_1(z) \Rightarrow \lambda_{\text{arx}} = -\ln\left(0.9601\right) = 4.07\times10^{-2} \tag{10}$$

After calculating all λs for all possible pairs (4 in this case), using the two approaches above, their values are compared and the one with smallest absolute value for each output (representing the slowest dynamics) is retained and used in the funnel design for the corresponding output. For example, for output 1, $\lambda_1 = 1.89 \times 10^{-2}$ is chosen, which is the smallest between 1.89×10^{-2} and 2.12×10^{-2}. Therefore, for this example, the following values of λ_i are selected:

$$\lambda_1 = 1.89\times10^{-2},\ \lambda_2 = 4.07\times10^{-2} \tag{11}$$

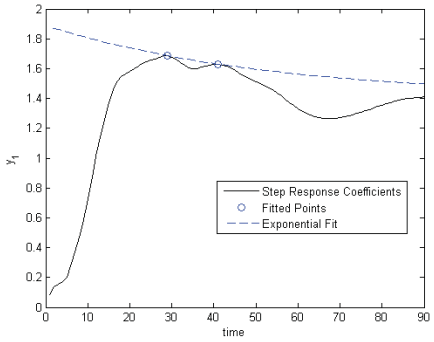


Fig. 1. Step Response Coefficients, Exponential Fit and Fitted Points for (y_1, u_1) pair.

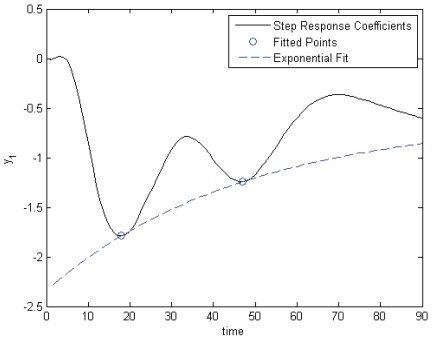


Fig. 2. Step Response Coefficients, Exponential Fit and Fitted Points for (y_1, d_1) pair.

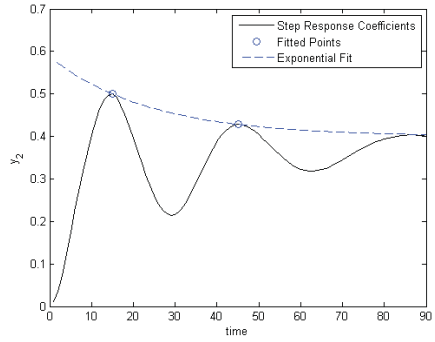


Fig. 3. Step Response Coefficients, Exponential Fit and Fitted Points for (y_2, d_1) pair.

Using the calculated amplitudes and decays, the following equation represents the funnel for each of the output variables:

$$f_i = \left(1+\alpha_f\right) a_i \exp\left[-\left(1+\beta_f\right)\lambda_i t\right] + y_{ss,i} \tag{12}$$

where $y_{ss,i}$ is one of the steady-state output constraints (upper or lower limit) for output i, which is calculated using the previously published interval operability approaches (Lima and Georgakis, 2008a; Lima and Georgakis, 2008b). Also, α_f

and β_f are adjustable tuning parameters, associated with amplitude and decay, respectively, and are independent of the output selected. As explained above for a_i, depending on the magnitude of the disturbance inserted, which in this case takes either its maximum or minimum value, $y_{ss,i}$ will have different values. For the case study here, if $d_1 = 1$, then $y_{ss,1} = -0.464$ and $y_{ss,2} = 0.464$. When $d_1 = -1$, $y_{ss,1} = 0.464$ and $y_{ss,2} = -0.464$. These values were extracted from the steady-state operability results presented in the ADCHEM 2006 paper by Lima and Georgakis (see case 2, section 3). Thus, each output envelope actually has upper and lower limits that start at different points and end at the upper and lower calculated steady-state constraints, respectively. The same decay holds for both cases. Therefore, for the 2-D case study above, selecting $\alpha_f = 0$ and $\beta_f = -0.74$, the following funnels are obtained for each output for the two extreme values of the disturbance:

$$\text{For } d_1 = 1:\ y_{ss,1} = -0.464,\ y_{ss,2} = 0.464$$

$$f_1 = -0.6\exp\left[-(1-0.74)1.89\times10^{-2}t\right] - 0.464 \quad (13)$$

$$f_2 = 0.4\exp\left[-(1-0.74)4.07\times10^{-2}t\right] + 0.464 \quad (14)$$

$$\text{For } d_1 = -1:\ y_{ss,1} = 0.464,\ y_{ss,2} = -0.464$$

$$f_1 = 0.6\exp\left[-(1-0.74)1.89\times10^{-2}t\right] + 0.464 \quad (15)$$

$$f_2 = -0.4\exp\left[-(1-0.74)4.07\times10^{-2}t\right] - 0.464 \quad (16)$$

The funnels for outputs 1 (eqs. 13 and 15) and 2 (eqs. 14 and 16) are plotted in Figs. 5 and 6, respectively, along with the DMCplus™ (Dynamic Matrix Control - AspenTech, a multivariable constrained controller) trend obtained for each case. For all cases, the controller is operating in closed-loop mode and the disturbance was inserted at time = 0.

4. HIGH-DIMENSIONAL INDUSTRIAL SYSTEM

The design of output constraints during transient for the Steam Methane Reformer (SMR; Vinson, 2000) process example will now be performed using the output funnels defined above. This process has 9 outputs, 4 inputs and 1 disturbance variable and it is defined by the following set of steady-state equations and constraining sets (see information on process dynamics below):

$$\begin{pmatrix}\delta y_1\\ \delta y_2\\ \delta y_3\\ \delta y_4\\ \delta y_5\\ \delta y_6\\ \delta y_7\\ \delta y_8\\ \delta y_9\end{pmatrix} = \begin{pmatrix} 1.00 & 0 & 3.99 & 2.22\\ 2.50 & 0.25 & 16.51 & -11.80\\ -0.14 & 0 & -1.53 & -1.45\\ 0.55 & 0 & 2.76 & 0.69\\ -0.04 & 0.02 & 0 & 0\\ 2.34 & -0.66 & 36.71 & 5.09\\ 3.96 & -0.19 & 44.40 & 5.79\\ -0.04 & 0 & -3.11 & -1.71\\ 0.36 & 0.05 & 1.10 & -16.54\end{pmatrix}\begin{pmatrix}\delta u_1\\ \delta u_2\\ \delta u_3\\ \delta u_4\end{pmatrix} + \begin{pmatrix}0.40\\ 1.65\\ -0.15\\ 0\\ 0\\ 3.67\\ 4.44\\ -0.31\\ 0.11\end{pmatrix}\delta d_1 \quad (17)$$

$$\mathbf{AIS} = \left\{\begin{array}{l}\mathbf{u}\in\Re^4 \mid 10\le u_1\le 48;\ \ 60\le u_2\le 140;\\ \qquad 0.2\le u_3\le 2.0;\ \ -2.4\le u_4\le -0.7\end{array}\right\} \quad (18)$$

$$\mathbf{EDS} = \{d_1 \mid -4\le d_1\le 4\}$$

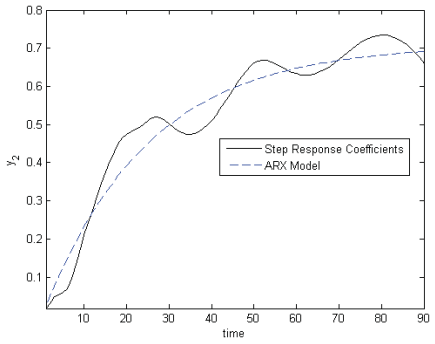


Fig. 4. Step Response Coefficients, Exponential Fit and First-order ARX Model for (y_2, u_1) pair.

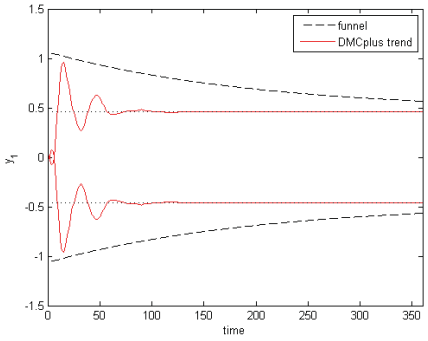


Fig. 5. Funnel Design and DMCplus trend for output y_1 with (α_f, β_f) = (0, -0.74).

where $\delta\mathbf{y}$, $\delta\mathbf{u}$ and δd_1 are deviation variables from the steady-state values for the outputs ($\mathbf{y}_{ss}$), the inputs ($\mathbf{u}_{ss}$), and the disturbance ($d_{1,ss}$), respectively. These steady-state values are given by:

$$\mathbf{y}_{ss} = (44.35,\ 94.10,\ 1.50,\ 21.5,\ 1.80,\ 431.45,\ 510.75,\ 5.35,\ 37.1)^T \quad (19)$$
$$\mathbf{u}_{ss} = (29.00,\ 100.00,\ 1.10,\ -1.55)^T;\ d_{1,ss} = 0$$

Also, the original output constraints (**DOS**), lower and upper limits, are given in Table 1. The SMR process has underdamped dynamics for several input-output/disturbance-output pairs, which are represented by the step response coefficients obtained by plant testing that are shown in Fig. 7. For this case, the disturbance gains, given in eq. (17), and the designed output constraints at the steady-state in Table 1 (from Lima, Georgakis, Smith, Vinson and Schnelle, 2009) will be used here to define each output funnel.

Following the same procedure as in the illustrative example above, exponential and ARX fits were obtained and λs calculated for all the input-output and disturbance-output pairs. The calculated λs for all these pairs are presented in Table 2, where the smallest absolute values of λ for each output, which will be used in the funnel design, are highlighted. Thus, using eq. (12), the funnel equations (20) and (21) are calculated for all outputs when d_1 moves from 0 to ± 4 (extreme cases) and (α_f, β_f) = (1.00, 0.36).

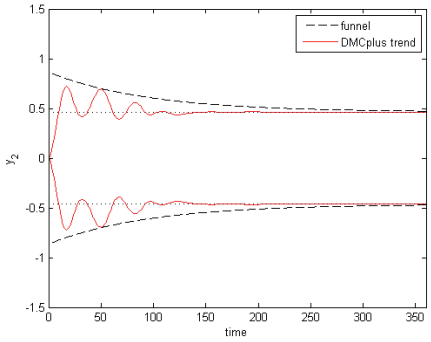


Fig. 6. Funnel Design and DMCplus trend for output y_2 with $(\alpha_f, \beta_f) = (0, -0.74)$.

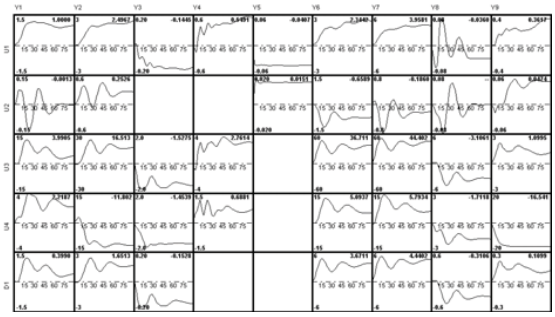


Fig. 7. Step Response Model for the SMR Problem. Responses for outputs y_1- y_9 to a step in inputs u_1 - u_4 and disturbance d_1. Empty boxes represent that there is no interaction between the input-output or disturbance-output pair.

Table 1. SMR Example: original and designed set of output constraints at the steady-state (Lima, Georgakis, Smith, Vinson and Schnelle, 2009).

Process Outputs	Original Lower Bound	Original Upper Bound	Designed Lower Bound	Designed Upper Bound
y_1	43.00	45.70	43.98	44.72
y_2	26.90	161.30	84.89	103.31
y_3	0.80	2.20	1.02	1.98
y_4	0	43.00	20.91	22.09
y_5	1.70	1.90	1.73	1.87
y_6	424.70	438.20	426.82	436.08
y_7	430.10	591.40	455.48	566.02
y_8	3.20	7.50	3.87	6.83
y_9	21.50	52.70	26.41	47.79

For $d_1 = 4$:

$$
\begin{aligned}
f_1 &= (1+1)1.60\exp\left[-(1+0.36)1.51\times10^{-2}t\right]+44.72\\
f_2 &= (1+1)6.60\exp\left[-(1+0.36)2.86\times10^{-2}t\right]+103.31\\
f_3 &= (1+1)(-0.60)\exp\left[-(1+0.36)1.40\times10^{-2}t\right]+1.02\\
f_4 &= 22.09\\
f_5 &= 1.87\\
f_6 &= (1+1)14.68\exp\left[-(1+0.36)1.95\times10^{-2}t\right]+436.08\\
f_7 &= (1+1)17.76\exp\left[-(1+0.36)2.21\times10^{-2}t\right]+566.02\\
f_8 &= (1+1)(-1.24)\exp\left[-(1+0.36)2.48\times10^{-2}t\right]+3.87\\
f_9 &= (1+1)0.44\exp\left[-(1+0.36)1.10\times10^{-2}t\right]+47.79
\end{aligned}
\tag{20}
$$

For $d_1 = -4$:

$$
\begin{aligned}
f_1 &= (1+1)(-1.60)\exp\left[-(1+0.36)1.51\times10^{-2}t\right]+43.98\\
f_2 &= (1+1)(-6.60)\exp\left[-(1+0.36)2.86\times10^{-2}t\right]+84.89\\
f_3 &= (1+1)0.60\exp\left[-(1+0.36)1.40\times10^{-2}t\right]+1.98\\
f_4 &= 20.91\\
f_5 &= 1.73\\
f_6 &= (1+1)(-14.68)\exp\left[-(1+0.36)1.95\times10^{-2}t\right]+426.82\\
f_7 &= (1+1)(-17.76)\exp\left[-(1+0.36)2.21\times10^{-2}t\right]+455.48\\
f_8 &= (1+1)1.24\exp\left[-(1+0.36)2.48\times10^{-2}t\right]+6.83\\
f_9 &= (1+1)(-0.44)\exp\left[-(1+0.36)1.10\times10^{-2}t\right]+26.41
\end{aligned}
\tag{21}
$$

Note that, for outputs y_4 and y_5, the steady-state disturbance gains are 0, and thus, their funnel's upper and lower bounds are constants at their upper and lower designed steady-state limits, respectively. Figs. 8, 9, and 10 show the DMCplus trends for y_1, y_2, and y_3, respectively, as well as the funnels for each case, where the controller is operating in closed-loop mode and the disturbance was inserted at time = 0. The funnels for the other outputs are not shown here due to space limitations of the manuscript.

Table 2. Calculated λs for each input-output and disturbance-output pairs for the SMR Example (smallest λ for each of the outputs in bold; dash (-) for pairs with no model)

y/ u, d	u_1	u_2	u_3	u_4	d_1
y_1	0.0189	0.0323	0.0212	**0.0151**	0.0212
y_2	0.0335	0.0699	0.0288	**0.0286**	0.0288
y_3	0.0794	-	**0.0140**	0.0847	**0.0140**
y_4	0.0537	-	0.0939	**0.0364**	-
y_5	1.4437	**0.5177**	-	-	-
y_6	0.0407	0.0453	0.0422	**0.0195**	0.0422
y_7	0.0530	0.0391	0.0681	**0.0221**	0.0681
y_8	**0.0248**	0.0280	0.0316	0.0706	0.0316
y_9	0.0243	0.0290	**0.0110**	0.1213	**0.0110**

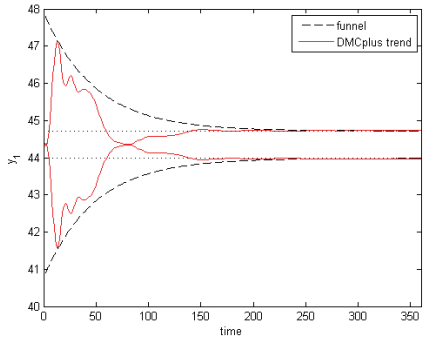


Figure 8: SMR Example: funnel design and DMCplus trend for output y_1 with $(\alpha_f, \beta_f) = (1.00, 0.36)$.

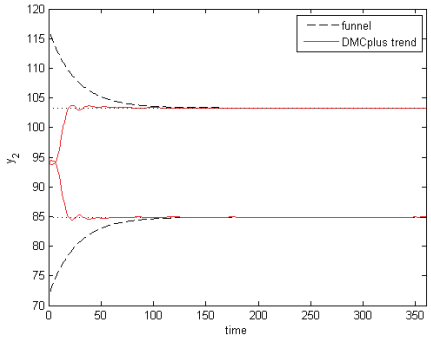


Figure 9: SMR Example: funnel design and DMCplus trend for output y_2 with $(\alpha_f, \beta_f) = (1.00, 0.36)$.

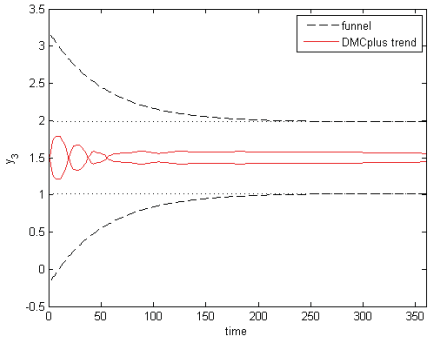


Figure 10: SMR Example: funnel design and DMCplus trend for output y_3 with $(\alpha_f, \beta_f) = (1.00, 0.36)$.

6. CONCLUSIONS

In this paper we have presented an extension of the previously developed steady-state interval operability approach to dynamical systems. Through the detailed examination of an illustrative case study we have motivated the calculation of output funnels for the design of output constraints during transient operation. The developed methodology was then applied to determine feasible output constraints for the Steam Methane Reformer industrial process. The analysis presented here provides a starting point for the verification of the achievability of control objectives in the entire control horizon. As potential future directions, an extension of this framework to address systems with multiple disturbances is necessary. Moreover, a moving horizon operability approach could be developed, where operability calculations would be performed online at each time instant, as the control horizon advances.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support from PRF-ACS through grant # 45400-AC9. We also wish to acknowledge David R. Vinson (APCI) for providing the process model and William M. Canney (AspenTech) for providing the DMCplus™ software.

REFERENCES

Dimitriadis, V.D. and Pistikopoulos, E.N. (1995). Flexibility analysis of dynamic systems. *Ind. Eng. Chem. Res*., 34, 4451-4462.

Lima, F.V. and Georgakis, C. (2006). Operability of multivariable non-square systems. *In Proceedings of the 2006 IFAC International Symposium on Advanced Control of Chemical Processes (ADCHEM)*, 989-994.

Lima, F.V. and Georgakis, C. (2008a). Design of output constraints for model-based non-square controllers using interval operability. *J. Process Control*, 18, 610-620.

Lima, F.V. and Georgakis, C. (2008b). Input-output operability of control systems. *J. Process Control, submitted for publication*.

Lima, F.V., Georgakis, C., Smith, J.F., Vinson, D.R. and Schnelle, P.D. (2009). Operability-based determination of feasible constraints for the control of several high-dimensional non-square industrial processes. *AIChE Journal*, *accepted for publication.*

Lima, F.V., Georgakis, C., Smith, J.F. and Schnelle, P.D. (2008). Analysis of the constraint characteristics of a sheet forming control problem using interval operability concepts. *Computer Aided Chemical Engineering*, 25, 387-392.

Ljung, L. (1999). *System identification: Theory for the user.* 2nd ed., Prentice Hall, Upper Saddle River, NJ.

Maciejowski, J.M. (2002). *Predictive control with constraints.* Prentice Hall, Englewood Cliffs, NJ.

Qin, S.J. and Badgwell, T.A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice,* 11, 733-764.

Rawlings, J.B. (2000). Tutorial overview of model predictive control. *IEEE Control Syst. Magazine*, 20, 38-52.

Vinson, D. R. (2000). *A new measure of process operability for the improved steady-state design of chemical processes*. Ph.D. Thesis, Lehigh University, USA.

Vinson, D.R. and Georgakis, C. (2000). A new measure of process output controllability. *J. Process Control*, 10, 185-194.

Computation of the Infinite Horizon Continuous Time Constrained Linear Quadratic Regulator ⋆

Gabriele Pannocchia * **James B. Rawlings** ** **David Q. Mayne** *** **Wolfgang Marquardt** ****

* *Dept. Chem. Eng., Ind. Chem. & Sc. Mat. – Univ. of Pisa, Italy (email: g.pannocchia@ing.unipi.it)*
** *Dept. Chem. & Biol. Eng. – Univ. of Wisconsin. Madison (WI), USA (emai: rawlings@eng.wisc.edu)*
*** *Dept. Elec. & Electronic Eng. – Imperial College London, UK (email: d.mayne@imperial.ac.uk)*
**** *AVT–Process Systems Engineering, RWTH Aachen University, Germany (email: Wolfgang.marquardt@avt.rwth-aachen.de)*

Abstract: We present a method for computing the solution to the infinite horizon continuous-time constrained linear quadratic regulator (CLQR). The method relies on two main features: a multi-grid method for placing a finite number of time intervals, and a piece-wise linear parameterization of the input within the intervals. The input values at the grid points and slopes within the time intervals are computed via quadratic programs (QPs). The grids are gradually refined to efficiently improve the accuracy of the solution, and the required matrices and vectors for all QPs are computed offline and stored to improve the online efficiency. We present two examples, a single-input single-output unstable system and a three-input three-output stable system, to show the main characteristics of the proposed computation method.

Keywords: Constrained Linear Quadratic Regulation, Continuous Time Systems, Model Predictive Control, Optimal Control

1. MOTIVATIONS FOR CONTINUOUS TIME MODEL PREDICTIVE CONTROL

The scope of applications of model predictive control (MPC) has expanded well beyond its original starting point in the process industries. With this increased scope has come the need to evaluate in real time the solution to the MPC optimal control problem for systems with fast open-loop dynamics. It is reasonable to anticipate that this trend to faster applications may culminate with a return to the continuous time description of the system model. The previous widespread adoption of discrete time models to represent the system dynamics made perfect sense. The typical sample time in earlier applications was small compared to the closed-loop dynamic response of the system (seconds compared to minutes) so there was essentially no loss in model accuracy. Moreover, the earlier analysis of the closed-loop properties and computational strategies to approximate infinite horizon control laws was simplified using discrete time models (Mayne et al., 2000).

In today's application environment, it is no longer safe to assume that some fixed sample rate can be chosen very small compared to the closed-loop system dynamics. Next, given the rapid development of MPC theory for discrete time models over the last 15 years, there is no real difficulty in establishing properties of interest in a continuous time setting. Finally, the actuator hardware has become "smarter" and it is now becoming appropriate to assume that any reasonable time function may be sent to the actuator, and it is actuator hardware's job to accurately track this signal. If the application has fast dynamics, obviously a requirement of the process design is the selection of sensors and actuators that are fast enough to keep up.

In this paper we would like to remove all issues of sampling and address directly the MPC problem for the continuous-time model (Yuz et al., 2005). The job of the MPC controller in this context is to send its solution as a time signal to the actuators until a measurement becomes available and a new state initial condition is available to the controller. This context has become popular in the nonlinear MPC area, where nonlinear models from physical principles are almost always continuous time nonlinear differential equations [see (Diehl et al., 2008) and references therein]. In this paper we would like to explore what efficiencies can be gained when we restrict attention to *linear* continuous time models.

2. PROBLEM DEFINITION

We consider linear time-invariant continuous-time systems

$$\dot{x} = Ax + Bu\,, \tag{1}$$

in which $x \in \mathbb{R}^n$ is the state and $u \in \mathbb{R}^m$ is the input. We define the following cost function for a given initial state $x_0 \in \mathbb{R}^n$ and infinite-time input $\mathbf{u}$:

$$V(x_0, \mathbf{u}) = \frac{1}{2}\int_0^\infty (x'Qx + u'Ru)\,dt, \quad \text{s.t. (1) and } x(0) = x_0\,. \tag{2}$$

in which we use the following notation. Given a function $u : \overline{\mathbb{R}} \to \mathbb{R}^m$, we define $\mathbf{u} = \{u(t) | t \geq 0\}$. The aim of this work is to compute, given the current initial state x_0, the solution of the

⋆ This research was supported by National Science Foundation (Grant CTS-0456694).

following infinite horizon optimal constrained linear quadratic regulation (CLQR) problem

$$\mathbf{u}^*(x_0) = \underset{\mathbf{u}}{\operatorname{argmin}}\, V(x_0, \mathbf{u}), \quad \text{s.t. } Du(t) \leq d\,, \qquad (3)$$

in which $D = \begin{bmatrix} D_1' & D_2' \end{bmatrix}'$ and $d = \begin{bmatrix} d_1' & d_2' \end{bmatrix}'$ with $d_1 > 0$ and $d_2 = 0$ (in element-wise sense). Notice that we allow the possibility of either (D_1, d_1) or (D_2, d_2) being empty.

We make the following assumptions.

Assumption 1. Given a matrix $T \in \mathbb{R}^{m \times r}$ with rank r such that $D_2 T = 0$, the pair (A, BT) is stabilizable, R is positive definite, Q is positive semi-definite and (A, Q) is detectable.

It is important to point out that the constraint $D_2 u \leq 0$ is active at the equilibrium point ($u = 0$). Assumption 1 states that the system must be stabilizable under the restricted control $D_2 u = 0$. We can write the restricted optimal control as $u = \bar{K}x = -(T'RT)^{-1}(BT)\bar{P}x$, with $\bar{P}$ solution of the continuous-time Riccati equation for the system matrices (A, BT) with penalties $(Q, T'RT)$ [see (Rao and Rawlings, 1999; Pannocchia et al., 2003) for further details]. Such control is feasible for all x in the positively invariant set $\mathcal{X} = \{x | D_1 \bar{K} e^{(A+B\bar{K})t} x \leq d_1 \text{ for all } t\}$.

Compared to the discrete-time counterpart, the continuous-time CLQR problem has received much less attention, although several results are available. Cannon and Kouvaritakis (2000) proposed a method for single-input single-output systems, using basis functions, in which input constraint satisfaction is ensured by a backoff strategy. Kojima and Morari (2004) propose a design method that is based on the singular value decomposition (SVD) of the finite horizon linear system and that guarantees in the limit constraint satisfaction and convergence to the optimal solution. For low dimensional linear systems, Sakizlis et al. (2005) present an approach for computing the explicit solution to the finite horizon continuous-time CLQR problem, by merging variational analysis with parametric optimization tools. Goebel and Subbotin (2007) present an approach based on the solution of the backward Hamiltonian system; optimal trajectories are stored for subsequent on-line suboptimal evaluation.

In this work, we propose a novel approach based on the direct evaluation of a suboptimal solution to the infinite horizon CLQR problem that requires, like the discrete-time case, the solution of only quadratic programs (QPs). Thus, we compute the control input in a number of grid points that can be efficiently adapted on-line to improve the accuracy of the solution. Two appropriate continuous-time input parameterizations (holds) are proposed, which show markedly improved convergence towards the optimal continuous-time solution compared to the usual piece-wise constant parameterization, and still guarantees satisfaction of the input constraints. Other related details, including proofs of the results, can be found in (Pannocchia et al., 2009).

3. INPUT PARAMETERIZATION, EXACT COST EVALUATION AND CONVERGENCE ANALYSIS

We consider three different input parameterization methods (also referred to as holds), and for each hold type a discrete-time system realization is obtained in a way that the continuous-time and the discrete-time state and input match at given grid points $(t_0, t_1, \ldots)$ assumed, in general, not evenly spaced. Moreover,

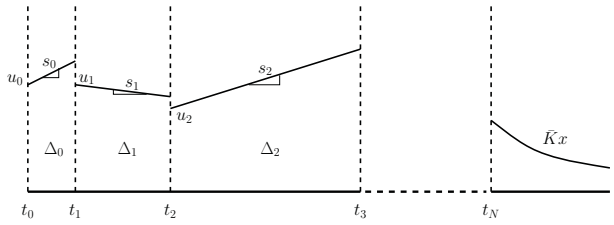


Fig. 1. Piece-wise linear input parameterization and time gridding scheme.

for each hold type we compute the corresponding discrete-time LQR cost matrices in a way that the cost in (2) is exactly evaluated. We recall that the solution to (1) is given by:

$$x(t) = e^{At}x_0 + \int_0^t e^{A(t-\tau)} B u(\tau) d\tau\,. \qquad (4)$$

3.1 Three input parameterization methods

First we summarize the properties of the well known zero-order hold (ZOH), defined as:

$$u(t) = u_k, \qquad t_k \leq t < t_{k+1}\,, \qquad (5)$$

with $t_{k+1} = t_k + \Delta_k$. Notice that we do not assume, however, that Δ_k is constant. From (4) and (5), we obtain the time-varying discrete-time system: $x_{k+1} = A_k^0 x_k + B_k^0 u_k$, in which the matrices (A_k^0, B_k^0) are defined as: $A_k^0 = e^{A\Delta_k}$, $B_k^0 = I_0(\Delta_k)B$, with $I_0(t) = \int_0^t e^{A\tau} d\tau$.

We consider a second input parameterization referred to as "piece-wise linear" hold (PWLH), and defined as:

$$u(t) = u_k + s_k(t - t_k), \qquad t_k \leq t < t_{k+1}\,, \qquad (6)$$

in which $s \in \mathbb{R}^m$ defines the "slope" of u between grid points. In the sake of clarity, we sketch the PWLH input parameterization in Figure 1, in which we also emphasize the uneven time gridding scheme that we consider later in Section 4. From (4) and (6), the time-varying discrete-time system $x_{k+1} = A_k^{\mathrm{I}} x_k + B_k^{\mathrm{I}} u_k^{\mathrm{I}}$ is obtained, in which $u^{\mathrm{I}} \in \mathbb{R}^{2m} = \begin{bmatrix} u' & s' \end{bmatrix}'$ is the augmented input, and the matrices $(A_k^{\mathrm{I}}, B_k^{\mathrm{I}})$ are: $A_k^{\mathrm{I}} = A_k^0 = e^{A\Delta_k}$, $B_k^{\mathrm{I}} = [I_0(\Delta_k)B \quad I_1(\Delta_k)B]$, with $I_0(t) = \int_0^t e^{A\tau} d\tau$ and $I_1(t) = \int_0^t e^{A(t-\tau)} \tau d\tau$.

Finally, we consider a third input parameterization, which also assumes that the input varies linearly between discrete times, but with a slope equal to forward finite input difference, i.e.

$$u(t) = u_k + \left(\frac{u_{k+1} - u_k}{\Delta_k} \right)(t - t_k), \qquad t_k \leq t < t_{k+1}\,. \quad (7)$$

In the sequel, this input parameterization is referred to as "forward first-order" hold (FFOH). From (4) and (7), we obtain the time-varying discrete-time system: $x_{k+1}^{\mathrm{II}} = A_k^{\mathrm{II}} x_k^{\mathrm{II}} + B_k^{\mathrm{II}} u_k^{\mathrm{II}}$, in which the augmented state $x^{\mathrm{II}} \in \mathbb{R}^{n+m}$, the shifted input $u^{\mathrm{II}} \in \mathbb{R}^m$, and $(A_k^{\mathrm{II}}, B_k^{\mathrm{II}})$ are:

$$x_k^{\mathrm{II}} = \begin{bmatrix} x_k \\ u_k \end{bmatrix}, \quad u_k^{\mathrm{II}} = u_{k+1},$$

$$A_k^{\mathrm{II}} = \begin{bmatrix} A_k^0 & I_0(\Delta_k)B - \dfrac{I_1(\Delta_k)}{\Delta_k} B \\ 0 & 0 \end{bmatrix}, \quad B_k^{\mathrm{II}} = \begin{bmatrix} \dfrac{I_1(\Delta_k)}{\Delta_k} B \\ I \end{bmatrix}.$$

We emphasize that ZOH generates a $u(t)$ constant between discrete times and discontinuous at the discrete times, PWLH generates a $u(t)$ linear between discrete times and discontinuous at the discrete times, FFOH generates a $u(t)$ continuous at all times and linear between discrete times.

3.2 Exact continuous time cost computation

Lemma 2. (Exact cost matrices for ZOH). If the continuous-time input is given by (5), then

$$\int_{t_k}^{t_{k+1}} (x'Qx + u'Ru)dt = x_k'Q_k^0x_k + u_k'R_k^0u_k + 2x_k'M_k^0u_k \ ,$$

with: $Q_k^0 = \int_0^{\Delta_k}(e^{At})'Qe^{At}dt$, $R_k^0 = \int_0^{\Delta_k}\left(R + (I_0B)'QI_0B\right)dt$, $M_k^0 = \int_0^{\Delta_k}(e^{At})'QI_0Bdt$.

It is interesting to notice that this quadrature result is known (Kwakernaak and Sivan, 1972, p.549), (Yuz et al., 2005, Sec.2.2), but most of the literature on optimal control of continuous-time systems usually ignores the mixed state-input cost term and assumes $Q_k^0 = Q\Delta_k$, $R_k^0 = R\Delta_k$, thus introducing an inherent quadrature error.

From the previous result it immediately follows that, given an infinite discrete-time input sequence $(u_0, u_1, u_2, \ldots)$, assuming that the continuous-time input $\mathbf{u}$ is defined in (5), then

$$V(x_0, \mathbf{u}) = \frac{1}{2}\sum_{k=0}^{\infty} x_k'Q_k^0x_k + u_k'R_k^0u_k + 2x_k'M_k^0u_k \ .$$

Lemma 3. (Exact cost matrices for PWLH). If the continuous-time input is given by (6), then $\int_{t_k}^{t_{k+1}}(x'Qx + u'Ru)dt = x_k'Q_k^{\mathrm{I}}x_k + (u_k^{\mathrm{I}})'R_k^{\mathrm{I}}u_k^{\mathrm{I}} + 2x_k'M_k^{\mathrm{I}}u_k^{\mathrm{I}}$, with:

$Q_k^{\mathrm{I}} = \int_0^{\Delta_k}(e^{At})'Qe^{At}dt$,

$R_k^{\mathrm{I}} = \int_0^{\Delta_k}\begin{bmatrix} R + (I_0B)'Q(I_0B) & (I_0B)'Q(I_1B) + Rt \\ (I_1B)'Q(I_0B) + Rt & (I_1B)'Q(I_1B) + Rt^2 \end{bmatrix}dt$,

$M_k^{\mathrm{I}} = \int_0^{\Delta_k}\begin{bmatrix}(e^{At})'Q(I_0B) & (e^{At})'Q(I_1B)\end{bmatrix}dt$.

From this results it follows that, given infinite discrete-time input sequence and slope sequence $(u_0, u_1, u_2, \ldots)$, $(s_0, s_1, s_2, \ldots)$, assuming that the continuous-time input $\mathbf{u}$ is defined in (6), then

$$V(x_0, \mathbf{u}) = \frac{1}{2}\sum_{k=0}^{\infty} x_k'Q_k^{\mathrm{I}}x_k + (u_k^{\mathrm{I}})'R_k^{\mathrm{I}}u_k^{\mathrm{I}} + 2x_k'M_k^{\mathrm{I}}u_k^{\mathrm{I}} \ .$$

Lemma 4. (Exact cost matrices for FFOH). If the continuous-time input is given by (7), then $\int_{t_k}^{t_{k+1}}(x'Qx + u'Ru)dt = (x_k^{\mathrm{II}})'Q_k^{\mathrm{II}}x_k^{\mathrm{II}} + (u_k^{\mathrm{II}})'R_k^{\mathrm{II}}u_k^{\mathrm{II}} + 2(x_k^{\mathrm{II}})'M_k^{\mathrm{II}}u_k^{\mathrm{II}}$, with:

$Q_k^{\mathrm{II}} = \int_0^{\Delta_k}\begin{bmatrix} (e^{At})'Qe^{At} & (e^{At})'Q(I_0 - \frac{I_1}{\Delta_k})B \\ ((I_0 - \frac{I_1}{\Delta_k})B)'Qe^{At} & ((I_0 - \frac{I_1}{\Delta_k})B)'Q(I_0 - \frac{I_1}{\Delta_k})B + R\left(1 - \frac{t}{\Delta_k}\right)^2\end{bmatrix}dt$,

$R_k^{\mathrm{II}} = \int_0^{\Delta_k}\left(R\left(\frac{t}{\Delta_k}\right)^2 + \left(\frac{I_1}{\Delta_k}B\right)'Q\left(\frac{I_1}{\Delta_k}B\right)\right)dt$,

$M_k^{\mathrm{II}} = \int_0^{\Delta_k}\begin{bmatrix} (e^{At})'Q\frac{I_1}{\Delta_k}B \\ ((I_0 - \frac{I_1}{\Delta_k})B)'Q\frac{I_1}{\Delta_k}B + R\frac{t}{\Delta_k}\left(1 - \frac{t}{\Delta_k}\right)\end{bmatrix}dt$.

Clearly, it follows that given an infinite discrete-time input sequence $(u_0, u_1, u_2, \ldots)$, assuming that the continuous-time input $\mathbf{u}$ is defined in (7), then

$$V(x_0, \mathbf{u}) = \frac{1}{2}\sum_{k=0}^{\infty}(x_k^{\mathrm{II}})'Q_k^{\mathrm{II}}x_k^{\mathrm{II}} + (u_k^{\mathrm{II}})'R_k^{\mathrm{II}}u_k^{\mathrm{II}} + 2(x_k^{\mathrm{II}})'M_k^{\mathrm{II}}u_k^{\mathrm{II}} \ .$$

3.3 Unconstrained convergence analysis for the three holds

In this section, we evaluate the unconstrained optimal cost that is achieved by using the three different input parameterizations, in the case of evenly spaced points, i.e. $\Delta_0 = \Delta_1 = \cdots = \Delta$, and we compare the order of convergence to the optimal continuous-time cost as Δ goes to zero. [1] We first recall the following well-known results for unconstrained LQR problems.

Lemma 5. The optimal cost-function value for the unconstrained continuous-time LQR problem $\min_{\mathbf{u}} V(x_0, \mathbf{u})$ is given by $\frac{1}{2}x_0'Px_0$ in which P is the positive semi-definite solution of the Riccati equation:

$$0 = Q + A'P + PA - PBR^{-1}B'P \ . \tag{8}$$

Lemma 6. The following discrete-time LQR problem with mixed state-input terms:

$$\min_{(u_0,u_1,\ldots)} \frac{1}{2}\sum_{k=0}^{\infty} x_k'\bar{Q}x_k + u_k'\bar{R}u_k + 2x_k'\bar{M}u_k \ , \quad \text{s.t.}$$
$$x_{k+1} = \bar{A}x_k + \bar{B}u_k \ ,$$

is equivalent to the following discrete-time LQR problem without mixed state-input terms:

$$\min_{(u_0,u_1,\ldots)} \frac{1}{2}\sum_{k=0}^{\infty} x_k'\bar{Q}x_k + u_k'\bar{R}u_k, \quad \text{s.t.}$$
$$x_{k+1} = \bar{A}x_k + \bar{B}u_k \ , \tag{9}$$

in which the following change of variables is considered: $u_k \leftarrow u_k - \bar{R}^{-1}M'x_k$, $\bar{A} \leftarrow \bar{A} - \bar{B}\bar{R}^{-1}\bar{M}'$, $\bar{Q} \leftarrow \bar{Q} - \bar{M}\bar{R}^{-1}\bar{M}'$.

Lemma 7. The optimal cost-function value for the unconstrained discrete-time LQR problem (9) is $\frac{1}{2}x_0'\Pi x_0$, in which Π is the positive semi-definite solution of the Riccati equation:

$$0 = -\Pi + \bar{Q} + \bar{A}'\Pi\bar{A} - \bar{A}'\Pi\bar{B}(\bar{R} + \bar{B}'\Pi\bar{B})^{-1}\bar{B}\Pi\bar{A} \ . \tag{10}$$

If P is the solution of the continuous-time Riccati equation (8) and $\Pi(\Delta)$ is the solution of the discrete-time Riccati equation (10) using a given hold and a fixed discrete-time interval Δ, it is straightforward to show that $P \preccurlyeq \Pi(\Delta)$, which is equivalent to saying $\Pi(\Delta) - P$ is positive semidefinite. Clearly, it is desirable for an input parameterization to have convergence $\Pi(\Delta) \to P$ as $\Delta \to 0$. We define the order of convergence ℓ for a given hold implementation as the smallest non-negative integer for which $\lim_{\Delta\to 0}\frac{\Pi(\Delta)-P}{\Delta^\ell} \neq 0$.

We next establish the following results about the convergence order the LQR cost using different holds [2] . Notice that each hold defines a discrete-time optimal control problem, in which the decision variables are $\{u_k\}_{k=0}^{\infty}$ for ZOH and FFOH, and are $\{(u_k, s_k)\}_{k=0}^{\infty}$ for PWLH.

Theorem 8. (Second order convergence of ZOH). The convergence of $\Pi^0(\Delta)$ to P is second order for system matrices $\bar{A} = A^0 - B^0(R^0)^{-1}(M^0)'$, $\bar{B} = B^0$ and cost matrices $\bar{Q} = Q^0 - M^0(R^0)^{-1}(M^0)'$, $\bar{R} = R^0$.

It is interesting to notice that if inexact cost matrices are used in ZOH, the order of convergence is less than 2. For instance, if one simply chooses $Q^0 = Q\Delta$, $R^0 = R\Delta$, $M^0 = 0$, the convergence order is 1.

Theorem 9. (Fourth order convergence of PWLH). The convergence of $\Pi^{\mathrm{I}}(\Delta)$ to P is fourth order for system matrices $\bar{A} = A^{\mathrm{I}} - B^{\mathrm{I}}(R^{\mathrm{I}})^{-1}(M^{\mathrm{I}})'$, $\bar{B} = B^{\mathrm{I}}$ and cost matrices $\bar{Q} = Q^{\mathrm{I}} - M^{\mathrm{I}}(R^{\mathrm{I}})^{-1}(M^{\mathrm{I}})'$, $\bar{R} = R^{\mathrm{I}}$.

Before presenting the convergence result for FFOH, it is important to recall that, in system (3.1), the state is aug-

[1] The proof for ZOH is reported in (Pannocchia et al., 2009), and follows Taylor expansions of all terms in the discrete algebraic Riccati equation. For PWLH and FFOH, symbolic manipulation software may be useful.

[2] Since the time interval is fixed, all the discrete-time matrices are time-invariant. Thus, we drop the subscript k in this section.

mented. Thus, the corresponding solution of (10) is in the form $\Pi = \begin{bmatrix} \Pi_{xx} & \Pi_{xu} \\ \Pi'_{xu} & \Pi_{uu} \end{bmatrix}$, and, hence, the unconstrained cost for any given initial state x_0 is given by $\frac{1}{2}(x_0'\Pi_{xx}x_0 + 2x_0'\Pi_{xu}u_0 + u_0'\Pi_{uu}u_0)$. Since the input u_0 is a decision variable, the optimal unconstrained cost for any given initial state x_0 using FFOH is easily obtained as $\frac{1}{2}x_0'(\Pi_{xx} - \Pi_{xu}\Pi_{uu}^{-1}\Pi'_{xu})x_0$.

Theorem 10. (Fourth order convergence of FFOH). The convergence of $\Pi^{\rm II}(\Delta) = \Pi_{xx} - \Pi_{xu}\Pi_{uu}^{-1}\Pi'_{xu}$ to P is fourth order for system matrices $\bar{A} = A^{\rm II} - B^{\rm II}(R^{\rm II})^{-1}(M^{\rm II})'$, $\bar{B} = B^{\rm II}$ and cost matrices $\bar{Q} = Q^{\rm II} - M^{\rm II}(R^{\rm II})^{-1}(M^{\rm II})'$, $\bar{R} = R^{\rm II}$.

We now show how the optimal discrete-time cost matrices for different holds are "ordered".

Theorem 11. (Cost comparison). The following linear matrix inequalities hold: $\Pi^{\rm I}(\Delta) \preccurlyeq \Pi^{0}(\Delta)$, $\Pi^{\rm I}(\Delta) \preccurlyeq \Pi^{\rm II}(\Delta)$.

The reader may naturally expect the following ordering also to hold, $\Pi^{\rm II}(\Delta) \preccurlyeq \Pi^{0}(\Delta)$, but this is not valid for arbitrary Δ. The discontinuities allowed in ZOH may provide better performance than the continuous FFOH for large Δ. Of course, due to the different convergence orders, the ordering does hold for sufficiently small Δ.

4. ALGORITHM FOR COMPUTATION OF THE CONTINUOUS-TIME CLQR

4.1 Introduction and main definitions

Motivated by the nice convergence results of the PWLH and FFOH input parameterizations, we propose computing a suboptimal solution to problem (3) in terms of an appropriate finite number of decision variables, namely the inputs $(u_0, \ldots, u_{N-1})$, and for the PWLH case also the slopes $(s_0, \ldots, s_{N-1})$. As shown in this section, we can write the infinite-horizon continuous-time CLQR problem as a finite dimensional Quadratic Program, whose complexity is the same as that of the discrete-time CLQR problem. Furthermore, we define a procedure for placing the grid points $(t_0, \ldots, t_N)$ in a way that the number of decision variables is kept small while the accuracy of the solution is improved.

Let $(t_0, \ldots, t_N)$ be a sequence of $N+1$ grid points with $t_0 = 0$, and consider the following suboptimal CLQR problems:

$$\mathbf{u}^{\rm I} = \underset{\mathbf{u}}{\operatorname{argmin}}\ V(x_0, \mathbf{u}) \quad \text{s.t. } Du(t) \le d, \text{ (6), and } D_2u(t) = 0, \quad \text{for } t \ge t_N\,, \quad (11)$$

$$\mathbf{u}^{\rm II} = \underset{\mathbf{u}}{\operatorname{argmin}}\ V(x_0, \mathbf{u}) \quad \text{s.t. } Du(t) \le d, \text{ (7), and } D_2u(t) = 0, \quad \text{for } t \ge t_N\,, \quad (12)$$

From what presented so far, we can rewrite the problem (11) as follows:

$$\min_{(u_0^{\rm I}, u_1^{\rm I}, \ldots, u_{N-1}^{\rm I})} \frac{1}{2}x_N'P^{\rm I}x_N + \frac{1}{2}\sum_{k=0}^{N-1} x_k'Q_k^{\rm I}x_k + (u_k^{\rm I})'R_k^{\rm I}u_k^{\rm I} + 2x_k'M_k^{\rm I}u_k^{\rm I}, \quad \text{s.t.} \quad (13a)$$

$$x_{k+1} = A_k^{\rm I}x_k + B_k^{\rm I}u_k^{\rm I},\ x_0 = x(0),\ \begin{bmatrix} D & 0 \\ D & D\Delta_k \end{bmatrix} u_k^{\rm I} \le \begin{bmatrix} d \\ d \end{bmatrix}\,, \quad (13b)$$

in which $P^{\rm I} = \bar{P}$, provided that $x_N \in \mathcal{X}$.

Similarly, the problem (12) can be rewritten as follows:

$$\min_{(x_0^{\rm II}, u_0^{\rm II}, u_1^{\rm II}, \ldots, u_{N-1}^{\rm II})} \frac{1}{2}(x_N^{\rm II})'P^{\rm II}x_N^{\rm II} + \frac{1}{2}\sum_{k=0}^{N-1} (x_k^{\rm II})'Q_k^{\rm II}x_k^{\rm II} + (u_k^{\rm II})'R_k^{\rm II}u_k^{\rm II} + 2(x_k^{\rm II})'M_k^{\rm II}u_k^{\rm II}, \quad \text{s.t.} \quad (14a)$$

$$x_{k+1}^{\rm II} = A_k^{\rm II}x_k^{\rm II} + B_k^{\rm II}u_k^{\rm II},\ \mathcal{I}x_0^{\rm II} = x(0),\quad D[0,\ I]x_0^{\rm II} \le d, \quad Du_k^{\rm II} \le d\,, \quad (14b)$$

in which $P^{\rm II} = \mathcal{I}'\bar{P}\mathcal{I}$ and $\mathcal{I} = [I,\ 0]$, provided that $\mathcal{I}x_N^{\rm II} \in \mathcal{X}$.

After elimination of the state variables using (13b), problem (13) can be written as a convex QP in $2mN$ decision variables. Similarly, using (14b), the problem (14) can be written as a convex QP in $m(N+1)$ decision variables. Notice that $x_0^{\rm II} = [x_0',\ u_0']'$ is a decision variable in (14), but the initial constraint in (14b) implies that only u_0 is a free variable.

Notice that if $V^{\rm I}(x_0) = V(x_0, \mathbf{u}^{\rm I})$ and $V^{\rm II}(x_0) = V(x_0, \mathbf{u}^{\rm II})$ denote the optimal cost-function values for the problems (11) and (12), respectively, obtained with the same grid points $(t_0, \ldots, t_N)$, then it follows that $V^{\rm II}(x_0) \ge V^{\rm I}(x_0) \ge V^*(x_0)$.

4.2 Offline and online computations

The degree of suboptimality of the solution depends on the number of grid points, where it is clear that a larger number of intervals would result in a more accurate solution. However, it is important to observe that, in order to improve the solution, more grid points need to be placed where input and states are far away from the origin, whereas when they approach the origin, many intervals are unnecessary. During this work we tested several strategies for deciding the number and size of the intervals, with two desired goals in mind: (i) at each iteration the solution accuracy is improved, i.e. the computed objective function is decreased; (ii) the total number of intervals, and hence the number of QP decision variables, is kept small. We next present the simplest algorithm that is used offline to generate the QP associated with (13).

Algorithm 1. Data: maximum time t_N, initial number of intervals Θ, number of halving loops Θ_s.

(1) Compute the Θ intervals $(\Delta_0, \ldots, \Delta_{\Theta-1})$ such that $\Delta_k/\Delta_{k+1} = 0.5$ and $\sum_{k=0}^{\Theta-1}\Delta_k = t_N$.
(2) Define the initial sequence of $N_1 = \Theta$ intervals as $\mathcal{P}_1 = (\Delta_0^1, \ldots, \Delta_{N_1-1}^1) = (\Delta_0, \ldots, \Delta_{\Theta-1})$.
(3) Set $j \leftarrow j+1$ and define the next sequence $\mathcal{P}_j$ of $N_j = 2N_{j-1}$ intervals by halving each interval of $\mathcal{P}_{j-1}$.
(4) If $j < \Theta_s + 1$ go to 3. Otherwise, for each interval sequence $\mathcal{P}_j$ with $j = 1, \ldots, \Theta_s+1$, compute the matrices $(A_k^{\rm I}, B_k^{\rm I}, Q_k^{\rm I}, R_K^{\rm I}, M_k^{\rm I})$ in (13), build the associated QP in the form:

$$\mathbb{QP}_j: \qquad \min_{\mathbf{v}} \frac{1}{2}\mathbf{v}'\mathbf{H}\mathbf{v} + \mathbf{v}'\mathbf{Q}x_0, \quad \text{s.t. } \mathbf{A}\mathbf{v} \le \mathbf{b} \quad (15)$$

with $\mathbf{v} = (u_0^{\rm I}, \ldots, u_{N_j-1}^{\rm I})$, storing $\mathbf{H}$, $\mathbf{Q}$, $\mathbf{A}$, $\mathbf{b}$.

For open-loop unstable systems, in order to avoid ill-conditioning of $\mathbf{H}$, the input variable re-parameterization discussed in (Rossiter et al., 1998) is recommended.

The next algorithm describes the operations that are performed online to compute $\mathbf{u}^I(x_0)$.

Algorithm 2. Given x_0, the relative cost decrease tolerance $\mu > 0$. Initialize $j = 1$.

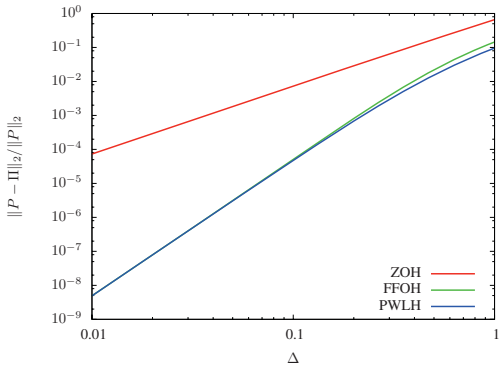


Fig. 2. System 1: Relative error between P and discrete-time cost Π for ZOH, FFOH, PWLH *vs* Δ.

(1) Solve $\mathbb{QP}_j$, let $V_j^{\mathrm{I}}(x_0)$ be the associated optimal cost and $\mathbf{u}_j^{\mathrm{I}}(x_0)$ the optimal input.
(2) If $j = 1$, set $j \leftarrow j + 1$ and go to 1. Otherwise,
(3) If the relative cost decrease satisfies $\frac{V_{j-1}^{\mathrm{I}}(x_0) - V_j^{\mathrm{I}}(x_0)}{V_{j-1}^{\mathrm{I}}(x_0)} < \mu$ or $j = \Theta_s + 1$, set $\mathbf{u}^{\mathrm{I}}(x_0) = \mathbf{u}_j^{\mathrm{I}}(x_0)$ and stop. Otherwise, set $j \leftarrow j + 1$ and go to 1.

Notice that if at any iteration j, the solution to $\mathbb{QP}_j$ is such that $x_N \notin \mathcal{X}$, then an additional interval of the largest size is added, i.e. $t_N \leftarrow t_N + \Delta_{N-1}$ and $\mathbb{QP}_j$ is solved again $x_N \notin \mathcal{X}$. Such additional intervals are retained also for subsequent iterations. For efficient online computation it is advised that the matrices associated with each $\mathbb{QP}_j$ are built and stored for different increasing t_N, so that the online CPU time is required only for solving the QPs.

We have the following important result.

Theorem 12. For each iteration $j > 1$ of Algorithm 2, we have that:

$$V_j^{\mathrm{I}}(x_0) \leq V_{j-1}^{\mathrm{I}}(x_0), \quad \text{for all } j > 0\,. \tag{16}$$

It is important to remark that the same cost decrease property holds true if Algorithm 2 is applied to compute $\mathbf{u}^{\mathrm{II}}(x_0)$.

5. CASE STUDIES

To illustrate the main features of the proposed method, we present two examples. The first example is a SISO unstable system, whose transfer function is shown below:

$$g_1(s) = -\frac{6.512s + 1.628}{-2.4390s^2 + 3.9756s + 1}\,.$$

The second example is the 3 input, 3 output Shell Control Problem (Prett and Morari, 1987), for which we use a 10 state continuous-time model. For both examples we use $Q = I$, $R = 0.1I$.

In Figure 2 we show, for the first example, the relative error (evaluated using the 2-norm) between the continuous-time LQR cost matrix P and the corresponding discrete-time cost matrix Π obtained with ZOH, FFOH, PWLH as a function of Δ. As expected from Theorems 8, 9, 10, the orders of convergence for ZOH is 2, whereas it is 4 for PWLH and FFOH.

Unless otherwise specified, in the subsequent studies we consider input constraints $-1 \leq u \leq 1$, the optimal input $\mathbf{u}^{\mathrm{II}}(x_0)$ is computed with Algorithm 2 using FFOH input parameterization and using a relative tolerance of $\mu = 10^{-4}$. For

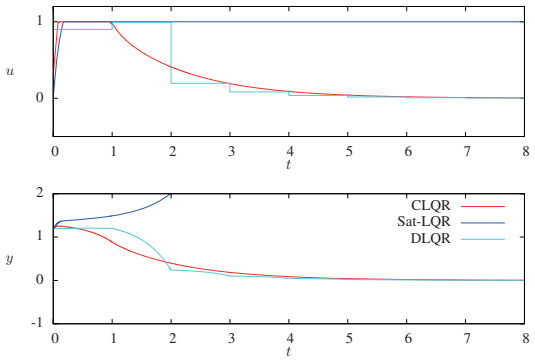


Fig. 3. System 1. Input and output closed-loop response using CLQR, "saturated" LQR and DLQR.

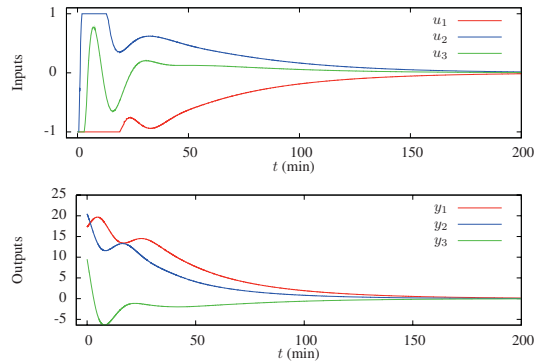


Fig. 4. System 2. Closed-loop inputs and outputs using CLQR at decision times $0, 1, \ldots$.

System 1 we show in the top plot of Figure 3 the optimal closed-loop input and output, computed by solving the optimal control problem with the proposed algorithm at decision times $0, 1, \ldots, 7$, and implementing the computed infinite horizon input in a receding horizon fashion. For comparison, we also show the results obtained by: (i) infinite horizon discrete-time constrained LQR (DLQR), (ii) "saturated" continuous-time LQR law $u = \mathrm{sat}\,(Kx)$. We observe that DLQR generates a stable closed-loop response that is fairly different from the optimal one obtained with CLQR. We also note that Kx_0 is feasible, but nonetheless the saturated LQR makes the closed-loop system unstable, whereas CLQR and DLQR stabilize the system in closed-loop (due to the infinite horizon formulation). Inputs and outputs for System 2 computed with constrained LQR at the decision times $0, 1, \ldots, 200$ are shown in Figure 4.

Finally, for System 1, we report in Table 1 the relative cost error, the number of required intervals, and CPU-time for solving all the QPs [3] for different values of μ. In computing the relative cost error, we approximate $V^*(x)$ with the value obtained using PWLH-CLQR with $\mu = 10^{-8}$. The data reported in Table 1 refer to the computation of $\mathbf{u}(x_0)$ for the same initial state considered in Figure 3. We report in Table 2 the same computational study for System 2. We can observe that, when the number of intervals is the same, PWLH-CLQR achieves a slightly lower cost than FFOH-CLQR. However, this comes at the expense of a higher CPU time because PWLH-CLQR

[3] Using GNU Octave on an AMD Athlon$^{\mathrm{TM}}$ 64 X2 Dual Core Processor 4400+ running Debian Linux.

Table 1. System 1. Comparison of cost relative error, number of intervals of the final QP and overall CPU-time for solving all QPs vs. relative tolerance μ for PWLH-CLQR and FFOH-CLQR

	PWLH-CLQR Algorithm			FFOH-CLQR Algorithm		
μ	$\frac{V^{\mathrm{I}}(x_0)-V^*(x_0)}{V^*(x_0)}$	N	CPU-time (s)	$\frac{V^{\mathrm{II}}(x_0)-V^*(x_0)}{V^*(x_0)}$	N	CPU-time (s)
10^{-2}	$1.755 \cdot 10^{-3}$	6	0.00012	$7.689 \cdot 10^{-4}$	12	0.00020
10^{-3}	$1.841 \cdot 10^{-4}$	24	0.00120	$1.883 \cdot 10^{-4}$	24	0.00040
10^{-4}	$4.495 \cdot 10^{-8}$	96	0.0376	$4.926 \cdot 10^{-8}$	96	0.0075
10^{-5}	$4.495 \cdot 10^{-8}$	96	0.0376	$4.926 \cdot 10^{-8}$	96	0.0075
10^{-6}	$4.495 \cdot 10^{-8}$	96	0.0376	$4.926 \cdot 10^{-8}$	96	0.0075
10^{-7}	$6.408 \cdot 10^{-9}$	192	0.360	$8.841 \cdot 10^{-9}$	192	0.0440
10^{-8}	—	384	3.264	$1.468 \cdot 10^{-9}$	384	0.444

Table 2. System 2. Comparison of cost relative error, number of intervals of the final QP and overall CPU-time for solving all QPs vs. relative tolerance μ for PWLH-CLQR and FFOH-CLQR

	PWLH-CLQR Algorithm			FFOH-CLQR Algorithm		
μ	$\frac{V^{\mathrm{I}}(x_0)-V^*(x_0)}{V^*(x_0)}$	N	CPU-time (s)	$\frac{V^{\mathrm{II}}(x_0)-V^*(x_0)}{V^*(x_0)}$	N	CPU-time (s)
10^{-2}	$3.3681 \cdot 10^{-3}$	6	0.00040	$6.4084 \cdot 10^{-3}$	6	0.00027
10^{-3}	$5.0657 \cdot 10^{-5}$	48	0.334	$1.1411 \cdot 10^{-4}$	48	0.0267
10^{-4}	$8.8931 \cdot 10^{-7}$	96	2.40	$1.0786 \cdot 10^{-6}$	192	2.29
10^{-5}	0	192	15.8	$1.0786 \cdot 10^{-6}$	192	2.29

optimizes over the input and the slope in each interval, while FFOH-CLQR optimizes only over the input with the slope fixed by continuity at the end point of the interval. For instance, for System 1, with $\mu = 10^{-3}$ PWLH-CLQR optimizes (in the last QP) over $2mN = 48$ decision variables while FFOH-CLQR over $m(N+1) = 25$ variables. Also notice that a relative tolerance between 10^{-2} and 10^{-3} results in small suboptimality. For these reasons, the most effective algorithm appears to be FFOH-CLQR, which achieves a solution in 0.40 ms with a relative cost error of about $1 \cdot 10^{-4}$. For System 2, a reasonable value for μ is also between 10^{-2} and 10^{-3}. For instance using FFOH-CLQR with $\mu = 10^{-3}$, we compute a solution in 27 ms with a relative cost error less than $2 \cdot 10^{-4}$. Such a computational time is clearly negligible compared to the systems dynamics. The resolution of the actuator in the application also implies a fairly loose tolerance ($\sim 10^{-3}$) on the solution should be used. It obviously makes little sense to compute an optimal input more accurately than the actuator hardware can resolve.

6. CONCLUSIONS

In this paper we presented a method for solving the infinite horizon continuous-time constrained linear quadratic regulator, by solving a finite number of finite dimensional quadratic programs. A number of unevenly spaced grid points are selected and adapted on-line to achieve a tolerance specification in the controller cost. The input parameterization is piecewise linear on the chosen time intervals; we examined both continuous and discontinuous parameterizations. The parameterizations guarantee exact input constraint satisfaction. Both of these input parameterizations converge to the optimal solution much more quickly than piecewise constant inputs, allowing a reduction in the final number of intervals and, therefore, decision variables. Furthermore, we derived exact discrete-time matrices and penalties to avoid quadrature errors, and moved offline all the computation required for solving ODEs and creating the Hessian and linear terms in the QPs solved online. Finally, we presented simulation results for two examples to illustrate the main ideas. The online complexity of solving the infinite horizon continuous-time CLQR with the proposed method is no larger than that required for solving an infinite horizon discrete-time CLQR problem of the same size.

REFERENCES

Cannon, M. and Kouvaritakis, B. (2000). Infinite horizon predictive control of constrained linear systems. *Automatica*, 36, 943–955.

Diehl, M., Ferreau, H.J., and Haverbeke, N. (2008). Efficient numerical methods for nonlinear MPC and moving horizon estimation. In *Proceedings of the International Workshop on Assessment and Future Directions of NMPC*. Pavia, Italy.

Goebel, R. and Subbotin, M. (2007). Continuous time linear quadratic regulator with control constraints via convex duality. *IEEE Trans. Auto. Contr.*, 52(5), 886–892.

Kojima, A. and Morari, M. (2004). LQ control of constrained continuous-time systems. *Automatica*, 40, 1143–1155.

Kwakernaak, H. and Sivan, R. (1972). *Linear Optimal Control Systems*. John Wiley & Sons.

Mayne, D.Q., Rawlings, J.B., Rao, C.V., and Scokaert, P.O.M. (2000). Constrained model predictive control: Stability and optimality. *Automatica*, 36(6), 789–814.

Pannocchia, G., Wright, S.J., and Rawlings, J.B. (2003). Existence and computation of infinite horizon model predictive control with active steady-state input constraints. *IEEE Trans. Auto. Contr.*, 48(6), 1002–1006.

Pannocchia, G., Rawlings, J.B., Mayne, D.Q., and Marquardt, W. (2009). On computing the solutions to the continuous time constrained linear quadratic regulator. Submitted for publication in IEEE Trans. Auto. Cont.

Prett, D.M. and Morari, M. (1987). *The Shell Process Control Workshop*. Butterworth Publishers.

Rao, C.V. and Rawlings, J.B. (1999). Steady states and constraints in model predictive control. *AIChE J.*, 45, 1266–1278.

Rossiter, J.A., Kouvaritakis, B., and Rice, M.J. (1998). A numerically robust state-space approach to stable-predictive control strategies. *Automatica*, 34, 65–73.

Sakizlis, V., Perkins, J.D., and Pistikopoulos, E.N. (2005). Explicit solutions to optimal contol problems for contrained continuous-time linear systems. *IEE Proc.-Control Theory Appl.*, 152(4), 443–452.

Yuz, J., Goodwin, G., Feuer, A., and De Doná, J. (2005). Control of constrained linear systems using fast sampling rates. *Syst. Contr. Lett.*, 54, 981–990.

Explicit Robust Model Predictive Control ⋆

Efstratios N. Pistikopoulos * **Nuno P. Faísca** **
Konstantinos I. Kouramas * **Christos Panos** *

* *Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, SW7 2AZ London, UK (e-mail: e.pistikopoulos, k.kouramas, christos.panos08@imperial.ac.uk)*
** *Process Systems Enterprise Ltd, London, UK (e-mail: pitigrilli@gmail.com)*

Abstract: Explicit robust multi–parametric feedback control laws are designed for constrained dynamic systems involving uncertainty in the left-hand side(LHS) of the underlying MPC optimization model. Our proposed procedure features: (i) a robust reformulation/optimization step, (ii) a dynamic programming framework for the model predictive control (MPC) problem formulation, and (iii) a multi-parametric programming solution step.

Keywords: Robust model predictive control, multi-parametric programming, dynamic programming, multi–parametric control, explicit control.

1. INTRODUCTION

Robust model predictive control (Robust MPC) is an important class of constrained, model–based control methods that can explicitly account for the presence of modeling uncertainties in the controlled process, which has received significant attention in control systems research–an indicative list of related publications is given in (Bemporad and Morari, 1999; Mayne et al., 2000; Sakizlis et al., 2004; Wang and Rawlings, 2004; Pistikopoulos et al., 2007a) and references within. On the other hand, explicit MPC, which has also received equal attention recently (Pistikopoulos et al., 2002, 2007a), is a control method where the online MPC optimization problem is solved off–line with multi–parametric programming methods to obtain the optimal control actions as a set of functions of the system states. The MPC controller can then be implemented online as a set of simple feedback control laws based on function evaluations instead of using online optimization with complex and increased computational demands.

Despite these significant advances, explicit, robust MPC is still an important area of research. It is evident from the relevant literature (Bemporad et al., 2003; Wang and Rawlings, 2004; Pistikopoulos et al., 2007a) that, even for the case of linear MPC, the underlying optimization model of the MPC is nonlinear due to the uncertainties appearing both in the left–hand side and right–hand side of the optimization constraints (Borrelli, 2003; Pistikopoulos et al., 2007a). This imposes difficulties for the application of the existing multi–parametric programming techniques and special treatment is required to ensure that the constraints are always satisfied (Bemporad et al., 2003; Kouramas et al., 2009).

⋆ This work is supported by EPSRC (GR/T02560/01, EP/E047017/1) and European Commission (PRISM ToK project, Contact No: MTKI-CT-2004-512233 and DIAMANTE ToK project, Contract No: MTKI-CT-2005-IAP-029544)

Explicit robust MPC was investigated in Sakizlis et al. (2004) for the case of linear dynamic systems with additive state disturbances (right–hand side uncertainty in the optimization model). A dynamic programming based method, for linear dynamic systems with linear objective costs and uncertainties in left–hand side of the optimization model was studied in Bemporad et al. (2003). Furthermore, an explicit robust MPC with a quadratic objective and left–hand side uncertainties, based on robust optimization methods (Ben-Tal and Nemirovski, 2000; Lin et al., 2004), was presented in Kouramas et al. (2009) where the MPC optimization is treated as a robust multi–parametric optimization problem. Explicit robust MPC problems with quadratic costs have not yet been fully studied since the underlying multi–parametric optimization problem becomes nonlinear due to the uncertain coefficients in the constraints (Kouramas et al., 2009). On the other hand, employing dynamic programming methods for even the simple case of explicit MPC (with no uncertainties) results either into solving a demanding global optimization problem (Faísca et al., 2008) at each stage of the dynamic programming procedure or overlapping critical regions in the explicit solution.

This work presents a novel method for *Explicit Robust Model Predictive Control* based on dynamic programming methods (Bellman (2003); Faísca et al. (2008)) and robust optimization techniques (Ben-Tal and Nemirovski, 2000; Lin et al., 2004) that (i) allows the use of quadratic objective functions, (ii) accounts for the uncertainties in the left–hand side of the underlying MPC optimization problem, and (iii) overcomes the limitations of previous methods and the need for global optimization at each stage of the dynamic programming.

We focus on the following explicit robust MPC problem

$$
\begin{aligned}
V^*(x) &= \min_U J(U,x) \\
&= \min_U \sum_{k=0}^{N-1} \{x_k^T Q x_k + u_k^T R u_k\} + x_N^T P x_N \qquad (1)\\
&\text{s.t. } x_{k+1} = Ax_k + Bu_k,\ \forall\, \Delta A \in \mathcal{A},\ \Delta B \in \mathcal{B} \\
&\quad Cx_k \le d,\ \ k = 0,1,\ldots,N \\
&\quad Mu_k \le \mu,\ \ k = 0,1,\ldots,N-1 \\
&\quad Tx_N \le \tau \\
&\quad x = x_0
\end{aligned}
$$

where $x \in \mathbb{R}^n$ is the system state, $u \in \mathbb{R}^m$ is the system input and N the prediction horizon. We assume that the underlying system is uncertain in that the system matrices are described as

$$
\begin{aligned}
&x_{k+1} = Ax_k + Bu_k,\ A = A_0 + \Delta A,\ B = B_0 + \Delta B \qquad (2)\\
&\Delta A \in \mathcal{A} = \{\Delta A \in \mathbb{R}^{n\times n} | -\varepsilon_a |A_0| \le \Delta A \le \varepsilon_a |A_0|\} \\
&\Delta B \in \mathcal{B} = \{\Delta B \in \mathbb{R}^{n\times n} | -\varepsilon_\beta |B_0| \le \Delta B \le \varepsilon_\beta |B_0|\}
\end{aligned}
$$

where A_0, B_0 are of known constant values but the values of matrices $\Delta A, \Delta B$ are not known but are bounded as given in (2) and $\epsilon_a, \epsilon_\beta \in [0,1)$. The system states and inputs are also subject to the following linear constraints

$$x \in \mathcal{X} = \{x \in \mathbb{R}^n | Cx \le d\} \qquad (3)$$

$$u \in \mathcal{U} = \{u \in \mathbb{R}^m | Mu \le \mu\} \qquad (4)$$

where the sets $\mathcal{X}$,$\mathcal{U}$ are assumed to be compact, non-empty polytopic sets that include the origin in their interior and with $C \in \mathbb{R}^{n_c \times n}$, $d \in \mathbb{R}^{n_c}$, $M \in \mathbb{R}^{m_M \times m}$ and $\mu \in \mathbb{R}^{m_M}$. The proposed approach and the underlying mathematical framework for solving (1) will be discussed in detail in the following sections.

2. EXPLICIT ROBUST MODEL PREDICTIVE CONTROL

The proposed approach is realized in three key steps:

(1) dynamic programming: the MPC optimization is recast in a multi-stage optimization setting,
(2) robust reformulation: the constraints at each stage are reformulated to account for the worst-case uncertainty, and
(3) multi-parametric programming: each one of the reformulated stages is solved as multi-parametric programming problems where the optimization variables are the incumbent control inputs, given the optimal solutions of the previous steps.

These steps are described in detail in the following.

2.1 Dynamic Programming – Multi-stage optimization

The robust MPC problem (1) can be expressed as a multi–stage optimization problem since it involves a discrete-time dynamic system and a stage–additive quadratic objective function. The same procedure was applied for the nominal system case (where $\varepsilon_a, \varepsilon_\beta = 0$) in Faísca et al. (2008) Dynamic programming techniques (Bellman, 2003) can be applied to decompose (1) into a set of *stage-wise* problems of smaller dimensions, significantly reducing the complexity of the initial problem (Bellman (2003) and Faísca et al. (2008)) - at each stage k the following optimization problem is considered

$$
\begin{aligned}
V_k(x_k) &= \min_{u_k \in \mathcal{U}} J_k(u_k, x_k) \\
&= \min_{u_k \in \mathcal{U}} \sum_{i=k}^{N-1} \{x_i^T Q x_i + u_i^T R u_i\} + x_N^T R x_N \qquad (5)\\
&\text{s.t. } x_{i+1} = Ax_i + Bu_i,\ i = k,\ldots,N \\
&\quad Cx_k \le d,\ \ Cx_{k+1} \le d,\ \ Mu_k \le \mu, \\
&\quad \forall\, \Delta A \in \mathcal{A},\ \Delta B \in \mathcal{B}
\end{aligned}
$$

The optimization is taken only on the current stage input u_k and only the constraints on x_k and x_{k+1} have to be considered. The main idea is to solve the single–stage optimization problem (5) as a robust mp–QP problem and obtain the control variable u_k at each stage as an explicit function of current state x_k

$$u_k = f_k^*(x_k) \qquad (6)$$

or

$$u_k = K_k^i x_k + c_k^i \quad \textbf{if } \ x_k \in CR_k^i\ ,\ \ i = 1,\ldots,L_k$$

A method for solving (5) as a robust mp–QP problem and deriving (6) is presented in the following sections. The proposed procedure for solving (1) as a multi-stage problem is the following: starting from time $k = N-1$, problem (5) is solved iteratively at each time k until $k = 0$ where the procedure stops. At the initial stage $k = N-1$ the extra terminal constraint $Tx_N \le \tau$ should also be added in (5).

In order to ensure that a feasible solution u_k exists for all $k = 0,1,\ldots N-1$ an extra feasibility constraint is introduced in each of the single stage problems (5)

$$x_{k+1} \in \mathcal{X}^{k+1}\ ,\ \ \mathcal{X}^{k+1} = \cup_{i=1}^{L_{k+1}} CR_{k+1}^i \qquad (7)$$

where $\mathcal{X}^{k+1}$ is the union of all critical regions of the explicit solution $u_{k+1} = f_{k+1}^*(x_{k+1})$ from the previous stage $k+1$ i.e. $\mathcal{X}^{k+1}$ is the set of states x_{k+1} for which the optimization problem at the stage $k+1$ has a feasible solution. Since the set of all critical regions is a convex polyhedral set (Pistikopoulos et al., 2002), the set $\mathcal{X}^{k+1}$ is given by a set of linear inequalities

$$\mathcal{X}^{k+1} = \{x \in \mathbb{R}^n | H^{k+1} x \le h^{k+1}\} \qquad (8)$$

Adding the constraints (8) in (5) will ensure that the future state x_{k+1} lies in the set $\mathcal{X}^{k+1}$ and hence one of the critical regions CR_{k+1}^i, and therefore a feasible control input $u_{k+1} = f_{k+1}^*(x_{k+1})$ at time $k+1$ can be obtained. For simplicity the inequalities $Cx_{k+1} \le d$ and (8) will be replaced by the single inequality

$$\mathcal{G}^k x_{k+1} \le b^k$$

where $\mathcal{G}^k = [C^T \ \ {H^{k+1}}^T]^T$ and $b^k = [d^T \ \ {h^{k+1}}^T]^T$.

We will now proceed to describe how to reformulate (5) to a robust mp–QP problem. Considering u_k as the optimization variable and $\theta_k = \left[x_k^T\ u_{k+1}^T \ldots u_{N-1}^T\right]^T$ as the vector of parameters, and by incorporating the system dynamics $x_{k+1} = Ax_k + Bu_k$ into the objective and constraints, one obtains the following multi-parametric optimization problem

$$
\begin{aligned}
V_k(x_k) &= \min J_k(u_k, \theta_k) \\
&= \min_{u_k \in \mathcal{U}} \left\{ \frac{1}{2} u_k^T H u_k + \theta_k^T F u_k \right\} + \theta_k^T Y \theta_k \qquad (9)\\
&\text{s.t. } \mathcal{G}^k A x_k + \mathcal{G}^k B u_k \le b^k,\ \ Cx_k \le d,\ \ Mu_k \le \mu \\
&\quad \forall\, \Delta A \in \mathcal{A},\ \Delta B \in \mathcal{B}
\end{aligned}
$$

where the matrices H, F, Y are functions of the matrices A, B, Q and R. When there is no uncertainty in the underlying system dynamics $\varepsilon_a = \varepsilon_\beta = 0$, (5) is a simple mp–QP problem and can be solved with the known mp–QP method (Pistikopoulos et al., 2007b). However, in the presence of uncertainty (when $\varepsilon_a, \varepsilon_\beta \in [0, 1)$ are non-zeros) special treatment of (9) is required to reformulate it into mp–QP problem.

Remark 1. In conventional dynamic programming, the optimal value $u_{k+1} = f^*_{k+1}(x_{k+1})$ would have been incorporated into the formulation of (14) to create an optimization problem where only u_k is the optimization variable and x_k the parameter. However, even for the simple case with no uncertainties, this would have resulted into a nonlinear multi–parametric programming problem (since $u_{k+1} = f^*_{k+1}(x_{k+1})$ is a piecewise affine function) that would need to employ global optimization methods to be solved (Borrelli, 2003; Faísca et al., 2008). Our approach is based on the work of Faísca et al. (2008) for the case of explicit MPC with no uncertainties, where this issue is overcome by substituting previous solutions u_{k+1} in the current solution u_k after the multi–parametric programming has been solved.

2.2 Robustification Step

The main issue for applying multi-parametric optimization techniques for the solution of (9) is the presence of the uncertain matrices A, B in the objective and the inequalities of (9). The objective function can be set to penalize only the behaviour of the nominal system $x_{k+1} = A_0 x_k + B_0 u_k$, that is to say the objective function in (9) is formed by replacing $x_{k+1} = A_0 x_k + B_0 u_k$ in the objective (5) and H, F, Y are constant matrices. However, it is very important to guarantee the feasibility of the constraints in the presence of the uncertainty. Problem (9) can then be recast as

$$V_k(x_k) = \min_{u_k \in \mathcal{U}} \left\{ \frac{1}{2} u_k^T H u_k + \theta_k^T F u_k \right\} + \theta_k^T Y \theta_k \qquad (10)$$
$$\text{s.t. } \mathcal{G}^k A_0 x_k + \mathcal{G}^k \Delta A x_k + \mathcal{G}^k B_0 u_k + \mathcal{G}^k \Delta B u_k \le b^k$$
$$C x_k \le d, \;\; M u_k \le \mu, \;\; \forall\, \Delta A \in \mathcal{A}, \; \Delta B \in \mathcal{B}$$

It is obvious from (10) that due to variations of $\Delta A, \Delta B$ constraint violations might occur. Solving (10) is a robust multi-parametric optimization problem where u_k is the optimization variable and θ_k is the vector of parameters. The objective is to find a solution $u^*_k(\theta_k)$ which can guarantee constraint satisfaction for all admissible values of the uncertainty i.e. for all $\Delta A \in \mathcal{A}$ and $\Delta B \in \mathcal{B}$.

Definition 2.1. A solution $u^*_k(\theta_k)$ of robust mp-QP problem (10) is a *robust* or *reliable* solution if it is feasible for (10) both for the nominal system ($A = A_0$, $B = B_0$) and the uncertain system i.e. if it is feasible for all admissible values of the uncertainty i.e. for all $\Delta A \in \mathcal{A}$ and $\Delta B \in \mathcal{B}$.

In order to avoid constraint violations, the constraints have to be *immunized* against the model uncertainty (see Ben-Tal and Nemirovski (2000) and Lin et al. (2004)). In order to account for the uncertainty in (10), the inequality constraints of (10) are replaced by the following two inequalities

$$\mathcal{G}^k A_0 x_k + \mathcal{G}^k B_0 u_k \le d \qquad (11)$$

$$\mathcal{G}^k A_0 x_k + \varepsilon_a |\mathcal{G}^k| |A_0| |x_k| + \mathcal{G}^k B_0 u_k + \varepsilon_\beta |\mathcal{G}^k| |B_0| |u_k| \le b^k + \delta \max\{1, |d|\} \qquad (12)$$

The first inequality ensures that the problem is feasible for the nominal system case while the second inequality represents the realisation of the constraint for the worst-case value of the uncertainty. The newly introduced variable δ is a measure of the infeasibility tolerance for the constraint in the problem i.e. how much the constraint can be relaxed to ensure a feasible solution. If no infeasibility is allowed then $\delta = 0$.

Replacing the new constraints (11)–(12) into (10) results into a multi–parametric nonlinear programming problem. To overcome this, (12) is replaced by the following linear inequalities

$$\begin{aligned} &\mathcal{G}^k A_0 x_k + \varepsilon_a |\mathcal{G}^k| |A_0| z_k + \mathcal{G}^k B_0 u_k \\ &+ \varepsilon_\beta |\mathcal{G}^k| |B_0| \omega_k \le b^k + \delta \max\{1, |d|\} \\ &- z_k \le x_k \le z_k, \;\; -\omega_k \le u_k \le \omega_k, \;\; z_k, \omega_k \ge 0 \end{aligned} \qquad (13)$$

It is obvious that if a pair x_k, u_k satisfies (13) then, since $|x_k| \le z_k$ and $|u_k| \le \omega_k$, it also satisfies (12). By replacing (13) in (10) the new robust mp–QP formulation is obtained for each stage

$$V_k(x_k) = \min_{u_k, z_k, \omega_k} \left\{ \frac{1}{2} u_k^T H u_k + \theta_k^T F u_k \right\} + \theta_k^T Y \theta_k \qquad (14)$$
$$\begin{aligned} &\text{s.t. } \mathcal{G}^k A_0 x_k + \mathcal{G}^k B_0 u_k \le b^k \\ &\mathcal{G}^k A_0 x_k + \varepsilon_a |\mathcal{G}^k| |A_0| z_k + \mathcal{G}^k B_0 u_k \\ &\quad + \varepsilon_\beta |\mathcal{G}^k| |B_0| \omega_k \le b^k + \delta \max\{1, |d|\} \\ &\quad - z_k \le x_k \le z_k, \;\; -\omega_k \le u_k \le \omega_k, \;\; z_k, \omega_k \ge 0 \\ &\quad C x_k \le d, \;\; M u_k \le \mu \end{aligned}$$

where now the parameters are θ_k, the optimization variable is $\pi_k = \left[u_k^T, z_k^T, \omega_k^T\right]^T$, the objective function is a quadratic function and the constraints are all linear inequalities. The new formulation (14) is an mp–QP problem and can be solved by employing the mp–QP methods of Pistikopoulos et al. (2002) and Pistikopoulos et al. (2007b) which is discussed next.

2.3 Multi–Parametric Quadratic Programming

In order to solve (14) as an mp–QP problem, the following three steps have to be followed

Step 1. The Karush–Kuhn–Tucker (KKT) conditions are first applied for problem (14) (see Bazaraa and Shetty (1979)):

$$\nabla \mathcal{L}(\pi_k, \lambda, \theta_k) = 0, \;\; \lambda_i \psi_i(\pi_k, \theta_k) = 0, \quad \forall \;\; i = 1, \ldots, p$$
$$\mathcal{L} = J_k(\pi_k, \theta_k) + \sum_{i=1}^{p} \lambda_i \psi_i(\pi_k, \theta_k) \qquad (15)$$

where $J_k(\pi_k, \theta_k)$ is the objective function of (14), $\psi(\pi_k, \theta_k) \le 0$ is the vector of the inequality constraints in (14) and λ is the vector of the Lagrange multipliers.

Step 2. The basic sensitivity theorem (Fiacco (1976)) is then applied to the KKT conditions (15). For simplicity we set $\theta = \theta_k$ and $\pi = \pi_k$.

Theorem 2. Let θ_0 be a vector of parameter values and $(\pi_0, \lambda_0, \mu_0)$ a KKT triple corresponding to (15), where

λ_0 is nonnegative and π_0 is feasible in (14). Also assume that (i) strict complementary slackness (SCS) holds, (ii) the binding constraint gradients are linearly independent (LICQ: Linear Independence Constraint Qualification), and (iii) the second-order sufficiency conditions (SOSC) hold. Then, in neighbourhood of θ_0, there exists a unique, once continuously differentiable function, $z(\theta) = [\pi(\theta), \lambda(\theta)]$, satisfying (15) with $z(\theta_0) = [\pi(\theta_0), \lambda(\theta_0)]$ where $\pi(\theta)$ is a unique isolated minimiser for (14), and

$$\begin{pmatrix} d\pi(\theta_0)/d\theta \\ d\lambda(\theta_0)/d\theta \end{pmatrix} = -(M_0)^{-1} N_0, \tag{16}$$

where, M_0 and N_0 are the Jacobians of system (15) with respect to z and θ (Fiacco, 1983, pp. 80–81), (Pistikopoulos et al., 2002).

Step 3. A general analytic expression for π_k is then derived by applying the following corollary of Dua et al. (2002)

Corollary 3. First-order estimation of $\pi(\theta)$, $\lambda(\theta)$, near $\theta = \theta_0$ (Fiacco, 1983): Under the assumptions of Theorem 2, a first-order approximation of $[\pi(\theta), \lambda(\theta)]$ in a neighbourhood of θ_0 is,

$$\begin{bmatrix} \pi(\theta) \\ \lambda(\theta) \end{bmatrix} = \begin{bmatrix} \pi_0 \\ \lambda_0 \end{bmatrix} + (M_0)^{-1} \cdot N_0 \cdot \theta + o(||\theta||), \tag{17}$$

where $(\pi_0, \lambda_0) = [\pi(\theta_0), \lambda(\theta_0)]$, $M_0 = M(\theta_0)$, $N_0 = N(\theta_0)$, and $\phi(\theta) = o(||\theta||)$ means that $\phi(\theta)/||\theta|| \to 0$ as $\theta \to \theta_0$.

The critical region (set of θ) where (17) remains optimal can then be obtained as follows (Dua et al., 2002). If $\breve{\psi}$ corresponds to the non-active constraints, and $\tilde{\lambda}$ corresponds to the active constraints then each critical region is defined by

$$\begin{array}{ll} \breve{\psi}(u(\theta_k), \theta_k) \leq 0 & \text{(Feasibility conditions)}, \\ \tilde{\lambda}(\theta_k) \geq 0 & \text{(Optimality conditions)}. \end{array} \tag{18}$$

It is obvious from step 1.–3. and corollary 3 that the the explicit solution π_k^* of (14) is given by a conditional piecewise linear function (Dua et al. (2002) and Pistikopoulos et al. (2007a)) i.e. $\pi_k = f_k^*(\theta_k)$. Consequently, the control u_k is also obtained as an explicit function of the parameter θ_k as follows

$$u_k = f_k^*(\theta_k) = f_k^*(x_k, u_{k+1}, \ldots, u_{N-1}) \tag{19}$$

or

$$u_k = K_k^i \theta_k + c_k^i, \ \textbf{if} \ \ \theta_k \in CR_k^i, \ i = 1, \ldots, L_k \tag{20}$$

where K_k^i, c_k^i are matrices and vectors of appropriate dimensions and the critical regions $CR_k^i \subset \mathbb{R}^n$ are sets defined by (18). The same procedure repeats iteratively, starting at $k = N-1$ and stopping at $k = 0$ and hence the full profile of control policies $u_k(\theta_k)$, $k = 0, 1, \ldots, N-1$ is derived.

Although u_k is a function of θ_k, the objective is to obtain u_k as an explicit control function of the incumbent state x_k thus obtaining a feedback control strategy. We can overcome this issue by following an approach similar to Faísca et al. (2008) for the nominal explicit MPC case. As the procedure is repeated repetitively and backwards from $k = N-1$ to $k = 0$, the control inputs $u_{k+1}, \ldots, u_{N-1}$ before stage k are obtained as in (19)

$$\begin{array}{rl} u_{k+1} & = f_{k+1}^*(x_{k+1}, u_{k+2}, \ldots, u_{N-1}) \\ & \vdots \\ u_{N-1} & = f_{N-1}^*(x_{N-1}) \end{array} \tag{21}$$

All the above control inputs are piecewise linear functions of their arguments. Note also that since the control inputs $u_{k+1}, \ldots, u_{N-1}$ are functions of the future states $x_{k+1}, \ldots, x_{N-1}$ they are also functions of the incumbent input u_k and state x_k. By incorporating the previous solutions (21) into (19) and by performing algebraic manipulation we obtain the explicit control law $u_k = f_k^*(x_k)$ (see for more details Faísca et al. (2008)). The final critical regions of $u_k = f_k^*(x_k)$ are defined as a union of the inequalities (of the critical regions) of (19) and of each of the critical regions of (21). This results in (i) realisable feasible sets of inequalities describing the feasible critical regions of $u_k = f_k^*(x_k)$ and (ii) empty sets of inequalities where no feasible solution exists. Feasibility tests, as the ones presented in Faísca et al. (2008), are finally performed, during the substitution of (21) into (19), to obtain the final feasible critical regions.

2.4 Algorithm for Robust mp–MPC

The dynamic programming based procedure that was described above is summarized in table 1. The Algorithm starts at $k = N-1$ and iterates through Steps 2 and 3 until $k = 0$. At the k^{th} stage of the algorithm, problem (14) is solved following the analysis in sections 2.1–2.3. Each of the inputs u_k is obtained as an explicit function of the corresponding state x_k i.e. $u_k = f_k(x_k)$ where $f_k(x_k)$ is a piecewise linear function similar to (6). At the termination of the algorithm a sequence of admissible control policies is obtained $u_0^* = f_0^*(x_0)$, $u_1^* = f_1^*(x_1)$, $\ldots$, $u_{N-1}^* = f_{N-1}^*(x_{N-1})$. Each of these control policies are *reliable* (or *robust*) control policies for each of the stage problems (14). Since each control policy also guarantees that the state and input constraints $x_k \in \mathcal{X}$ and $u_k \in \mathcal{U}$ at each stage are satisfied, then the control sequence $U = \{u_0^*, u_1^* \ldots u_{N-1}^*\}$ is also a robust solution for the initial robust mp-MPC problem (1). The following lemma can then be stated

Lemma 4. The control sequence $U = \{u_0^*, u_1^* \ldots u_{N-1}^*\}$, where u_k^*, $k = 0, 1, \ldots, N-1$ are the optimal control policies obtained by solving (14) iteratively using the algorithm in table 1, is a robust (or reliable solution) of (1).

Table 1. Algorithm for Robust Multi–Parametric MPC

Step 1.	Set $k = N-1$: solve the mp-QP problem (14) with x_{N-1} being the parameters and obtain $u_{N-1}^* = f_{N-1}^*(x_{N-1})$.
Step 2.	Set k to the current stage: solve the k^{th} stage–wise mp–QP problem (14) with $x_k, u_k, \ldots, u_{N-1}$ being the parameters and obtain $u_k^* = f_k^*(x_k, u_k, \ldots, u_{N-1})$.
Step 3.	Obtain the control law $u_k = f_k(x_k)$ by comparing the sets of solutions (19) and (21).
Step 4.	Set $k = k-1$: if $k = 0$ stop, else go to Step 2.

The main advantage of the proposed algorithm is that it can handle robust Model Predictive Control problems with quadratic objectives in the presence of uncertainties in the

LHS of the underlying optimization model at each stage of the proposed dynamic programming procedure. This is achieved by treating the optimization problem for each stage of the procedure as a convex robust mp–QP problem (14) with linear constraints, avoiding the nonlinearities introduced by the presence of uncertainty in (11).

The introduction of the two new variables z, ω also results in an increase of the number of constraints in the optimization as it can be seen from (13). The total number of optimization variables in the resulting robust mp–QP problem (14) is $2m + n$, while the total number of linear inequalities is $2n_c + m_M + n + m$. One can notice that both the number of optimization variables and inequalities for problem (14), after the robustification step, is linear with the number of system states and inputs. Thus the complexity of the mp–QP problem is not significantly increased. Finally, the number of parameters of the mp–QP (14) at each stage is equal to $n + (N - k - 1)m$, hence it increases as k decreases. This will have an important effect on the number of critical regions at each stage and eventually in the overall number of critical regions.

3. EXAMPLE

Consider the following robust MPC example where

$$A_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \ , \ B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \ , \ \varepsilon_a = \varepsilon_\beta = 0.2$$

$$\begin{bmatrix} -10 \\ -10 \end{bmatrix} \leq x_k \leq \begin{bmatrix} 10 \\ 10 \end{bmatrix} \ , \ -1 \leq u_k \leq 1$$

$$Q = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \ , \ R = 1 \ , \ P = \begin{bmatrix} 2.6005 & 2.081 \\ 2.081 & 3.3306 \end{bmatrix} \ , \ N = 3$$

The target set in this example is considered to be simply the set of state constraints while $\delta = 0$ is set equal to zero. The algorithm, presented in Table 1, is applied and the results can be seen in Figures 1, 2, 3. In the first iteration of the algorithm the robust multi-parametric programming problem (14) for $k = 2$ is solved, where the parameter is $\theta_2 = x_2$. The critical regions of the explicit solution $u_2 = f_2(x_2)$ are shown in Figure 1. Then, the procedure is repeated for the stages $k = 1$ and $k = 0$ to obtain the explicit controls $u_1 = f_1(x_1)$ and $u_0 = f_0(x_0)$. The critical regions for these stages are shown in Figures 2 and 3 respectively. One can notice that the area of the critical regions at each stage k decreases as k decreases. This happens since the set of states which can be driven to the target set (which here is the set of constraints) reduces as k reduces. Also, the number of critical regions increases since at each stage the number of parameters increases. Two different simulations for two different initial states of the explicit robust MPC control is shown in Figure 4 where the system matrices A, B are perturbed around their nominal values. Finally, table 1 shows some of the critical regions and corresponding control functions for the explicit solution at stage 0.

4. CONCLUDING REMARKS

A new algorithm for robust multi–parametric MPC was presented when uncertainty is introduced in the LHS of the

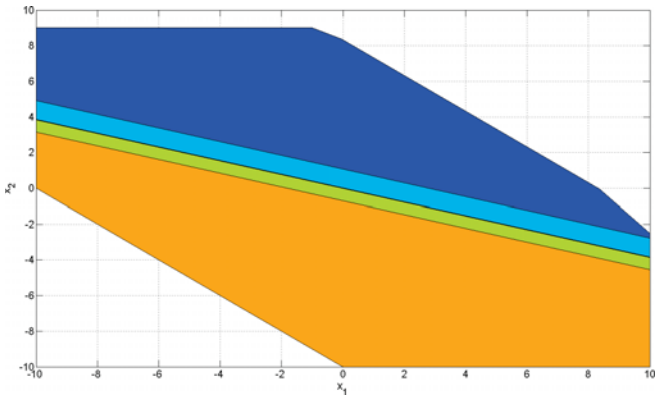


Fig. 1. Critical regions of the explicit robust MPC for stage 2, $u_2 = f_2(x_2)$

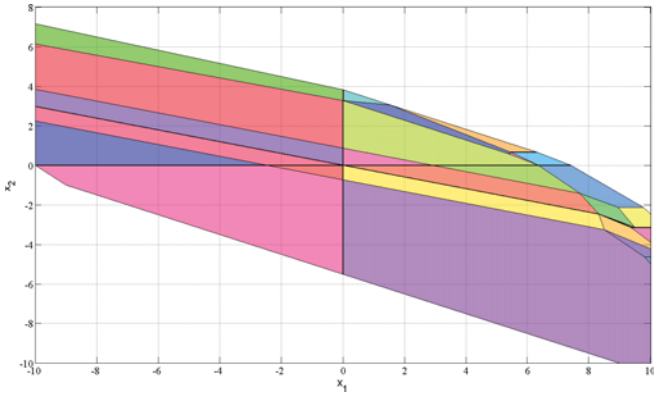


Fig. 2. Critical regions of the explicit robust MPC for stage 1, $u_1 = f_1(x_1)$

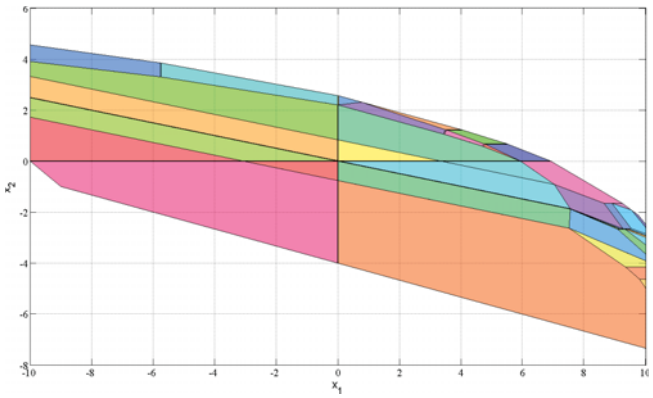


Fig. 3. Critical regions of the explicit robust MPC for stage 0, $u_0 = f_0(x_0)$

underlying MPC optimization model. Based on dynamic programming and robust optimization, the algorithm obtains the control input explicitly as function of the states by solving a set of convex mp–QP problems and avoid the need for employing multi–parametric global optimization. Current work is focusing on the generalisation of the presented results to the following problems: (i) explicit robust MPC of constrained dynamic systems with uncertainty

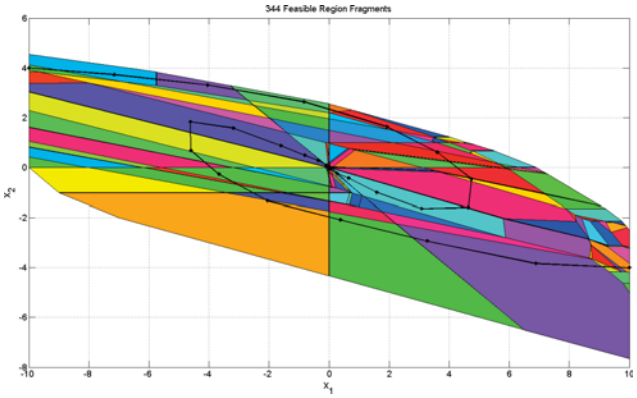


Fig. 4. Simulation of the uncertain system state trajectory with explicit robust MPC.

Table 2. Critical Regions and the corresponding Control Laws for stage 0

Critical Regions No.	Control Law	Critical Regions
1	$u = 1$	$\begin{bmatrix} 0.2174 & 1 \\ 0 & 1 \\ -0.5 & -1 \\ -0.3333 & -1 \\ 1 & 0 \end{bmatrix} x \leq \begin{bmatrix} -1.748 \\ -1.01 \\ 5.5 \\ 4.3333 \\ -0.01 \end{bmatrix}$
2	$u = 1$	$\begin{bmatrix} 0.2174 & 1 \\ -0.3333 & -1 \\ -1 & 0 \\ 1 & 0 \end{bmatrix} x \leq \begin{bmatrix} -1.7480 \\ 4.3333 \\ 0.01 \\ 0.0050 \end{bmatrix}$
3	$u = -1$	$\begin{bmatrix} -0.2513 & -1 \\ 0.9 & 1 \\ 0 & 1 \\ 0 & -1 \\ -0.7821 & -1 \\ -1 & -0.9524 \end{bmatrix} x \leq \begin{bmatrix} -0.1791 \\ 6.65 \\ -1.6562 \\ 2.1250 \\ -4.5301 \\ -6.3333 \end{bmatrix}$
4	$u = -0.5662x_1 - 1.3573x_2 + 1.0378$	$\begin{bmatrix} -1 & -0.2757 \\ 0.4172 & 1 \\ 1 & 0.631 \\ -0.4172 & -1 \\ -1 & -0.555 \end{bmatrix} x \leq \begin{bmatrix} -7.8743 \\ 1.5014 \\ 7.3523 \\ -0.7720 \\ -7.0852 \end{bmatrix}$
5	$u = -0.4701x_1 - 1.3476x_2 - 0.0001$	$\begin{bmatrix} 1 & 0.9042 \\ -1 & -1 \\ -1 & -0.7395 \\ -1 & -0.1585 \\ 1 & 1 \end{bmatrix} x \leq \begin{bmatrix} 0.079 \\ 0.01 \\ -0.0211 \\ -0.0579 \\ 0.005 \end{bmatrix}$

and additive disturbance, both in the LHS and RHS of the underlying multi–parametric optimization model (Sakizlis et al. (2004)), (ii) explicit robust MPC of hybrid systems – based on multi–parametric Mixed Integer Linear Programming (Faísca et al. (2009)) and (iii) multi–parametric Global Optimisation (Dua et al. (2004))..

REFERENCES

Bazaraa, M. and Shetty, C. (1979). *Nonlinear Programming–Theory and Algorithms*. Wiley, New York.

Bellman, R. (2003). *Dynamic programming*. Dover Publications.

Bemporad, A., Borrelli, F., and Morari, M. (2003). Min-max control of constrained uncertain discrete–time linear systems. *IEEE Trans. Aut. Con.*, 48, 1600–1606.

Bemporad, A. and Morari, M. (1999). Robustness in identification and control: A survey. In A. Garulli, A. Tesi, and A. Vicino (eds.), *Robustness in identification and control*. Springer-Verlag, Boston, USA.

Ben-Tal, A. and Nemirovski, A. (2000). Robust solutions of linear programming problems contaminated with uncertain data. *Math. Prog.*, 88, 411–424.

Borrelli, F. (2003). *Constrained optimal control of linear and hybrid systems*, volume 290 of *Lecture notes in control and information sciences*. Berlin: Springer.

Dua, V., Bozinis, A., and Pistikopoulos, E. (2002). A multiparametric programming approach for mixed-integer quadratic engineering problems. *Comput. Chem. Eng.*, 26, 715–733.

Dua, V., Papalexandri, K., and Pistikopoulos, E. (2004). Global optimization issues in multiparametric continuous and mixed–integer optimization problems. *Journal of Global Optimization*, 30, 59–89.

Faísca, N., Kosmidis, V., Rustem, B., and Pistikopoulos, E. (2009). Global optimization of multi–parametric milp problems. *Journal of Global Optimization*, to appear in Special Issue of JOGO for the Workshop on Global Optimization.

Faísca, N., Kouramas, K., Saraiva, P., Rustem, B., and Pistikopoulos, E. (2008). A multi-parametric programming approach for ocnstrained dynamic programming problems. *Optimization Letters*, 2, 267–280.

Fiacco, A. (1976). Sensitivity analysis for nonlinear programming using penalty methods. *Math. Prog.*, 10, 287–311.

Fiacco, A. (1983). *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*. Academic Press, New York.

Kouramas, K., Sakizlis, V., and Pistikopoulos, E. (2009). Design of robust model-based controllers via multiparametric programming. *Encyclopedia of Optimization*, 677–687.

Lin, X., Janak, S., and Floudas, C. (2004). A new robust optimization approach for scheduling under uncertainty: I. bounded uncertainty. *Comp. Chem. Eng.*, 28, 1069–1085.

Mayne, D., Rawlings, J., Rao, C., and Scokaert, P. (2000). Constrained model predictive control: stability and optimality. *Automatica*, 36, 789–814.

Pistikopoulos, E., Dua, V., Bozinis, N., Bemporad, A., and Morari, M. (2002). On-line optimization via off-line parametric optimization tools. *Comp. Chem. Eng.*, 26, 175–185.

Pistikopoulos, E., Georgiadis, M., and Dua, V. (2007a). *Multi-parametric model-based control: theory and applications*, volume 2 of *Process Systems Engineering Series*. Wiley-VCH, Weinheim.

Pistikopoulos, E., Georgiadis, M., and Dua, V. (2007b). *Multi-parametric Programming: Theory, Algorithms and Applications*, volume 1 of *Process Systems Engineering Series*. Wiley-VCH, Weinheim.

Sakizlis, V., Kakalis, N., Dua, V., Perkins, J., and Perkins, E. (2004). Design of robust model-based controllers via parametric programming. *Automatica*, 40, 189–201.

Wang, Y. and Rawlings, J. (2004). A new robust model predictive control method i: theory and computation. *J. of Proc. Con.*, 14, 231–247.

Robust Adaptive MPC for Systems with Exogeneous Disturbances ⋆

V. Adetola * **M. Guay** *

* *Department of Chemical Engineering, Queen's University, Kingston, Ontario, Canada (e-mail: martin.guay@chee.queensu.ca)*

Abstract: In this paper, we consider the problem of Adaptive model predictive control subject to exogenous disturbances. Using a novel set-based adaptive estimation, the problem of robust adaptive MPC is proposed and solved for a class of linearly parameterized uncertain nonlinear systems subject to state and input constraints. Two formulations of the adaptive MPC routine are proposed. A minmax approach is first considered. A Lipschitz-based formulation, amenable to real-time computations, is then proposed. A chemical reactor simulation example is presented that demonstrates the effectivenessof the technique.

Keywords: Adaptive control, Robust MPC, Nonlinear MPC

1. INTRODUCTION

Most physical systems possess consists of parametric and non-parametric uncertainties and the system dynamics can be influenced by exogeneous disturbances as well. Examples in chemical engineering include reaction rates, activation energies, fouling factors, and microbial growth rates. Since parametric uncertainty may degrade the performance of MPC, mechanisms to update the unknown or uncertain parameters are desirable in application. One possibility would be to use state measurements to update the model parameters off-line. A more attractive possibility is to apply adaptive extensions of MPC in which parameter estimation and control are performed online. In this paper, we extend an adaptive MPC framework to nonlinear systems with both constant parametric uncertainty and additive exogenous disturbances.

The literature contains very few results on the design of adaptive nonlinear MPC Adetola and Guay (2004); Mayne and Michalska (1993). Existing design techniques are restricted to systems that are linear in the unknown (constant) parameters and do not involve state constraints. Although MPC exhibits some degree of robustness to uncertainties, in reality, the degree of robustness provided by nominal models or certainty equivalent models may not be sufficient in practical applications. Parameter estimation error must be accounted for in the computation of the control law.

This paper is inspired by DeHaan and Guay (2007); DeHaan et al. (2007). While the focus in DeHaan and Guay (2007); DeHaan et al. (2007) is on the use of adaptation to reduce the conservatism of robust MPC controller, this study addresses the problem of adaptive MPC and incorporates robust features to guarantee closed-loop stability and constraint satisfaction. Simplicity is achieved here-in by generating a parameter estimator for the unknown parameter vector and parameterizing the control policy in terms of these estimates rather than adapting a parameter uncertainty set directly.

First, a min-max feedback nonlinear MPC scheme is combined with an adaptation mechanism. The parameter estimation routine are used to update the parameter uncertainty set, at certain time instants, in a manner that guarantees non-expansion of the set leading to a gradual reduction in the conservativeness or computational demands of the algorithms. The min-max formulation explicitly accounts for the effect of future parameter estimation and automatically injects some useful excitation into the closed-loop system to aid in parameter identification.

Second, the technique is extended to a less computationally demanding robust MPC algorithm. The nominal model rather than the unknown bounded system state is controlled, subject to conditions that ensure that given constraints are satisfied for all possible uncertainties. State prediction error bound is determined based on assumed Lipschitz continuity of the model. Using a nominal model prediction, it is impossible to predict the actual future behavior of the parameter estimation error as was possible in the min-max framework. It is shown how the future model improvement over the prediction horizon can be considered by developing a worst-case upper bound on the future parameter estimation error. The conservativeness of the algorithm reduces as the error bound decreases monotonically over time.

The paper is as follows. The problem description is given in section 2. The parameter estimation routine is presented in section 3.Two approaches to robust adpative model predictive control are detailed in section 4. This is followed by a simulation example in section 5 and brief conclusions in section 6.

2. PROBLEM SET-UP

Consider the uncertain nonlinear system

$$\dot{x} = f(x,u) + g(x,u)\theta + \vartheta \triangleq \mathcal{F}(x,u,\theta,\vartheta) \qquad (1)$$

⋆ The authors would like to acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada.

where the disturbance $\vartheta \in \mathcal{D} \subset \mathbb{R}^{n_d}$ is assumed to satisfy a known upper bound $\|\vartheta(t)\| \leq M_\vartheta < \infty$. The objective of the study is to (robustly) stabilize the plant to some target set $\Xi \subset \mathbb{R}^{n_x}$ while satisfying the pointwise constraints $x \in \mathbb{X} \in \mathbb{R}^{n_x}$ and $u \in \mathbb{U} \in \mathbb{R}^{n_u}$. The target set is a compact set, contains the origin and is robustly invariant under no control. It is assumed that θ is uniquely identifiable and lie within an initially known compact set $\Theta^0 = B(\theta_0, z_\theta)$ where θ_0 is a nominal parameter value, z_θ is the radius of the parameter uncertainty set.

3. PARAMETER AND UNCERTAINTY SET ESTIMATION

3.1 Parameter Adaptation

Let the estimator model for (1) be selected as

$$\dot{\hat{x}} = f(x,u) + g(x,u)\hat{\theta} + k_w\, e + w\dot{\hat{\theta}}, \quad k_w > 0 \tag{2}$$

$$\dot{w} = g(x,u) - k_w\, w, \qquad w(t_0) = 0. \tag{3}$$

resulting in state prediction error $e = x - \hat{x}$ and auxiliary variable $\eta = e - w\tilde{\theta}$ dynamics:

$$\dot{e} = g(x,u)\tilde{\theta} - k_w\, e - w\dot{\hat{\theta}} + \vartheta \qquad e(t_0) = x(t_0) - \hat{x}(t_0) \tag{4}$$

$$\dot{\eta} = -k_w\, \eta + \vartheta, \qquad \eta(t_0) = e(t_0). \tag{5}$$

Since ϑ is not known, an estimate of η is generated from

$$\dot{\hat{\eta}} = -k_w\, \hat{\eta}, \qquad \hat{\eta}(t_0) = e(t_0). \tag{6}$$

with resulting estimation error $\tilde{\eta} = \eta - \hat{\eta}$ dynamics

$$\dot{\tilde{\eta}} = -k_w\, \tilde{\eta} + \vartheta, \qquad \tilde{\eta}(t_0) = 0. \tag{7}$$

Let $\Sigma \in \mathbb{R}^{n_\theta \times n_\theta}$ be generated from

$$\dot{\Sigma} = w^T w, \qquad \Sigma(t_0) = \alpha\, I \succ 0, \tag{8}$$

based on equations (2), (3) and (6), thye preferred parameter update law is given by

$$\dot{\Sigma}^{-1} = -\Sigma^{-1} w^T w \Sigma^{-1}, \qquad \Sigma^{-1}(t_0) = \frac{1}{\alpha} I \tag{9a}$$

$$\dot{\hat{\theta}} = \mathrm{Proj}\left\{\gamma\, \Sigma^{-1} w^T (e - \hat{\eta}),\ \hat{\theta}\right\}, \qquad \hat{\theta}(t_0) = \theta^0 \in \Theta^0 \tag{9b}$$

where $\gamma = \gamma^T > 0$ and $\mathrm{Proj}\{\phi,\ \hat{\theta}\}$ denotes a Lipschitz projection operator such that

$$-\mathrm{Proj}\{\phi,\ \hat{\theta}\}^T \tilde{\theta} \leq -\phi^T \tilde{\theta}, \tag{10}$$

$$\hat{\theta}(t_0) \in \Theta^0 \ \Rightarrow\ \hat{\theta}(t) \in \Theta^0_\epsilon,\ \ \forall\, t \geq t_0. \tag{11}$$

where $\Theta^0_\epsilon \triangleq B(\theta^0, z^0_\theta + \epsilon)$, $\epsilon > 0$. More details on parameter projection can be found in Krstic et al. (1995). To proof the following lemma, we need the following result

Lemma 1. Desoer and Vidyasagar (1975) Consider the system

$$\dot{x}(t) = Ax(t) + u(t) \tag{12}$$

Suppose the equilibrium state $x_e = 0$ of the homogeneous equation is exponentially stable,

(1) if $u \in \mathrm{L}_p$ for $1 < p < \infty$, then $x \in \mathrm{L}_p$ and
(2) if $u \in \mathrm{L}_p$ for $p = 1$ or 2, then $x \to 0$ as $t \to \infty$.

Lemma 2. The identifier (9) is such that the estimation error $\tilde{\theta} = \theta - \hat{\theta}$ is bounded. Moreover, if

$$\vartheta \in \mathcal{L}_2 \qquad \text{or} \qquad \int_{t_0}^{\infty} \left[\|\tilde{\eta}\|^2 - \underline{\gamma}\, \|e - \hat{\eta}\|^2\right] d\tau < +\infty \tag{13}$$

with $\underline{\gamma} = \lambda_{\min}(\gamma)$ and the strong condition

$$\lim_{t\to\infty} \lambda_{\min}\big(\Sigma\big) = \infty \tag{14}$$

is satisfied, then $\tilde{\theta}$ converges to zero asymptotically.

Proof: Let $V_{\tilde{\theta}} = \tilde{\theta}^T \Sigma\, \tilde{\theta}$, it follows from (9) and the relationship $w\tilde{\theta} = e - \hat{\eta} - \tilde{\eta}$ that

$$\begin{aligned} \dot{V}_{\tilde{\theta}} &\leq -2\underline{\gamma}\, \tilde{\theta}^T w^T (e - \hat{\eta}) + \tilde{\theta}^T w^T w \tilde{\theta} \\ &= -\underline{\gamma}\, (e - \hat{\eta})^T (e - \hat{\eta}) + \|\tilde{\eta}\|^2, \end{aligned} \tag{15}$$

implying that $\tilde{\theta}$ is bounded. Moreover, it follows from (15) that

$$V_{\tilde{\theta}}(t) = V_{\tilde{\theta}}(t_0) + \int_{t_0}^{t} \dot{V}_{\tilde{\theta}}(\tau) d\tau \tag{16}$$

$$\leq V_{\tilde{\theta}}(t_0) - \underline{\gamma} \int_{t_0}^{t} \|e - \hat{\eta}\|^2\, d\tau + \int_{t_0}^{t} \|\tilde{\eta}\|^2\, d\tau \tag{17}$$

Considering the dynamics of (7), if $\vartheta \in \mathcal{L}_2$, then $\tilde{\eta} \in \mathcal{L}_2$ (Lemma 1). Hence, the right hand side of (17) is finite in view of (13), and by (14) we have
$\lim_{t\to\infty} \tilde{\theta}(t) = 0$ ■

3.2 Set Adaptation

An update law that measures the worst-case progress of the parameter identifier in the presence of disturbance is given by:

$$z_\theta = \sqrt{\frac{V_{z\theta}}{\lambda_{\min}(\Sigma)}} \tag{18a}$$

$$V_{z\theta}(t_0) = \lambda_{\max}\Big(\Sigma(t_0)\Big)\, (z^0_\theta)^2 \tag{18b}$$

$$\dot{V}_{z\theta} = -\underline{\gamma}\, (e - \hat{\eta})^T (e - \hat{\eta}) + \left(\frac{M_\vartheta}{k_w}\right)^2. \tag{18c}$$

Using the parameter estimator (9) and its error bound z_θ (18), the uncertain ball $\Theta \triangleq B(\hat{\theta}, z_\theta)$ is adapted online according to the following algorithm:

Algorithm 1. Beginning from time $t_{i-1} = t_0$, the parameter and set adaptation is implemented iteratively as follows:

1 **Initialize** $z_\theta(t_{i-1}) = z^0_\theta$, $\hat{\theta}(t_{i-1}) = \hat{\theta}^0$ and $\Theta(t_{i-1}) = B(\hat{\theta}(t_{i-1}), z_\theta(t_{i-1}))$.
2 At time t_i, using equations (9) and (18) **perform** the update

$$\Big(\hat{\theta},\ \Theta\Big) = \begin{cases} \Big(\hat{\theta}(t_i),\ \Theta(t_i)\Big), & \text{if } z_\theta(t_i) \leq z_\theta(t_{i-1}) \\ & \quad -\|\hat{\theta}(t_i) - \hat{\theta}(t_{i-1})\| \\ \Big(\hat{\theta}(t_{i-1}),\ \Theta(t_{i-1})\Big), & \text{otherwise} \end{cases} \tag{19}$$

3 **Iterate back** to step 2, **incrementing** $i = i + 1$.

The algorithm ensure that Θ is only updated when z_θ value has decreased by an amount which guarantees a contraction of the set. Moreover z_θ evolution as given in (18) ensures non-exclusion of θ as shown below.

Lemma 3. The evolution of $\Theta = B(\hat{\theta}, z_\theta)$ under (9), (18) and algorithm 1 is such that

i) $\Theta(t_2) \subseteq \Theta(t_1),\ t_0 \leq t_1 \leq t_2$
ii) $\theta \in \Theta(t_0) \Rightarrow \theta \in \Theta(t),\ \forall t \geq t_0$

Proof:

i) If $\Theta(t_{i+1}) \nsubseteq \Theta(t_i)$, then

$$\sup_{s \in \Theta(t_{i+1})} \|s - \hat{\theta}(t_i)\| \geq z_\theta(t_i). \tag{20}$$

However, it follows from triangle inequality and algorithm 1 that Θ, at update times, obeys

$$\begin{aligned}\sup_{s \in \Theta(t_{i+1})} &\|s - \hat{\theta}(t_i)\| \\ &\leq \sup_{s \in \Theta(t_{i+1})} \|s - \hat{\theta}(t_{i+1})\| + \|\hat{\theta}(t_{i+1}) - \hat{\theta}(t_i)\| \\ &\leq z_\theta(t_{i+1}) + \|\hat{\theta}(t_{i+1}) - \hat{\theta}(t_i)\| \ \leq\ z_\theta(t_i),\end{aligned}$$

which contradicts (20). Hence, Θ update guarantees $\Theta(t_{i+1}) \subseteq \Theta(t_i)$ and the strict contraction claim follows from the fact that Θ is held constant over update intervals $\tau \in (t_i, t_{i+1})$.

ii) We know that $V_{\tilde{\theta}}(t_0) \leq V_{z\theta}(t_0)$ (by definition) and it follows from (15) and (18c) that $\dot{V}_{\tilde{\theta}}(t) \leq \dot{V}_{z_\theta}(t)$. Hence, by the comparison lemma, we have

$$V_{\tilde{\theta}}(t) \ \leq\ V_{z\theta}(t), \quad \forall t \geq t_0. \tag{21}$$

and since $V_{\tilde{\theta}} = \tilde{\theta}^T \Sigma\, \tilde{\theta}$, it follows that

$$\|\tilde{\theta}(t)\|^2 \leq \frac{V_{z\theta}(t)}{\lambda_{\min}(\Sigma(t))} = z_\theta^2(t), \quad \forall t \geq t_0. \tag{22}$$

Hence, if $\theta \in \Theta(t_0)$, then $\theta \in B(\hat{\theta}(t), z_\theta(t)),\ \forall t \geq t_0$.

■

4. ROBUST ADAPTIVE MPC

4.1 A Min-max Approach

The formulation of the min-max MPC consists of maximizing a cost function with respect to $\theta \in \Theta$, $\vartheta \in \mathcal{D}$ and minimizing over feedback control policies κ. The robust receding horizon control law is

$$u = \kappa_{mpc}(x, \hat{\theta}, z_\theta) \triangleq \kappa^*(0, x, \hat{\theta}, z_\theta) \tag{23a}$$

$$\kappa^* \triangleq \arg\min_{\kappa(\cdot,\cdot,\cdot,\cdot)} J(x, \hat{\theta}, z_\theta, \kappa) \tag{23b}$$

where

$$J(x, \hat{\theta}, z_\theta, \kappa) \triangleq \max_{\theta \in \Theta,\ \vartheta \in \mathcal{D}} \int_0^T L(x^p, u^p) d\tau + W(x^p(T), \tilde{\theta}^p(T)) \tag{24a}$$

s.t. $\forall \tau \in [0,\ T]$

$$\dot{x}^p = f(x^p, u^p) + g(x^p, u^p)\theta + \vartheta,\ \ x^p(0) = x \tag{24b}$$

$$\dot{w^p} = g^T(x^p, u^p) - k_w\, w^p,\ \ w^p(0) = w \tag{24c}$$

$$(\dot{\Sigma}^{-1})^p = -(\Sigma^{-1})^p w^T w (\Sigma^{-1})^p, \quad (\Sigma^{-1})^p(0) = \Sigma^{-1} \tag{24d}$$

$$\dot{\hat{\theta}}^p = \text{Proj}\left\{\gamma\,(\Sigma^{-1})^p w^T (e - \hat{\eta}),\ \hat{\theta}\right\}, \quad \tilde{\theta}^p = \theta - \hat{\theta}^p,\ \ \hat{\theta}^p(0) = \hat{\theta} \tag{24e}$$

$$u^p(\tau) \triangleq \kappa(\tau, x^p(\tau), \hat{\theta}^p(\tau)) \in \mathbb{U} \tag{24f}$$

$$x^p(\tau) \in \mathbb{X},\ \ x^p(T) \in \mathbb{X}_f(\tilde{\theta}^p(T)) \tag{24g}$$

The effect of future parameter adaptation is also accounted for in this formulation. The conservativeness of the algorithm is reduced by parameterizing both W and $\mathbb{X}_f$ as functions of $\tilde{\theta}(T)$. While it is possible for the set Θ to contract upon θ over time, the robustness feature due to $\vartheta \in \mathcal{D}$ will still remain.

Algorithm 2. The MPC algorithm performs as follows: At sampling instant t_i

(1) **Measure** the current state of the plant $x(t)$ and obtain the current value of matrices w and Σ^{-1} from equations (3) and (9a) respectively
(2) Obtain the current value of parameter estimates $\hat{\theta}$ and uncertainty bound z_θ from (9b) and (18) respectively
If $z_\theta(t_i) \leq z_\theta(t_{i-1}) - \|\hat{\theta}(t_i) - \hat{\theta}(t_{i-1})\|$
$$\hat{\theta} = \hat{\theta}(t_i), \quad z_\theta = z_\theta(t_i)$$
,
Else
$$\hat{\theta} = \hat{\theta}(t_{i-1}), \quad z_\theta = z_\theta(t_{i-1})$$
End
(3) **Solve** the optimization problem (23) and apply the resulting feedback control law to the plant until the next sampling instant
(4) **Increment** $i = i+1$. Repeat the procedure from step 1 for the next sampling instant.

4.2 Lipschitz-based Approach

In this section, we present a Lipschitz-based method whereby the nominal model rather than the unknown bounded system state is controlled, subject to conditions that ensure that given constraints are satisfied for all possible uncertainties. State prediction error bound is determined based on the Lipschitz continuity of the model. A knowledge of appropriate Lipschitz bounds for the x-dependence of the dynamics $f(x,u)$ and $g(x,u)$ are assumed as follows:

Assumption 4. A set of functions $\mathcal{L}_j : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^+$, $j \in \{f, g\}$ are known which satisfy

$$\mathcal{L}_j(\mathbb{X}, u) \geq \min\left\{\mathcal{L}_j \ \Big|\ \sup_{x_1, x_2 \in \mathbb{X}} \Big(\|j(x_1,u) - j(x_2,u)\| - \mathcal{L}_j \|x_1 - x_2\|\Big) \leq 0\right\},$$

where for $j \equiv g$ is interpreted as an induced norm since $g(x,u)$ is a matrix.

Assuming a knowledge of the Lipschitz bounds for the x-dependence of the dynamics $f(x,u)$ and $g(x,u)$ as given in Assumption 4 and let $\Pi = z_\theta + \|\hat{\theta}\|$, a worst-case deviation $z_x^p \geq \max_{\theta \in \Theta} \|x - x^p\|$ can be generated from

$$\dot{z}_x^p = (\mathcal{L}_f + \mathcal{L}_g \Pi) z_x^p + \|g(x^p, u)\| z_\theta + M_\vartheta, \quad z_x^p(t_0) = 0. \tag{26}$$

Using this error bound, the robust Lipschitz-based MPC is given by

$$u = \kappa_{mpc}(x, \hat{\theta}, z_\theta) = u^*(0) \tag{27a}$$

$$u^*(.) \triangleq \arg\min_{u^p_{[0,T]}} J(x, \hat{\theta}, z_\theta, u^p) \tag{27b}$$

where

$$J(x,\hat{\theta},z_\theta,u^p)= \int_0^T L(x^p,u^p)d\tau + W(x^p(T),z_\theta^p) \tag{28a}$$

$$s.t.\ \forall\tau\in[0,\,T]$$

$$\dot{x}^p = f(x^p,u^p)+g(x^p,u^p)\hat{\theta},\quad x^p(0)=x \tag{28b}$$

$$\dot{z}_x^p = (\mathcal{L}_f+\mathcal{L}_g\Pi)z_x^p + \|g^p\|z_\theta + M_\vartheta,\ z_x^p(0)=0 \tag{28c}$$

$$X^p(\tau)\triangleq B(x^p(\tau),z_x^p(\tau))\subseteq\mathbb{X},\quad u^p(\tau)\in\mathbb{U} \tag{28d}$$

$$X^p(T)\subseteq\mathbb{X}_f(z_\theta^p) \tag{28e}$$

The effect of the disturbance is built into the uncertainty cone $B(x^p(\tau),z_x^p(\tau))$ via (28c). Since the uncertainty bound is no more monotonically decreasing in this case, the uncertainty radius z_θ which appears in (28c) and in the terminal expressions of (28a) and (28e) are held constant over the prediction horizon. However, the fact that they are updated at sampling instants when z_θ shrinks reduces the conservatism of the robust MPC and enlarges the terminal domain that would otherwise have been designed based on a large initial uncertainty $z_\theta(t_0)$.

Algorithm 3. The Lipschitz-based MPC algorithm performs as follows: At sampling instant t_i

(1) **Measure** the current state of the plant $x=x(t_i)$
(2) **Obtain** the current value of the parameter estimates $\hat{\theta}$ and uncertainty bound z_θ from equations (9) and (18) respectively,
If $z_\theta(t_i)\le z_\theta(t_{i-1})$
$$\hat{\theta}=\hat{\theta}(t_i),\quad z_\theta=z_\theta(t_i)$$
Else
$$\hat{\theta}=\hat{\theta}(t_{i-1}),\quad z_\theta=z_\theta(t_{i-1})$$
End
(3) **Solve** the optimization problem (27) and apply the resulting feedback control law to the plant until the next sampling instant
(4) **Increment** i:=i+1; repeat the procedure from step 1 for the next sampling instant.

5. CLOSED-LOOP ROBUST STABILITY

Robust stabilization to the target set Ξ is guaranteed by appropriate selection of the design parameters W and X_f. The robust stability conditions require the satisfaction of the following criteria.

Criterion 5. The terminal penalty function $W:\mathbb{X}_f\times\tilde{\Theta}^0\to[0,+\infty]$ and the terminal constraint function $\mathbb{X}_f:\tilde{\Theta}^0\to\mathbb{X}$ are such that for each $(\theta,\,\hat{\theta},\,\tilde{\theta})\in(\Theta^0\times\Theta^0\times\tilde{\Theta}^0_\epsilon)$, there exists a feedback $k_f(.,\hat{\theta}):\mathbb{X}_f\to\mathbb{U}$ satisfying

(1) $0\in\mathbb{X}_f(\tilde{\theta})\subseteq\mathbb{X},\quad\mathbb{X}_f(\tilde{\theta})$ closed
(2) $k_f(x,\hat{\theta})\in\mathbb{U},\ \forall x\in\mathbb{X}_f(\tilde{\theta})$
(3) $W(x,\tilde{\theta})$ is continuous with respect to $x\in\mathbb{R}^{n_x}$
(4) $\forall\,x\in\mathbb{X}_f(\tilde{\theta})\backslash\Xi$, $\mathbb{X}_f(\tilde{\theta})$ is strongly positively invariant under $k_f(x,\hat{\theta})$ with respect to $\dot{x}\in f(x,k_f(x,\hat{\theta}))+g(x,k_f(x,\hat{\theta}))\Theta+\mathcal{D}$
(5) $L(x,k_f(x,\hat{\theta}))+\frac{\partial W}{\partial x}\mathcal{F}(x,k_f(x,\hat{\theta}),\theta,\vartheta)\le 0,\ \forall x\in\mathbb{X}_f(\tilde{\theta})\backslash\Xi$.

Criterion 6. For any $\tilde{\theta}_1,\ \tilde{\theta}_2\in\tilde{\Theta}^0$ *s.t.* $\|\tilde{\theta}_2\|\le\|\tilde{\theta}_1\|$,

(1) $W(x,\tilde{\theta}_2)\le W(x,\tilde{\theta}_1),\ \forall x\in\mathbb{X}_f(\tilde{\theta}_1)$
(2) $\mathbb{X}_f(\tilde{\theta}_2)\supseteq\mathbb{X}_f(\tilde{\theta}_1)$

The revised condition C5 require W to be a local robust CLF for the uncertain system 1 with respect to $\theta\in\Theta$ and $\vartheta\in\mathcal{D}$.

5.1 Main Results

Theorem 7. Let $X_{d0}\triangleq X_{d0}(\Theta^0)\subseteq\mathbb{X}$ denote the set of initial states with uncertainty Θ^0 for which (23) has a solution. Assuming criteria 5 and 6 are satisfied, then the closed-loop system state x, given by (1,9,18,23), originating from any $x_0\in X_{d0}$ feasibly approaches the target set Ξ as $t\to+\infty$.

Proof: *Feasibility*: The closed-loop stability is based upon the feasibility of the control action at each sample time. Assuming, at time t, that an optimal solution $u^p_{[0,T]}$ to the optimization problem (23) exist and is found. Let Θ^p denote the estimated uncertainty set at time t and Θ^v denote the set at time $t+\delta$ that would result with the feedback implementation of $u_{[t,t+\delta]}=u^p_{[0,\delta]}$. Also, let x^p represents the worst case state trajectory originating from $x^p(0)=x(t)$ and x^v represents the trajectory originating from $x^v(0)=x+\delta v$ under the same feasible control input $u^v_{[\delta,T]}=u^p_{[\delta,T]}$. Moreover, let $X^a_{\Theta^b}\triangleq\{x^a|\,\dot{x}^a\in\mathcal{F}(x^a,u^p,\Theta^b)\triangleq f(x^a,u^p)+g(x^a,u^p)\Theta^b\}$.

Since the $u^p_{[0,T]}$ is optimal with respect to the worst case uncertainty scenario, it suffice to say that $u^p_{[0,T]}$ drives any trajectory $x^p\in X^p_{\Theta^p}$ into the terminal region $\mathbb{X}^p_f$. Since Θ is non-expanding over time, we have $\Theta^v\subseteq\Theta^p$ implying $x^v\in X^p_{\Theta^v}\subseteq X^p_{\Theta^p}$. The terminal region $\mathbb{X}^p_f$ is strongly positively invariant for the nonlinear system (1) under the feedback $k_f(.,.)$, the input constraint is satisfied in $\mathbb{X}^p_f$ and $\mathbb{X}^v_f\supseteq\mathbb{X}^p_f$ by criteria 2.2, 2.4 and 3.2 respectively. Hence, the input $u=[u^p_{[\delta,T]},k_{f[T,T+\delta]}]$ is a feasible solution of (23) at time $t+\delta$ and by induction, the optimization problem is feasible for all $t\ge0$.

Stability: The stability of the closed-loop system is established by proving strict decrease of the optimal cost $J^*(x,\hat{\theta},z_\theta)\triangleq J(x,\hat{\theta},z_\theta,\kappa^*)$. Let the trajectories $(x^p,\,\hat{\theta}^p,\,\tilde{\theta}^p,z_\theta^p)$ and control u^p correspond to any worst case minimizing solution of $J^*(x,\hat{\theta},z_\theta)$. If $x^p_{[0,T]}$ were extended to $\tau\in[0,T+\delta]$ by implementing the feedback $u(\tau)=k_f(x^p(\tau),\hat{\theta}^p(\tau))$ on $\tau\in[T,T+\delta]$, then criterion 5(5) guarantees the inequality

$$\int_T^{T+\delta}L(x^p,k_f(x^p,\hat{\theta}^p))d\tau+W(x^p_{T+\delta},\tilde{\theta}^p_T)-W(x^p_T,\tilde{\theta}^p_T)\le0 \tag{29}$$

where in (29) and in the remainder of the proof, $x^p_\sigma\triangleq x^p(\sigma)$, $\tilde{\theta}^p_\sigma\triangleq\tilde{\theta}^p(\sigma)$, for $\sigma=T,\,T+\delta$.
The optimal cost

$$J^*(x,\hat{\theta},z_\theta)=\int_0^T L(x^p,u^p)d\tau+W(x^p_T,\tilde{\theta}^p_T)$$

$$\ge\int_0^T L(x^p,u^p)d\tau+W(x^p_T,\tilde{\theta}^p_T) \tag{30}$$

$$+\int_T^{T+\delta} L(x^p, k_f(x^p,\hat{\theta}^p))d\tau + W(x^p_{T+\delta},\tilde{\theta}^p_T) - W(x^p_T,\tilde{\theta}^p_T) \tag{31}$$

$$\geq \int_0^{\delta} L(x^p,u^p)d\tau + \int_{\delta}^{T} L(x^p,u^p)d\tau \tag{32}$$

$$+\int_T^{T+\delta} L(x^p, k_f(x^p,\hat{\theta}^p))d\tau + W(x^p_{T+\delta},\tilde{\theta}^p_{T+\delta}) \tag{33}$$

$$\geq \int_0^{\delta} L(x^p,u^p)d\tau + J^*(x(\delta),\hat{\theta}(\delta),z_\theta(\delta)) \tag{34}$$

Then, it follows from (34) that

$$J^*(x(\delta),\hat{\theta}(\delta),z_\theta(\delta)) - J^*(x,\hat{\theta},z_\theta) \leq -\int_0^{\delta} L(x^p,u^p)d\tau$$
$$\leq -\int_0^{\delta} \mu_L(\|x,u\|)d\tau. \tag{35}$$

where μ_L is a class $\mathcal{K}_\infty$ function. Hence $x(t) \to 0$ asymptotically.

Remark 8. In the above proof,

- (31) is obtained using inequality (29)
- (33) follows from criterion 5.1 and the fact that $\|\tilde{\theta}\|$ is non-increasing
- (34) follows by noting that the last 3 terms in (33) is a (potentially) suboptimal cost on the interval $[\delta, T+\delta]$ starting from the point $(x^p(\delta),\hat{\theta}^p(\delta))$ with associated uncertainty set $B(\hat{\theta}^p(\delta), z^p_\theta(\delta))$.

The closed-loop stability is established by the feasibility of the control action at each sample time and the strict decrease of the optimal cost J^*. The proof follows from the fact that the control law is optimal with respect to the worst case uncertainty $(\theta,\ \vartheta) \in (\Theta,\ \mathcal{D})$ scenario and the terminal region $\mathbb{X}^p_f$ is strongly positively invariant for (1) under the (local) feedback $k_f(.,.)$. ■

Theorem 9. Let $X'_{d0} \triangleq X'_{d0}(\Theta^0) \subseteq \mathbb{X}$ denote the set of initial states for which (27) has a solution. Assuming Assumption 4 and Criteria 5 and 6 are satisfied, then the origin of the closed-loop system given by (1,9,18,27) is feasibly asymptotically stabilized from any $x_0 \in X'_{d0}$ to the target set Ξ.

The proof of the Lipschitz-based control law follows from that of theorem 7.

6. SIMULATION EXAMPLE

To illustrate the effectiveness of the proposed design, we consider the regulation of the CSTR subject to an additional disturbance on the temperature dynamic:

$$\dot{C}_A = \frac{q}{V}(C_{Ain} - C_A) - k_0 \exp\left(\frac{-E}{R\,T_r}\right) C_A$$
$$\dot{T}_r = \frac{q}{V}(T_{in} - T_r) - \frac{\Delta H}{\rho\, c_p}\, k_0 \exp\left(\frac{-E}{R\,T_r}\right) C_A$$
$$+ \frac{UA}{\rho\, c_p\, V}(T_c - T_r) + \vartheta$$

where $\vartheta(t)$ is an unknown function of time. We also assume that the reaction kinetic constant k_0 and ΔH are only nominally known.

It is assumed that reaction kinetic constant k_0 and heat of reaction ΔH are only nominally known and parameterized as $k_0 = \theta_1 \times 10^{10}\,\mathrm{min}^{-1}$ and $\Delta H\, k_0 = -\theta_2 \times 10^{15}$ J/mol min with the parameters satisfying $0.1 \leq \theta_1 \leq 10$ and $0.1 \leq \theta_2 \leq 10$. The objective is to adaptively regulate the unstable equilibrium $C_A^{eq} = 0.5$ mol/l, $T_r^{eq} = 350$ K, $T_c^{eq} = 300$ K while satisfying the constraints $0 \leq C_A \leq 1$, $280 \leq T_r \leq 370$ and $280 \leq T_c \leq 370$. The nominal operating conditions, which corresponds to the given unstable equilibrium are taken from Magni et al. (2001): $q = 100$ l/min, $V = 100$ l, $\rho = 1000$ g/l, $c_p = 0.239$ J/g K, E/R = 8750 K, $UA = 5\times10^4$ J/min K, $C_{Ain} = 1$ mol/l and $T_{in} = 350$ K.

The control objective is to robustly regulate the reactor temperature and concentration to the (open loop) unstable equilibrium $C_A^{eq} = 0.5$ mol/l, $T_r^{eq} = 350$ K, $T_c^{eq} = 300$ K by manipulating the temperature of the coolant stream T_c.

Defining $x = [\frac{C_A - C_A^{eq}}{0.5},\ \frac{T_r - T_r^{eq}}{20}]'$, $u = \frac{T_c - T_c^{eq}}{20}$, the stage cost $L(x,u)$ was selected as a quadratic function of its arguments:

$$L(x,u) = x^T\, Q_x\, x + u^T\, R_u\, u \tag{36a}$$
$$Q_x = \begin{bmatrix} 0.5 & 0 \\ 0 & 1.1429 \end{bmatrix} \qquad R_u = 1.333. \tag{36b}$$

The terminal penalty function used is a quadratic parameter-dependent Lyapunov function $W(x,\theta) = x^T P(\theta) x$ for the linearized system. Denoting the closed-loop system under a local robust stabilizing controller $u = k_f(\theta)\, x$ as $\dot{x} = A_{cl}(\theta)x$. The matrix $P(\theta) := P_0 + \theta_1 P_1 + \theta_2 P_2 + \ldots\ \theta_{n\theta} P_{n\theta}$ was selected to satisfy the Lyapunov system of LMIs

$$P(\theta) > 0$$
$$A_{cl}(\theta)^T P(\theta) + P(\theta) A_{cl}(\theta) < 0$$

for all admissible values of θ. Since θ lie between known extrema values, the task of finding $P(\theta)$ reduces to solving a finite set of linear matrix inequalities by introducing additional constraints Gahinet et al. (1996). For the initial nominal estimate $\theta^0 = 5.05$ and $z^0_\theta = 4.95$, the matrix $P(\theta^0)$ obtained is

$$P(\theta^0) = \begin{bmatrix} 0.6089 & 0.1134 \\ 0.1134 & 4.9122 \end{bmatrix} \tag{37}$$

and the corresponding terminal region is

$$\mathbb{X}_f = \{x\ :\ x^T P(\theta^0)\, x \leq 0.25\}. \tag{38}$$

For simulation purposes, the disturbance is selected as a fluctuation of the inlet temperate $\vartheta(t) = 0.01\, T_{in} \sin(3t)$ and the true values of the unknown parameters were also chosen as $k_0 = 7.2 \times 10^{10}\mathrm{min}^{-1}$ and ΔH = - 5.0 $\times$ 10^4 J/mol. The stage cost (36), terminal penalty (37) and terminal region (38) were used. The Lipschitz-based approach was used for the controller calculations and the result was implemented according to Algorithm 3. As depicted in Figures 1 to 3, the robust adaptive MPC drives the system to a neighborhood of the equilibrium while satisfying the imposed constraints and achieves parameter convergence. Figure 4 shows that the uncertainty bound z_θ also reduces over time.

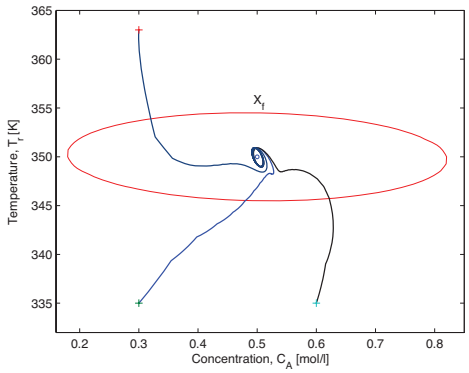


Fig. 1. Closed-loop reactor trajectories under additive disturbance $\vartheta(t)$

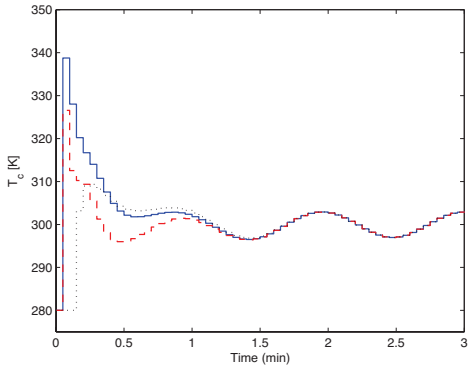


Fig. 2. Closed-loop input profiles for states starting at different initial conditions $(C_A(0),\ T_r(0))$: (0.3, 335) is solid line, (0.6, 335) is dashed line and (0.3, 363) is the dotted line

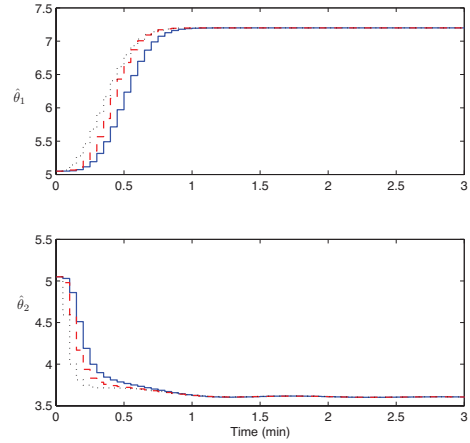


Fig. 3. Closed-loop parameter estimates profile for states starting at different initial conditions $(C_A(0),\ T_r(0))$: (0.3, 335) is solid line, (0.6, 335) is dashed line and (0.3, 363) is the dotted line

7. CONCLUSIONS

The adaptive MPC design technique is extended to constrained nonlinear systems with both parametric and time

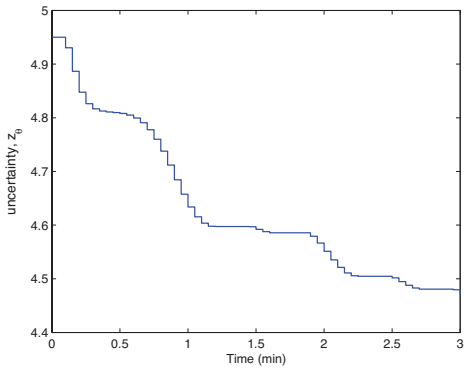


Fig. 4. Closed-loop uncertainty bound trajectories for initial condition $(C_A,\ T_r) = (0.3,\ 335)$

varying disturbances. The proposed robust controller updates the plant model online when model improvement is guaranteed. The embedded adaptation mechanism enables us to construct less conservative terminal design parameters based upon subsets of the original parametric uncertainty. While the introduced conservatism/computation complexity due to the parametric uncertainty reduces over time, the portion due to the disturbance $\vartheta \in \mathcal{D}$ remains active for all time.

REFERENCES

Adetola, V. and Guay, M. (2004). Adaptive receding horizon control of nonlinear systems. In *Proc. of IFAC Symposium on Nonlinear Control systems*, 1055–1060. Stuttgart.

DeHaan, D., Adetola, V., and Guay, M. (2007). Adaptive robust mpc: An eye towards computational simplicity. In *Proc. of IFAC Symposium on Nonlinear Control systems*. South Africa.

DeHaan, D. and Guay, M. (2007). Adaptive robust mpc: A minimally-conservative approach. In *Proc. of American Control Conference*.

Desoer, C. and Vidyasagar, M. (1975). *Feedback Systems: Input-Output Properties*. Academic Press, New York.

Gahinet, P., Apkarian, P., and Chilali, M. (1996). Affine parameter-dependent lyapunov functions and real. parametric uncertainty. *IEEE Transactions on Automatic Control*, 41(3), 436–442.

Krstic, M., Kanellakopoulos, I., and Kokotovic, P. (1995). *Nonlinear and Adaptive Control Design*. John Wiley and Sons Inc, Toronto.

Magni, L., De Nicolao, G., Magnani, L., and Scattolini, R. (2001). A stabilizing model-based predictive control algorithm for nonlinear systems. *Automatica*, 37, 1351–1362.

Mayne, D.Q. and Michalska, H. (1993). Adaptive receding horizon control for constrained nonlinear systems. In *In Proc. of IEEE Conference on Decision and Control*, 1286–1291. San Antonio, Texas.

Applications

Oral Session

Energy Consumption Optimization of RO Membrane Desalination Subject to Feed Salinity Fluctuation ⋆

Aihua Zhu, Panagiotis D. Christofides and Yoram Cohen

Department of Chemical and Biomolecular Engineering
Water Technology Research Center
University of California, Los Angeles, CA 90095-1592, U.S.A.

Abstract: We study the energy consumption optimization of a reverse osmosis water desalination process producing a constant permeate flow in the presence of feed concentration fluctuation. We propose a time-varying optimal operation strategy that can significantly reduce the specific energy consumption compared to time-invariant process operation.

Keywords: specific energy consumption optimization, reverse osmosis, thermodynamic restriction, feed concentration fluctuation

1. INTRODUCTION

Reverse osmosis (RO) membrane water desalination is now well established as a mature water desalination technology. However, energy consumption is a major portion of the total cost of water desalination and can reach as high as about 45% of the total permeate production cost (Manth et al. (2003); Busch and Mickols (2004); Wilf and Bartels (2005)). The energy cost per volume of produced permeate (i.e., the specific energy consumption or SEC) is significant in RO operation due to the high pressure requirement (up to about 1000 psi for seawater and in the range of 100-600 psi for brackish water desalting). Considerable effort, dating back to the initial days of RO development in the early 1960s (as reviewed in Zhu et al. (2008)), has been devoted to minimizing the specific energy consumption of water desalination. The introduction of highly permeable membranes in the mid 1990s with low salt passage (Wilf (1997)) has generated considerable interest (Zhu et al. (2008)), given their potential for reducing the energy required to attain a given permeate flow, since the operating pressure can be greatly reduced to approach the osmotic pressure difference at the exit of a membrane module (Wilf (1997), Song et al. (2003a); Song and Tay (2006)).

In a previous work (Zhu et al. (2008)), we systematically studied the effect of the thermodynamic restriction (i.e., the fact that the applied pressure cannot be lower than the osmotic pressure of the exit brine stream plus pressure losses across the membrane module) on the optimization of the specific energy consumption of an RO process. Specifically, we computed the optimum SEC, corresponding water recovery, and permeate flux for single-stage and two-stage RO membrane desalination systems. We also studied the effect of energy recovery device, membrane cost and brine disposal costs on SEC. The developed approach can also be utilized to evaluate the energy savings of a two-stage RO system over single-stage RO and the impact of extra membrane area consumption of two-stage over single-stage. In a recent work (Zhu et al. (2009)), we carried out a systematic study of the energy consumption of two-pass reverse osmosis membrane desalination accounting for key practical issues like membrane salt rejection, presence/absence of energy recovery devices and concentration polarization. We established that if the salt rejection level of the available membranes can achieve the desired permeate salt content, then a single-pass configuration is more energy favorable than a two-pass configuration for the same level of total water recovery and salt rejection. However, if it is not possible to obtain the desired permeate salt content with the available membranes, then a two-pass configuration has to be used, and in this case, the energy optimal solution is to operate the first-pass using the membranes with the maximum rejection.

In the present work, we extend our previous results to account for the effect of feed salinity fluctuation on energy consumption optimization. Due to seasonal rainfalls, the feed water salinity will fluctuate both for seawater and brackish water. For example, at one location in the central San Joaquin Valley, the total dissolved solid (TDS) deviated up to 52% from its annual average (McCool (2008)). The objective of the present work is to determine the optimal time-varying operating policy to produce a constant permeate flow in the presence of a given feed salinity fluctuation profile.

2. RO PROCESS

2.1 Description and Modeling

In order to illustrate the approach to energy cost optimization it is instructive to consider a membrane RO

⋆ The present work was supported in part by the International Desalination Association (Channabasappa Memorial Scholarship to Aihua Zhu), California Department of Water Resources, the Metropolitan Water District of Southern California, the University of California Water Resources Center, and the United States Environmental Protection Agency. Corresponding author: Panagiotis D. Christofides (e-mail: pdc@seas.ucla.edu)

process without the deployment of an energy recovery device (ERD) as shown schematically in Fig. 1.

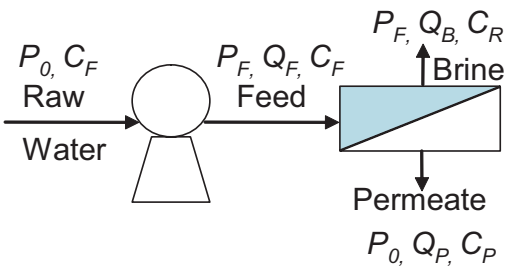


Figure 1. Schematic of simplified RO system.

The energy cost associated with RO desalination is evaluated in the present analysis as the specific energy consumption (SEC) defined as the electrical energy needed to produce a cubic meter of permeate. Pump efficiency can be included in the following analysis in a straightforward fashion as presented in Zhu et al. (2008). As a first step, however, in order to simplify the presentation of the approach, the required electrical energy is taken to be equal to the pump work, (i.e., assuming a pump efficiency of 100%). Accordingly, the SEC for the plant shown in Fig. 1 is given by:

$$SEC = \frac{\dot{W}_{pump}}{Q_p} \tag{1}$$

where Q_p is the permeate flow rate and $\dot{W}_{pump}$ is the rate of work done by the pump, given by:

$$\dot{W}_{pump} = \Delta P \times Q_f \tag{2}$$

in which

$$\Delta P = P_f - P_0 \tag{3}$$

where P_f is the water pressure at the entrance of the membrane module, P_0 is the pressure of the raw water which is assumed (for simplicity) to be the same as the permeate pressure, and Q_f is the volumetric feed flow rate. In order to simplify the analysis, we initially assume that the impact of the pressure drop (within the RO module) on locating the minimum SEC is negligible; this issue is addressed further in Zhu et al. (2008). It is acknowledged that, fouling and scaling will impact the selection of practical RO process operating conditions and feed pretreatment. However, the inclusion of such effects is beyond the scope of the present paper.

The permeate product water recovery for the RO process, Y, is an important measure of the process productivity, defined as:

$$Y = \frac{Q_p}{Q_f} \tag{4}$$

and combining Eqs. (1), (2) and (4), the SEC can be rewritten as follows:

$$SEC = \frac{\Delta P}{Y} \tag{5}$$

The permeate flow rate can be approximated by the classical reverse osmosis flux equation Mulder (1997):

$$Q_p = A_m L_p(\Delta P - \sigma\overline{\Delta\pi}) = A_m L_p(\overline{NDP}) \tag{6}$$

where A_m is the active membrane area, L_p is the membrane hydraulic permeability, σ is the reflection coefficient (typically assumed to be about unity for high rejection RO membranes and in this study $\sigma = 1$), ΔP is the transmembrane pressure, $\overline{\Delta\pi}$ is the average osmotic pressure difference between the retentate and permeate stream along the membrane module, $(\Delta P - \sigma\overline{\Delta\pi})$ is the average trans-membrane net driving pressure designated as $\overline{NDP}$. We also invoke the typical approximation in Mulder (1997) that the osmotic pressure varies linearly with concentration (i.e., $\pi = f_{os}C$ where f_{os} is the osmotic pressure coefficient and C is the solution salt concentration). For the purpose of the present analysis and motivated by our focus on RO processes that utilize highly permeable membranes, the average osmotic pressure difference (up to the desired level of product water recovery), $\overline{\Delta\pi}$, can be approximated as the log-mean average along the membrane (ASTM (2000)) as confirmed in a previous work Zhu et al. (2008),

$$\overline{\Delta\pi} = f_{os}C_f\frac{ln(\frac{1}{1-Y})}{Y} \tag{7}$$

where C_f is the salt concentration of the feed to the membrane module. The osmotic pressures at the entrance and the exit of the membrane module, relative to the permeate stream, are approximate by:

$$\Delta\pi_{entrance} = f_{os}C_f - \pi_p \tag{8}$$

$$\Delta\pi_{exit} = f_{os}C_r - \pi_p \tag{9}$$

where C_r is the salt concentration of the exit brine (i.e., concentrate) stream. For sufficiently high rejection level, the osmotic pressure of the permeate can be taken to be negligible relative to the feed or concentrate streams and C_r can be approximated by:

$$C_r = \frac{C_f}{1-Y} \tag{10}$$

Thus, by combining Eqs. (8)–(10), the osmotic pressure difference between the retentate and permeate stream at the exit of the module can be expressed as:

$$\Delta\pi_{exit} = \frac{\pi_0}{1-Y} \tag{11}$$

where $\pi_0 = f_{os}C_f$ is the feed osmotic pressure. Eq. (11) is a simple relationship that illustrates that the well known inherent difficulty in reaching high recovery in RO desalting is due to the rapid rise in osmotic pressure with increased recovery.

2.2 Thermodynamic Restriction of Cross-flow RO Operation

In the process of RO desalting, an external pressure is applied to overcome the osmotic pressure, and pure water is recovered from the feed solution through the use of a semipermeable membrane. Assuming that the permeate pressure is the same as the raw water pressure, P_0, the applied pressure (ΔP) needed to obtain a water recovery of Y should be no less than the osmotic pressure difference at the exit region (Wilf (1997); Song et al. (2003b)), which is given by Eq. (11). Therefore, in order to ensure permeate

productivity along the entire RO module (or stage), the following lower bound is imposed on the applied pressure:

$$\Delta P \geq \Delta\pi_{exit} = \frac{\pi_0}{1-Y} \tag{12}$$

This is the so-called thermodynamic restriction of cross-flow RO (Song et al. (2003a); Song and Tay (2006)) and referred to as the "thermodynamic restriction" in the current work. The equality on the right-hand-side of Eq. (12) is the condition at the "limit of thermodynamic restriction" in the exit of the membrane module and is attained at the limit of infinite membrane permeability for a finite membrane area. It is particularly important from a practical point of view when a highly-permeable membrane is used for water desalination at low pressures. It is emphasized that the constraint of Eq. (12) arises when one wants to ensure that the entire membrane area is utilized for permeate production.

The specific energy consumption (SEC) for the RO desalting process can be derived by combining Eqs. (1)–(4) and (12), to obtain:

$$SEC \geq \frac{\pi_0}{Y(1-Y)} \tag{13}$$

where SEC is in pressure units. It is convenient to normalize the SEC, at the limit of thermodynamic restriction (i.e., operation up to the point in which the applied pressure equals the osmotic pressure difference between the concentrate and permeate at the exit of the membrane module), with respect to the feed osmotic pressure such that:

$$SEC_{tr,norm} = \frac{SEC_{tr}}{\pi_0} = \frac{1}{Y(1-Y)} \tag{14}$$

and this dependence is plotted in Fig. 2 showing that there is a global minimum. In order to obtain the analytical global minimum $SEC_{tr,norm}$, with respect to the water recovery, one can set $(dSEC_{tr,norm})/(dY) = 0$ from which it can be shown that the minimum $SEC_{tr,norm}$ occurs at a fractional recovery of $Y = 0.5$ (or 50%) where $(SEC_{tr,norm})_{min} = 4$ (i.e., four times the feed osmotic pressure). The above condition, i.e., $(SEC_{tr,norm})_{min} = 4$ at $Y = 0.5$, represents the global minimum SEC (represented by the equality in Eq. 13). In order to achieve this global minimum energy cost, the RO process should be operated at a water recovery of 50% with an applied pressure equivalent to $2\pi_0$ (i.e., double that the feed osmotic pressure).

2.3 Feed Salinity Fluctuation

For the purpose of illustration of the proposed optimal operation approach, we consider a simple feed salinity fluctuation profile shown in Fig. 3. Specifically, we consider a 20-hour time window in which the feed osmotic pressure in the first 10 hours is 500 *psi*, and it is then reduced to 200 *psi* for the remaining 10 hours. For a single-stage RO system with constant feed flow rate Q_f, the average feed osmotic pressure is 350 *psi*. We will study the minimum specific energy consumption (SEC) of two difference cases. In case 1, the operating pressure is a constant, while in case 2, it will change with the instantaneous feed osmotic pressure and will always be double that of the instantaneous feed osmotic pressure. Both cases are operated at the limit of thermodynamic restriction.

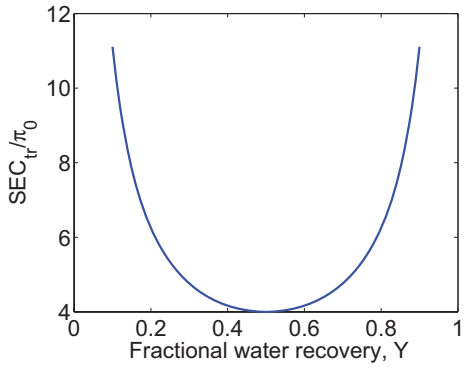


Figure 2. Variation of the normalized SEC with water recovery for a single-stage RO at the limit of thermodynamic restriction.

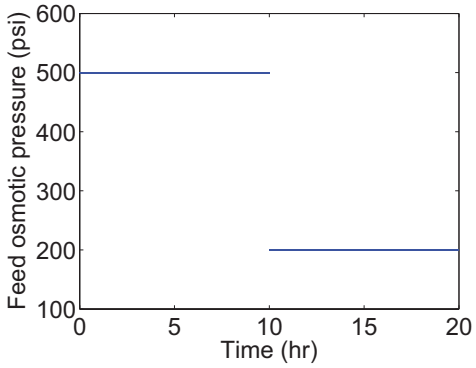


Figure 3. Feed osmotic pressure profile within 20 hours.

3. RESULTS

In the presence of the feed salinity fluctuation of Fig. 3,the following two operating strategies may be considered.

- Operating strategy A: The transmembrane pressure is maintained at double that of the average (over the whole 20-hour time window) feed osmotic pressure, i.e. 700 *psi*.
- Operating strategy B: The transmembrane pressure is maintained at double that of the instantaneous feed osmotic pressure.

For a built plant to produce the same amount of permeate volume for both operating strategy A and operating strategy B, the permeate flow rates in the first 10 hrs and the last 10 hrs have to be the same. The specific energy consumption (SEC) comparison of operating strategy A and operating strategy B will be first done for an RO process without an energy recovery device (see Fig. 1) and the case of an RO process with an energy recovery device (see Fig. 4) will be then addressed. In Fig. 4, P_e and P_p are the brine discharge and permeate pressure, respectively, which are assumed here to be equal to P_0.

The rate of work done by the pump on the raw water, in the presence of an ERD, is given by:

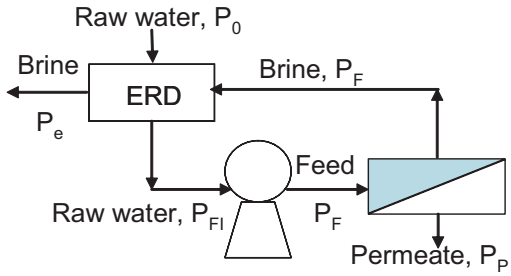


Figure 4. Simplified RO system with an energy recovery device (ERD).

$$\dot{W}_{pump} = \Delta P \times (Q_f - \eta Q_b) \tag{15}$$

where η is the efficiency of the energy recovery device.

3.1 RO Process without ERD

Operating strategy A. At the limit of thermodynamic restriction, according to Eq. 11, the water recovery in the first 10 hrs, $Y_1 = 1 - \frac{500}{700} = \frac{2}{7}$ and the water recovery in the last 10 hrs, $Y_2 = 1 - \frac{200}{700} = \frac{5}{7}$. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be 2.5 times that of the feed flow rate in the last 10 hrs ($Q_{f,2}$). Therefore, the permeate flow and energy consumption in the first and last 10 hrs are:

$$V_{p,1} = 2.5 \times Q_{f,2} \times \frac{2}{7} \times 10\,hr = \frac{50}{7} Q_{f,2} \times hr \tag{16}$$

$$W_1 = \Delta P_1 \times V_{f,1} = 17500\, Q_{f,2} \cdot psi \cdot hr \tag{17}$$

$$V_{p,2} = Q_{f,2} \times \frac{5}{7} \times 10\,hr = \frac{50}{7} Q_{f,2} \times hr \tag{18}$$

$$W_2 = \Delta P_2 \times V_{p,2} = 7000\, Q_{f,2} \cdot psi \cdot hr \tag{19}$$

Therefore, the average SEC is:

$$\overline{SEC}^A = \frac{W_1 + W_2}{V_{p,1} + V_{p,2}} = 1715\, psi \tag{20}$$

Operating strategy B. The water recovery in the last 10 hrs is the same as the water recovery in the first 10 hrs (both at 50%). In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs should be the same as the feed flow rate in the last 10 hrs ($Q'_{f,2}$). The permeate flow and energy consumption in the first and last 10 hrs are:

$$V'_{p,1} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \tag{21}$$

$$W'_1 = \Delta P'_1 \times V'_{f,1} = 10000\, Q'_{f,2} \cdot psi \cdot hr \tag{22}$$

$$V'_{p,2} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \tag{23}$$

$$W'_2 = \Delta P'_2 \times V'_{f,2} = 4000\, Q'_{f,2} \cdot psi \cdot hr \tag{24}$$

Therefore, the average SEC is:

$$\overline{SEC}^B = \frac{W'_1 + W'_2}{V'_{p,1} + V'_{p,2}} = 1400\, psi \tag{25}$$

From Eq. 20 and Eq. 25, we see that the operating strategy A has a higher SEC than operating strategy B about 22.5% ($\frac{1715-1400}{1400} = 22.5\%$). Furthermore, in order to equate the total permeate volume in operating strategy A and operating strategy B, $Q'_{f,2} = \frac{10}{7} Q_{f,2}$. Thus, the total feed volume in operating strategy B is $2 \times \frac{10}{7} Q_{f,2} = \frac{20}{7} Q_{f,2}$, while the total feed volume in operating strategy A is $(2.5 + 1) Q_{f,2} = 3.5 Q_{f,2}$. Therefore, in order to get the same amount of permeate volume, operating strategy A requires a higher amount of feed water, and thus, it has a lower overall water recovery.

3.2 RO Process with ERD: Efficiency is 100%

Operating strategy A. The water recovery in the last 10 hrs is 2.5 times that of the water recovery in the first 10 hrs. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be 2.5 times that of the feed flow rate in the last 10 hrs ($Q_{f,2}$). Therefore, the permeate flow and energy consumption in the first and last 10 hrs are:

$$V_{p,1} = 2.5 \times Q_{f,2} \times \frac{2}{7} \times 10\,hr = \frac{50}{7} Q_{f,2} \times hr \tag{26}$$

$$W_1^{ERD} = \Delta P_1 \times V_{p,1} = 5000\, Q_{f,2} \cdot psi \cdot hr \tag{27}$$

$$V_{p,2} = Q_{f,2} \times \frac{5}{7} \times 10\,hr = \frac{50}{7} Q_{f,2} \times hr \tag{28}$$

$$W_2^{ERD} = \Delta P_2 \times V_{p,2} = 5000\, Q_{f,2} \cdot psi \cdot hr \tag{29}$$

Therefore, the average SEC is:

$$\overline{SEC}^A = \frac{W_1^{ERD} + W_2^{ERD}}{V_{p,1} + V_{p,2}} = 700\, psi \tag{30}$$

Operating strategy B. The water recovery in the last 10 hrs is the same as the water recovery in the first 10 hrs. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be the same as that the feed flow rate in the last 10 hrs ($Q_{f,2}$'). The permeate flow and energy consumption in the first and last 10 hrs are:

$$V'_{p,1} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \tag{31}$$

$$W_1^{'ERD} = \Delta P'_1 \times V'_{p,1} = 5000\, Q'_{f,2} \cdot psi \cdot hr \tag{32}$$

$$V'_{p,2} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \tag{33}$$

$$W_2^{'ERD} = \Delta P_2 \times V'_{p,2} = 2000\, Q_{f,2} \cdot psi \cdot hr \tag{34}$$

Therefore, the average SEC is:

$$\overline{SEC}^B = \frac{W_1 + W_2}{V'_{p,1} + V'_{p,2}} = 700\, psi \tag{35}$$

From Eq. 30 and Eq. 35, we see that in the presence of an ERD with a 100% efficiency, operating strategy A and operating strategy B have the same SEC. Furthermore, in order to equate the total permeate volume in operating strategy A and operating strategy B, $Q'_{f,2} = \frac{10}{7} Q_{f,2}$. Thus, the total feed volume in operating strategy B is $2 \times \frac{10}{7} Q_{f,2} = \frac{20}{7} Q_{f,2}$, while the total feed volume in operating strategy A is $(2.5 + 1) Q_{f,2} = 3.5 Q_{f,2}$. Therefore, in order to get the same amount of permeate volume, operating strategy A requires a higher amount of feed water, and thus, it has a lower overall water recovery.

3.3 ERD Efficiency between 0 and 1

Operating strategy A. The water recovery in the last 10 hrs is 2.5 times that of the water recovery in the first 10 hrs. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be 2.5 times the feed flow rate in the last 10 hrs ($Q_{f,2}$). Therefore, the permeate flow and energy consumption in the first and last 10 hrs are:

$$V_{p,1} = 2.5 \times Q_{f,2} \times \frac{2}{7} \times 10\,hr = \frac{50}{7} Q_{f,2} \times hr \quad (36)$$

$$W_1^{ERD} = \Delta P_1 \times (V_{f,1} - \eta(V_{f,1} - V_{p,1})) \quad (37)$$

$$V_{p,2} = Q_{f,2} \times \frac{5}{7} \times 10\,hr = \frac{50}{7} Q_{f,2} \times hr \quad (38)$$

$$W_2^{ERD} = \Delta P_2 \times (V_{f,2} - \eta(V_{f,2} - V_{p,2})) \quad (39)$$

Therefore, the average SEC is:

$$\overline{SEC}^A_{ERD} = \frac{W_1^{ERD} + W_2^{ERD}}{V_{p,1} + V_{p,2}} = (1715 - 1015\eta)\,psi \quad (40)$$

Operating strategy B. The water recovery in the last 10 hrs is the same as the water recovery in the first 10 hrs. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be the same as that the feed flow rate in the last 10 hrs ($Q_{f,2}$'). Therefore, the permeate flow and energy consumption in the first and last 10 hrs are:

$$V'_{p,1} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \quad (41)$$

$$W_1^{'ERD} = \Delta P'_1 \times (V'_{f,1} - \eta(V'_{f,1} - V'_{p,1})) \quad (42)$$

$$V'_{p,2} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \quad (43)$$

$$W_2^{'ERD} = \Delta P_2 \times (V'_{f,2} - \eta(V'_{f,2} - V'_{p,2})) \quad (44)$$

Therefore, the average SEC is:

$$\overline{SEC}^B_{ERD} = \frac{W_1^{'ERD} + W_2^{'ERD}}{V'_{p,1} + V'_{p,2}} = 700(2 - \eta)\,psi \quad (45)$$

The SEC difference between operating strategy A and operating strategy B is $(1715 - 1015\eta) - 700(2 - \eta)\,psi = 315(1 - \eta)\,psi$. Thus, when $0 < \eta < 1$, the SEC of operating strategy A will be always greater than the SEC of operating strategy B. The fractional SEC increase is,

$$\frac{\overline{SEC}^A_{ERD} - \overline{SEC}^B_{ERD}}{\overline{SEC}^B_{ERD}} = \frac{315}{700} \frac{(1 - \eta)}{[1 + (1 - \eta)]} \quad (46)$$

which is plotted in Fig. 5. For example, when the ERD efficiency is 90%, the fractional SEC increase is 4.1%. Furthermore, in order to equate the total permeate volume in operating strategy A and operating strategy B, $Q'_{f,2} = \frac{10}{7} Q_{f,2}$. Thus, the total feed volume in operating strategy B is $2 \times \frac{10}{7} Q_{f,2} = \frac{20}{7} Q_{f,2}$, while the total feed volume in operating strategy A is $(2.5+1)Q_{f,2} = 3.5Q_{f,2}$. Therefore, in order to get the same amount of permeate volume, operating strategy A requires a higher amount of feed water, and thus, it has a lower overall water recovery.

In summary, operating strategy A is worse since we need to process more feed water to obtain the same permeate and has a higher SEC. In others words, by adjusting operating pressure to be double that of the instantaneous feed osmotic pressure, the system needs to process less volume of feed water to produce the same amount of permeate water and has a lower SEC.

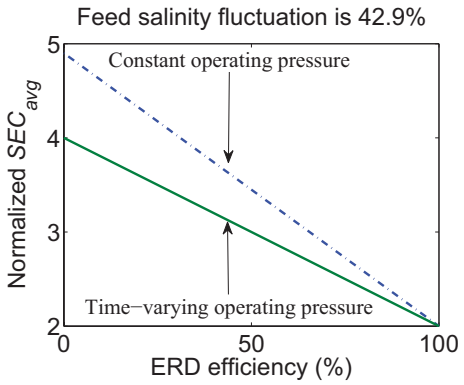


Figure 5. Percentage SEC increase when feed pressure is not adjusted. vs. ERD efficiency.

4. DISCUSSION

4.1 Effect of the Feed salinity Fluctuation Percentage on Energy Savings

The effect of the fluctuation amplitude on energy savings can be studied following the same procedure presented in Section 3.3. Assuming the average osmotic pressure is π_0, the osmotic pressure in the first 10 hrs is $(1 + \sigma)\pi_0$ $(0 < \sigma < 1)$, and the osmotic pressure in the last 10 hrs is $(1 - \sigma)\pi_0$. Therefore, the feed fractional fluctuation is σ. Similarly, the following two operating strategies may be considered.

- Operating strategy A: The transmembrane pressure is maintained at double that of the average feed osmotic pressure, i.e. $2\pi_0$.
- Operating strategy B: The transmembrane pressure is maintained at double that of the instantaneous feed osmotic pressure.

Operating strategy A. The water recovery in the last 10 hrs, $Y_1 = 1 - \frac{(1+\sigma)\pi_0}{2\pi_0} = \frac{1-\sigma}{2}$, and in the last 10 hrs, $Y_2 = 1 - \frac{(1-\sigma)\pi_0}{2\pi_0} = \frac{1+\sigma}{2}$. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be $\frac{1+\sigma}{1-\sigma}$ times that of the feed flow rate in the last 10 hrs ($Q_{f,2}$). The permeate flow and energy consumption in the first and last 10 hrs are:

$$V_{p,1} = \frac{1 + \sigma}{1 - \sigma} \cdot Q_{f,2} \cdot \frac{1 - \sigma}{2} \cdot 10\,hr = 5(1 + \sigma) \cdot Q_{f,2} \cdot hr \quad (47)$$

$$W_1^{ERD} = \Delta P_1 \times (V_{f,1} - \eta(V_{f,1} - V_{p,1})) \quad (48)$$

$$V_{p,2} = Q_{f,2} \times \frac{1 + \sigma}{2} \times 10\,hr = 5(1 + \sigma) \cdot Q_{f,2} \cdot hr \quad (49)$$

$$W_2^{ERD} = \Delta P_2 \times (V_{f,2} - \eta(V_{f,2} - V_{p,2})) \quad (50)$$

Therefore, the average SEC is:

$$\overline{SEC}^A_{ERD} = \frac{W_1^{ERD} + W_2^{ERD}}{V_{p,1} + V_{p,2}} = 2\pi_0[\frac{(1 - \eta)}{1 - \sigma} + \frac{(1 + \eta\sigma)}{1 + \sigma}] \quad (51)$$

Operating strategy B. The water recovery in the last 10 hrs is the same as the water recovery in the first 10 hrs. In order to produce the same amount of permeate volume, the feed flow rate in the first 10 hrs has to be the same as the feed flow rate in the last 10 hrs ($Q_{f,2}$'). The permeate flow and energy consumption in the first and last 10 hrs are:

$$V'_{p,1} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \tag{52}$$

$$W_1^{'ERD} = \Delta P'_1 \times (V'_{f,1} - \eta(V'_{f,1} - V'_{p,1})) \tag{53}$$

$$V'_{p,2} = Q'_{f,2} \times \frac{1}{2} \times 10\,hr = 5Q'_{f,2} \times hr \tag{54}$$

$$W_2^{'ERD} = \Delta P_2 \times (V'_{f,2} - \eta(V'_{f,2} - V'_{p,2})) \tag{55}$$

Therefore, the average SEC is:

$$\overline{SEC}^B_{ERD} = \frac{W_1^{'ERD} + W_2^{'ERD}}{V'_{p,1} + V'_{p,2}} = 2(2-\eta) \cdot \pi_0 \tag{56}$$

The SEC difference of operating strategy A from operating strategy B is $(2[\frac{(1-\eta)}{1-\sigma} + \frac{(1+\eta\sigma)}{1+\sigma}] - 2(2-\eta)) \cdot \pi_0$. When $0 < \eta < 1$, the SEC of operating strategy A will be always greater than the SEC of operating strategy B. The fractional SEC increase is:

$$\frac{\overline{SEC}^A_{ERD} - \overline{SEC}^B_{ERD}}{\overline{SEC}^B_{ERD}} = \frac{[\frac{(1-\eta)}{1-\sigma} + \frac{(1+\eta\sigma)}{1+\sigma}]}{(2-\eta)} - 1 \tag{57}$$

which is plotted in Fig. 6 when the efficiency of the ERD is set to be 90%. Fig. 6 shows that as feed salinity fluctuation percentage increases, time-invariant operation increases SEC more remarkably. Even in some cases there is only marginal energy savings, it is still worthwhile to adopt the proposed operating strategy accompanied by the control algorithms developed at UCLA M3 group regarding reverse osmosis water desalination system (McFall et al. (2008); Bartman et al. (2008)) since we will not be able to know what the future salinity profile would exactly be. Furthermore, in order to equate the total permeate volume in operating strategy A and operating strategy B, $Q'_{f,2} = (1+\sigma)Q_{f,2}$. Thus, the total feed volume in operating strategy B is $2(1+\sigma) \cdot Q_{f,2}$, while the total feed volume in operating strategy A is $(\frac{1+\sigma}{1-\sigma} + 1)Q_{f,2} = (1 + \frac{2\sigma}{1-\sigma} + 1)Q_{f,2} > (1+2\sigma+1)Q_{f,2}$. Therefore, in order to get the same amount of permeate volume, operating strategy A requires a higher amount of feed water, and thus, it has a lower overall water recovery.

5. CONCLUSION

Based on a model for a reverse osmosis membrane desalination plant and the feed concentration fluctuation (which is common in both seawater and brackish water desalination) profile, the proposed approach requires less amount of feed water and decreases specific energy consumption by as much as 22%, providing the same permeate flow. Experimental results confirming the proposed operating policy will be presented at the conference.

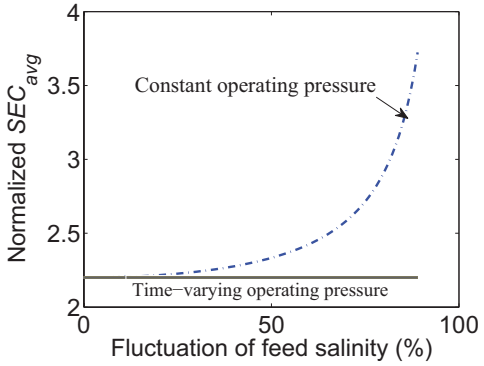


Figure 6. Percentage SEC increase without adjusting the operating pressure vs. feed concentration fluctuation percentage.

REFERENCES

Manual. Standard Practice for Standardizing Reverse Osmosis Performance Data, *ASTM international, April 2000.*

Bartman, A., McFall, C.W., Christofides, P.D., and Cohen, Y. (2009). Model-predictive control of feed flow reversal in a reverse osmosis desalination process. *Journal of Process Control*, 19, 433–442.

Busch, M. and Mickols, W. (2004). Reducing energy consumption in seawater desalination. *Desalination*, 165, 299–312.

Manth, T., Gabor, M., and Oklejas, E. (2003). Minimizing RO energy consumption under variable conditions of operation. *Desalination*, 157(1-3), 9–21.

McCool, B.C. (2008). The feasibility of reverse osmosis desalination of brackish water in the San Joaquin Valley. UCLA master thesis.

McFall, C.W., Bartman, A., Christofides, P.D., and Cohen, Y. (2008). Control and monitoring of a high recovery reverse osmosis desalination process. *Industrial & Engineering Chemistry Research*, 47(17):6698–6710.

Mulder, M. (1997). *Basic Principles of Membrane Technology.* Kluwer Academic Publishers (Boston).

Song, L., Hu, J.Y., Ong, S.L., Ng, W.J., Elimelech, M., and Wilf, M. (2003a). Emergence of thermodynamic restriction and its implications for full-scale reverse osmosis processes. *Desalination*, 155(3), 213–228.

Song, L., Hu, J.Y., Ong, S.L., Ng, W.J., Elimelech, M., and Wilf, M. (2003b). Performance limitation of the full-scale reverse osmosis process. *Journal of Membrane Science*, 214(2), 239–244.

Song, L. and Tay, K.G. (2006). Performance prediction of a long crossflow reverse osmosis membrane channel. *Journal of Membrane Science*, 281(1-2), 163–169.

Wilf, M. (1997). Design consequences of recent improvements in membrane performance. *Desalination*, 113(2-3), 157–163.

Wilf, M. and Bartels, C. (2005). Optimization of seawater RO systems design. *Desalination*, 173(1), 1–12.

Zhu, A., Christofides, P.D., and Cohen, Y. (2008). Effect of thermodynamic restriction on energy cost optimization of RO membrane water desalination. *Industrial & Engineering Chemistry Research*, Publication Date (Web): August 29, 2008 (DOI: 10.1021/ie800735q).

Zhu, A., Christofides, P.D., and Cohen, Y. (2009). Minimization of Energy Consumption for a Two-Pass Membrane Desalination: Effect of Energy Recovery, Membrane Rejection and Retentate Recycling. *Journal of Membrane Science*, doi:10.1016/j.memsci.2009.04.039.

A novel image based algorithm for interface level detection in a separation cell

Phanindra Jampana * **Sirish Shah** **

* *Department of Chemical Engineering, University of Alberta, Edmonton, Canada (e-mail: pjampana@ualberta.ca)*
** *Department of Chemical Engineering, University of Alberta, Edmonton, Canada (e-mail: slshah@ualberta.ca)*

Abstract: Controlling the interface between Bitumen-froth and Middlings in separation cells in the oil sands industry is important for economical and environmental reasons. Traditional sensors do not provide reliable measurements of this interface level and image based sensors are being used to alleviate this problem. Previous work in this area has focussed on separation cells with a single side-view glass. The current work describes a new image based algorithm for interface level detection and confidence estimation based on the concept of image differencing. The algorithm can be extended in a straight-forward manner to separation cells with arbitrary number of side-view glasses. Off-line and on-line results show that the algorithm accurately detects the interface level in normal process conditions and outputs correct confidence values in other situations with very low false positive and negative rates.

Keywords: Bitumen-froth Middlings interface, image sensors, image differencing

1. INTRODUCTION

The control of Bitumen-froth and Middlings interface using image based sensors has been approached previously (Jampana et al., 2008) via particle filtering techniques. Images obtained from a *side-view* glass camera are processed in real time for estimates of the interface level and its quality. These estimates are used subsequently for automatic control. A typical camera image from this setup is shown in Fig 1. For separation cells with multiple side-view glasses (Fig 2) the algorithm described there does not generalize in a straight forward manner. The current work describes an interface level detection algorithm based on image differencing which can be easily generalized to arbitrary number of side-view glasses.

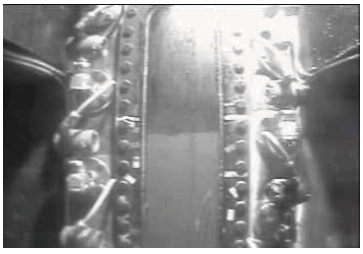


Fig. 1. Separation cell with single side view glass

The generalization property is achieved by computing a confidence estimate (in addition to the interface level estimate) for each side-view glass. This confidence estimate quantifies the chance of the presence of an interface. The final interface level estimate is obtained from the view glass with the highest quality. As confidence estimation is not entirely independent of the interface level estimation procedure, the interface level estimation procedure should facilitate the computation of quality values in an easy manner. The image differencing algorithm described in this paper is one such method.

⋆ Financial support from NSERC, Matrikon, Suncor and iCORE in the form of the Industrial Research Chair(IRC) program at the University of Alberta is gratefully acknowledged.

Fig. 2. Separation cell with three side view glasses

The image differencing method is based on the idea that the *change* from any previous video frame to the current video frame is maximum near the current interface, though this maximum need not be unique. This change is detected here through (absolute) image differencing. To ensure that the maximum change occurs very close to the current interface, (absolute) image differences between the current and many previous frames are used. The sum image of all these differenced images has maximum values located close to the current interface level for ideal interface images, i.e. images which are completely free from noise. The proof of this fact is given in section 2.

In reality, interface images are seldom noise free. This leads us into estimating a quality value which reflects whether the current interface level estimate is purely a result of

noise. To compute the final confidence estimate however, the noise based quality value alone would not suffice. This is because abnormal changes might occur in the separation cell, which cannot be ascribed to noise alone and which do not necessarily imply the existence of a true interface. Fig 3 shows an example of such a change where the noise based quality described above might be high but the interface is spurious.

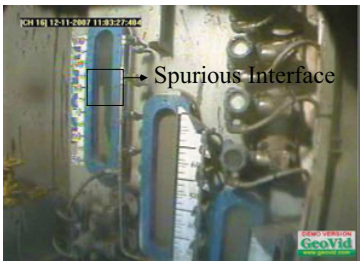


Fig. 3. Abnormal process condition resulting in a spurious interface

Therefore, apart from the noise based quality value, an edge quality is also estimated. This quality value quantifies the number of edges detected near the interface estimate. The edge detection method employed here is described in (Elder and Zucker, 1998). A combination of both these quality estimates suppresses most false negatives. In a few pathological cases however, both the noise based quality and the edge based quality can be high, even though the detected interface is spurious. To make the algorithm robust to these, a change based quality is estimated. The final confidence estimate is then based on the three values - noise based quality, edge based quality and change based quality.

The rest of the paper is organised as follows: Section 2 presents the image differencing based interface level detection algorithm in detail followed by section 3 which describes the confidence estimation procedure. Results are displayed in section 4 and section 5 gives the concluding remarks.

2. INTERFACE LEVEL DETECTION USING ABSOLUTE IMAGE DIFFERENCING

This section presents two results which describe the image differencing method for ideal images and also provide theoretical bounds for the interface level estimates obtained. The following notation is fixed first.

Let,

(1) I_t represent the video frame obtained at time t
(2) $D_{t_1,t_2} = I_{t_1} - I_{t_2}$, be the difference of two images at times t_1 and t_2
(3) $AD_{t_1,t_2} = abs(I_{t_1} - I_{t_2})$, be the absolute difference of two images at times t_1 and t_2
(4) $il(t)$ represent the interface level at time t (The interface level is always assumed to be on the Middlings side of the interface).
(5) $\mu_B(t), \mu_M(t)$ represent the average intensity values of pixels in the Bitumen-froth and Middlings regions at time t, respectively.
(6) W and H represent the width and height of the interface image
(7) C be the maximum change in the interface level between two successive video frames
(8) The origin of images is always assumed to be at the top-left corner

Lemma 1. Consider a noise free interface having dynamics such that it remains horizontal at all times and having homogeneous pixel intensities in the Bitumen and Middlings regions. Let $\{I_t, t = 0, 1, ...\}$ be a sequence of completely noiseless images from such an interface such that $\mu_B(0) = \mu_B(1) = \mu_B(2) = ...$ and $\mu_M(0) = \mu_M(1) = \mu_M(2) = ...$ If there is a change in the interface level in a time window $[t_0, t_N]$ and if:-

$$J_N(i,j) = \sum_{k=0}^{N-1} AD_{t_N,t_k}(i,j), \quad \forall i \in \{0,1,2,...,H-1\}, \quad j \in \{0,1,2,...,W-1\}$$

$$P_N(i) = \sum_{j=0}^{W-1} J_N(i,j), \forall i \in \{0,1,2,...,H-1\}$$

$$\hat{il}(t_N) = \inf(\arg\max_i P_N(i))$$

then,

(1) $P_N(i)$ is decreasing in $i \in [\hat{il}(t_N), H-1]$ and increasing in $i \in [0, \hat{il}(t_N)]$,
(2) $0 \leq il(t_N) - \hat{il}(t_N) \leq C$

In reality the interface is seldom horizontal. Lemma 2 guarantees similar bounds as above for the interface level estimate even for the more general case of non-horizontal interfaces:

Lemma 2. Consider the more general case of an interface having dynamics such that it can become non-horizontal. Let $ip(t,v)$ for $v \in [0, W-1]$ be the interface pixels at time t. If $|ip(t,v) - ip(t,m)| < Q$, for all $v, m \in [0, W-1]$, $t \in [t_0, t_N]$ and $|ip(t_1,v) - ip(t_2,v)| < C$ whenever $|t_1 - t_2| = 1$, $v \in [0, W-1]$ and if there is a change in the interface in the time window $[t_0, t_N]$ then it is true that $-C \leq \hat{il}(t_N) - ip(t_N, v) \leq C + Q$ for some $v \in [0, W-1]$.

(Proof for both proofs are omitted due to a lack of space but are available from the authors). The above lemma shows that in the absence of noise and non-homogeneties in images, the estimated interface level is close to the actual interface, especially if C and Q are small. However, when the images are corrupted by noise and other non-homogenties in pixel intensities, the estimated interface level might not be close to the actual interface. Hence a confidence value of the interface level estimate is computed.

3. CONFIDENCE ESTIMATION

The analysis above assumed that images obtained are completely noise free - an assumption that is never met in practice. Image noise is modelled to be additive, homogeneous and Gaussian with zero mean and variance σ^2.

In the presence of noise, it might no longer be true that $\hat{il}(t_N)$ will lie close to an interface point as predicted by

Lemmas 1 and 2. This is because the images observed are only instantiations of a (two dimensional) random field, which is completely described only by the noise statistics, the interface level $il(t_N)$ and the Middlings and Bitumen pixel intensities $\mu_M(t), \mu_B(t)$. Hence, each $P_N(i)$ for $i \in [0, H-1]$ now has a probability distribution. In the case of a horizontal interface, given the noise distribution, the probability that is of interest is the following:-

$$P(P_N(il(t_N)) > \max_{|j-il(t_N)|>G} P_N(j))$$

The above probability quantifies the chance of obtaining an interface level estimate (by following the differencing method described before), $\hat{il}(t_N)$, which satisfies $|\hat{il}(t_N) - il(t_N)| \leq G$. This probability can be used as the confidence value but it cannot be determined, as $il(t_N)$ cannot be known a priori.

As the theoretical confidence (the probability above) cannot be computed, a confidence estimate is obtained by heuristic methods. The confidence estimate is based on the following three quality values, which are explained subsequently:

- Noise based quality
- Edge based quality
- Change based quality

3.1 Noise based quality

Let $TP(t, i, j)$ represent the true (expected) pixel value in the image at time t and at the location i, j. Then the observed value of each pixel $I_t(i, j)$ can be written as $TP(t, i, j) + Y(t, i, j)$, where $Y(t, i, j)$ is a random variable whose distribution is the same as the estimated noise distribution. Using this, the following can be derived:

$$\begin{aligned} P_N(i) - \sum_{j=0}^{W-1}\sum_{k=0}^{N-1} |Y(t_N, i, j) - Y(t_k, i, j)| \\ \leq \sum_{j=0}^{W-1}\sum_{k=0}^{N-1} |TP(t_N, i, j) - TP(t_k, i, j)| \\ \leq P_N(i) + \sum_{j=0}^{W-1}\sum_{k=0}^{N-1} |Y(t_N, i, j) - Y(t_k, i, j)| \end{aligned}$$

The above inequality gives loose bounds on the actual values, $\sum_{j=0}^{W}\sum_{k=0}^{N-1} |TP(t_N, i, j) - TP(t_k, i, j)| \equiv M(i)$, i.e., the values which would have resulted if the images are noise free. In practice, only one instance of $P_N(i)$ is observed. From this value, the value of the corresponding instance of $R_N(i) \equiv \sum_{j=0}^{W}\sum_{k=0}^{N-1} |Y(t_N, i, j) - Y(t_k, i, j)|$ cannot be computed. Therefore the bounds above cannot be determined exactly.

Given $P_N(i) = \hat{P}_N(i)$, $R_N(i)$ follows the conditional probability distribution given by $P(R_N(i)|P_N(i) = \hat{P}_N(i))$. Considering the instances $(\hat{R}_N(i))$ of this distribution allows us to compute inequalities which are be obeyed with a certain degree of probability. For example, if $P_{R_N(i)|P_N(i)}(R_N(i) \leq \hat{R}_N(i)) = r(i)$, then the inequalities

$$\hat{P}_N(i) - \hat{R}_N(i) \leq M(i) \leq \hat{P}_N(i) + \hat{R}_N(i)$$

are true with a probability of $r(i)$. If $\hat{R}_N(i)$ are chosen such that $r(i)$ are very high, then the inequalities are very likely to be satisfied. On the other hand, if the $\hat{R}_N(i)$ are chosen such that $r(i)$ are very low, it is very unlikely that the inequalities will be correct. Given a choice of $\hat{R}_N(i)$, the noise based quality can be defined as:

$$Q_{noise}(t_N) = \begin{cases} 0; \text{if } \exists i, |i - \hat{il}(t_N)| > N_{TH}, \\ \hat{P}_N(i) + \hat{R}_N(i) > \\ \hat{P}_N(\hat{il}(t_N)) - \hat{R}_N(\hat{il}(t_N)) \\ 1; \text{otherwise} \end{cases}$$

This quality value penalizes the interface level estimates when the minimum bound of $M(\hat{il}(t_N))$ is less than the maximum bound of $M(i)$, for i far away ($N_{TH} > C$) from the current interface. In this case, the interface estimate is said to be obtained purely due to camera noise and other irregularities in the images.

As the conditional probability distribution cannot be estimated, the instance $\hat{R}_N(i)$, is chosen based on the unconditional one. The support of the unconditional distribution is a superset of the support of the conditional distribution. Hence, for high values of $\hat{R}_N(i)$ (based on the unconditional distribution) the inequalities obtained will very likely be true. But high values of $\hat{R}_N(i)$ make the bounds very loose and are not useful for noise based quality estimation as most quality estimates will be zero. On the other hand, for small values of $\hat{R}_N(i)$, the quality estimates might be high but the inequalities themselves are true only with a very small probability.

The problem is to obtain estimates $\hat{R}_N(i)$, for which the inequalities will be true with a high probability and are tight enough for use in noise based quality estimation. In the absence of any other information, the choice $E(R_N(i)) = \hat{R}_N(i)$, where E represents mathematical expectation can be considered a possible candidate. From basic probability and the properties of the Gaussian distribution, it can be computed that $E(R_N(i)) = NW\sigma\sqrt{(\frac{8}{\pi})}$.

The accuracy of the noise based quality estimates $Q_{noise}(t_N)$, obtained by the choice $\hat{R}_N(i) = E(R_N(i))$ depends on the absolute difference of average pixel intensities $|\mu_B(t_N) - \mu_M(t_N)|$, the size of the images and the noise standard deviation σ. Based on this dependence, false positive and false negative error rates for the noise based quality are estimated.

When σ is small and $|\mu_B(t_N) - \mu_M(t_N)|$ is high, the false positive rate is approximately $0 - 2\%$, which is small as expected. This rate increases with an increase in σ but decreases with an increase in $|\mu_B(t_N) - \mu_M(t_N)|$. The ratio $\frac{|\mu_B(t_N) - \mu_M(t_N)|}{\sigma}$ can be considered as an upper bound on the Signal to Noise ratio (SNR). If $\frac{|\mu_B(t_N) - \mu_M(t_N)|}{\sigma} = 10$, the false positive error rate is $7 - 8\%$ on an average.

For computing the false negative error rates, random interface images, which do not contain an interface are created. As these images do not contain any interface the percentage of time $Q_{noise}(t_N) = 1$ is considered an estimate of the false negative error rate. In a simulation

study using the same parameters as above (except that $\mu_B(t_N) = \mu_M(t_N)$), it has been found that there were no false negatives. As other type of examples cannot be readily created to study the false positive and negative error rates, they are estimated on real videos collected from a plant site. These are presented in Section 4.

3.2 Edge based quality

Noise based quality alone is not sufficient for estimating confidence. This is due to the fact that false negatives result when abnormal changes occur inside the separation cell (scenarios as shown in Fig 3) which cannot be explained by noise alone. Hence, an additional edge detection algorithm is used to aid in the estimation of the confidence.

The motivation for using edge detection to estimate a quality value is that the available information in images would be utilised in a very efficient manner as the edge based algorithm captures information which cannot be obtained by image differencing. Given only the difference images $I(t) - I(s)$ for $s < t$, it is impossible to recover the edge map of $I(t)$ and similarly given only the edge map of $I(t)$, it is impossible to estimate the difference images except in a few pathological cases. Using the image differencing and the edge detection algorithms simultaneously most false negatives (high confidence values when the interface level estimates are wrong), can be avoided.

The algorithm described in (Elder and Zucker, 1998) is used here with the already estimated variance σ^2 of the Gaussian noise distribution. The advantage of this particular edge detection algorithm over standard algorithms (Sobel, Canny etc.) is its ability to detect edges over a large blur scale and contrast. The Bitumen-Middlings interface tends to become fuzzy when the percent of sand in the oil sands ore is high. The chosen algorithm can detect edges under these situations and hence is suitable for the purpose. Another reason for the choice is that spurious edges that occur due to sensor noise are minimised because of statistical bound checking based on the sensor noise variance in the algorithm. This increases the efficiency of the edge based quality.

A simple heuristic based on the number of edge points in a predefined window near the detected interface level is used to estimate the edge based quality. If EI is the edge map returned by the edge detection algorithm, and if $nedges$ represent the number of edges in a predefined window near the detected interface level and E_{TH} is a given threshold then the edge based quality is defined as:

$$Q_{edge}(t_N) = \begin{cases} 0; nedges < E_{TH} \\ 1; \text{otherwise} \end{cases}$$

3.3 Change based quality

The edge detection algorithm, in most cases does not produce the exact edge map, EI. When spurious edges are detected (due to shadows, lighting glare etc.), the edge based quality might be high even when the interface level estimate is not correct. If the noise variance is under estimated, the noise based quality would also be high resulting in a wrong estimate of the interface level. False negatives in interface level detection can have an undesired effect on the overall process as the controller takes immediate corrective action based on these false readings.

To make the algorithm robust to such cases a quality based on the percent change near the interface is estimated. The change based quality analyzes the instance of P_N observed, $\hat{P}_N$. An example $\hat{P}_N$ (for a normal interface image sequence) is shown in Fig 4.

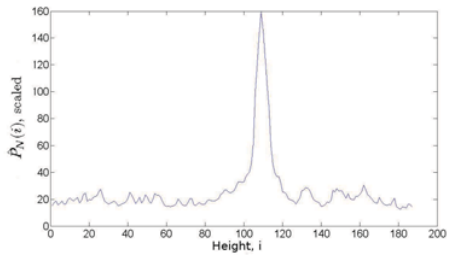


Fig. 4. An example of the profile, P_N obtained by the image differencing method

For a normal interface, based on test videos, the average and minimum values of $\hat{P}_N$ have been observed to be close to each other as shown in the figure above. The maximum value of P_N, $\hat{P}_N(\hat{il}(t_N))$, is in general high compared to both these values. Based on this, the change based quality is defined simply as:

$$Q_{change}(t_N) = \begin{cases} 0; \dfrac{\max(\hat{P}_N) - avg(\hat{P}_N)}{\max(\hat{P}_N) - ((1-\epsilon)\min(\hat{P}_N))} \\ \quad < C_{TH} \\ 1; \text{otherwise} \end{cases}$$

Here, $0 < \epsilon \approx 10^{-2} << \min(\hat{P}_N)$. The change based quality value would be high when the average value of $\hat{P}_N$ is close to the minimum value of $\hat{P}_N$. When the average is close to the maximum this quality value would be small.

The thresholds – N_{TH}, E_{TH}, C_{TH} and N determine the performance of the final algorithm. The value of E_{TH} is chosen as a percentage of the width of the image W and the value of $C_{TH} \in [0, 1]$. Hence both these thresholds are relative in nature. The value of N_{TH} is chosen based on the dynamics of the interface. Based on the three quality values, the final confidence is defined as

$$il_{conf}(t_N) = \begin{cases} 1; Q_{noise}(t_N) = 1, \\ \quad Q_{edge}(t_N) = 1, \\ \quad Q_{change}(t_N) = 1 \\ \quad 0; \text{otherwise} \end{cases}$$

4. RESULTS

4.1 Off-line results

The algorithm is first tested off-line on three videos recorded at the Suncor Energy Inc. plant site located at Fort McMurray, Alberta, Canada. The first video contained only one side view glass whereas the other two were equipped with three side view glasses. In the first video (Fig 5a) the view glass was wider and the interface was always present inside it. There was also significant lighting

glare present on the top of the glass window. The other two videos had considerably smaller view glasses. In one of these two videos (Fig 5c), the interface was only present in two of the three view glasses. In the other video (Fig 5b), spurious changes occurred (due to Bitumen sticking on the inside) in one of the glasses initially and the interface reappeared at the end.

The original videos were from colour cameras and for the purpose of analysis, they were converted to grayscale by averaging across all the three (RGB) colour channels. For single side view glass, the algorithm as described in the sections before can be applied directly. Whereas, in the case of three view glasses, the algorithm is extended in a straightforward manner. Each glass window is analysed separately and finally the window with the highest confidence value is chosen along with its interface level estimate. In cases where the interface is present in two or more glasses, more than one window can have a high confidence value. In such situations, the final interface level is chosen at random from these glasses, as all of the interface level estimates refer to the same interface.

In all the videos the same parameters, $N = 100$, $N_{TH} = 30$ pixels, $E_{TH} = \frac{W}{4}$, $C_{TH} = 0.75$ were used. Fig 6a shows the true and the estimated interface level values for the video with a single side view glass ($H = 188$ pixels, $W = 61$ pixels). It can be seen that the estimated value is very close to the actual value. The average absolute error was calculated to be approximately two pixels. This corresponds to an average error of less than one percent with respect to the height of the view glass. The confidence estimate was equal to one throughout(except at one frame where the edge based quality was zero). Noise standard deviation was estimated to be $\sigma = 1.0$ pixels and $|\mu_B(t_N) - \mu_N(t_N)| = 21.9$ intensity units. The corresponding false positive has been estimated to be zero which explains the fact that the noise based quality was equal to one throughout. Edge based quality was also high because the interface was clear and easily detectable by the edge detection algorithm. The change based quality was one throughout.

For the video with three side view glasses shown in Fig 5c, the results obtained are shown in Fig 6b. Note that in this case, the interface level estimate corresponds to the view glass with the highest confidence value. The average absolute error was calculated to be three pixels approximately, which corresponds to an average error of less than one percent with respect to the height of the view glass, as before. The confidence estimate was equal to one at all times except for three frames. The noise based quality was equal to one throughout but the edge based quality was zero at these three frames owing to significant fuzziness in the interface (not shown here). The change based quality was one throughout as before.

Finally, the video shown in Fig 5b is split into two segments. In the first part, the interface was either spurious or not present in the view glass. For this segment of the video the false negative rate obtained was equal to zero, i.e. the confidence value was identically zero all the time. Fig 6c shows the estimated and the actual interface level for the second part of the video, when the interface reappeared in the view glass. The average absolute difference was equal to three pixels which corresponds to an error of less than one percent with respect to the height of the view glasses. The false positive rate during this time was estimated to be 10%, due to zero edge based quality during those frames. The high false positive rate in this video can be attributed to following:-

- Loss of resolution from the original to the recorded video resulting in a poor quality of the video
- Highly fuzzy interfaces occur due to a high fines situation – too many sand particles in the Bitumen-froth

The false positive rate can be minimized by employing a simple filtering rule. In the industry, a single occurrence of a confidence value of zero triggers an alarm for operator intervention. As the confidence value is susceptible to sudden changes in the fuzziness of the interface it is reasonable to wait until the confidence value stabilizes. Hence, instead of signalling an alarm for a single occurrence, alarm is only signalled when the confidence value is zero for a sustained period of time ($\tau \approx 5s$). The interface level estimate used for control during this phase is the most recent estimate with a confidence value of one. This simple filtering rule has been observed to increase the efficiency of the algorithm.

4.2 On-line results

The algorithm described in this paper has been implemented on two separation cells (previously shown in Fig 5b and Fig 5c) at Plant 86, Suncor Energy Inc., Fort McMurray, Alberta, Canada. A frame grabber card is used to transfer the images from the analog cameras to the PC. Software has been built in in the C programming language based mainly on the Intel OpenCV library for image manipulation.

Fig 7a compares the true and estimated interface level values for the separation cell shown in Fig 5b. In this plot, hourly data is collected at random times and stitched together for the final result. A total of eight hours of data is used for comparison. On this data set, the average absolute error(in percentage) was calculated to be four percent. Similarly, Fig 7b compares the true and estimated interface level values for the separation cell shown in Fig 5c. The average absolute error was equal to three percent of the total height of the view glasses. These results suggest that the estimates from the vision sensor very closely reflect the true interface level values.

5. CONCLUSIONS

This work has presented a novel image differencing method for Bitumen-froth and Middlings interface level detection. It has been shown that in the case of noiseless images the estimation error is bounded. For nominal values of the dynamics of the separation cell, the bounds are very small.

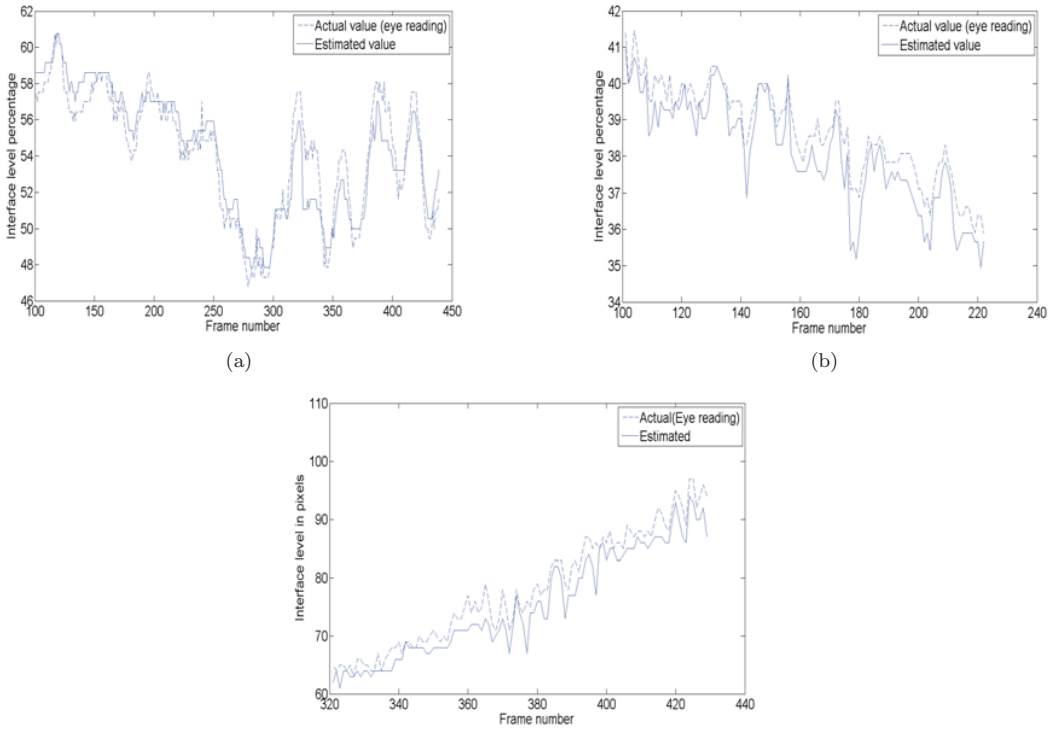
When noise is present in the images, a confidence value which estimates the correctness of the detection is computed. The confidence value is based on a novel noise based quality estimate along with simple edge and change criterion. Analysis and results show that the final algorithm accurately detected the interface level and exhibited

(a) Video with one side view glass

(b) Video with spurious changes

(c) Video with interface present in two glasses

Fig. 5. Interface levels in different separation cells

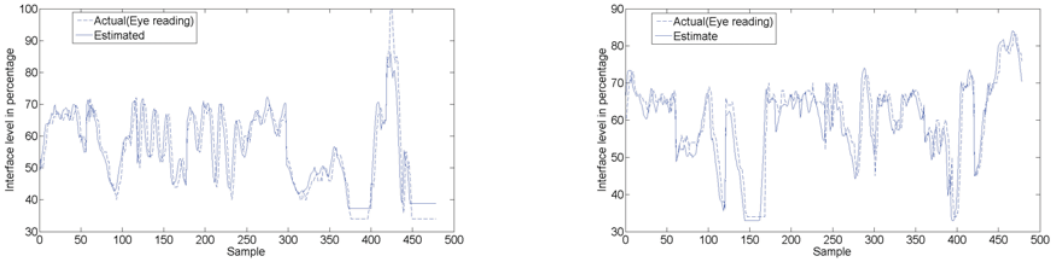


(a)

(b)

(c)

Fig. 6. True and estimated interface levels - Off-line results



(a) True and estimated interface levels for separation cell shown in Fig 5b

(b) True and estimated interface levels for separation cell shown in Fig 5c

Fig. 7. True and estimated interface levels - On-line results

very few false positive and negative error rates. The sensor has been installed at Suncor Energy. Inc, Fort McMurray, Canada and has been yielding highly satisfactory results.

REFERENCES

Elder and Zucker (1998). Local scale control for edge detection and blur estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(7), 699 – 716.

Jampana, P., Shah, S., and Kadali, R. (2008). Computer vision based interface level control in a separation cell. *IFAC world congress*.

Model-Based Dosing Control of a Pellet Softening Reactor

K.M. van Schagen [*,**] **R. Babuška** [**] **L.C. Rietveld** [***] **A.M.J. Veersma** [****]

[*] *DHV B.V., Amersfoort, The Netherlands (e-mail: kim.vanschagen@dhv.com)*
[**] *Delft Center for Systems and Control, Delft University of Technology, The Netherlands*
[***] *Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands*
[****] *Waternet, Amsterdam, The Netherlands*

Abstract: The control of a drinking-water treatment plant aims to produce the correct quantity of water, with a constant quality. Achieving constant water quality is not an obvious task, since the online water-quality measurements and possible control actions are limited. Applying model-based control improves disturbance rejection and online process optimisation. For the softening process step, the integral control scheme is shown with multiple controllers for different time scales and process detail. The dosing control is elaborated and verified using simulation experiments. The control is implemented and tested in the pilot plant of Weesperkarspel (Amsterdam). It shows that in the case of accurate state estimation, quick changes in setpoint can be tracked.

Keywords: multivariable control; MPC; crystallisation

INTRODUCTION

In the last decades, most drinking-water treatment plants have been automated. During these first automation realisations, the goal was to operate the treatment plant in the same way as the operators did before. Therefore the control configurations consisted of a heuristic control strategy, based on historical operator knowledge. The controls are designed for the static situation, including extra safety margins to take operator response into account. This was a logical and practical solution. However, this heuristic solution does not optimise the control of a treatment plant.

The heuristic control is based on static local control objectives, without taking the current state of the treatment plant into account. Therefore it is necessary to adopt a new control strategy, which can take into account quality-related and economic criteria and optimise the overall performance of the plant, based on the current state of the processes.

Since the treatment steps are coupled, local changes affect other treatment steps and therefore local optimisations should be considered in a global context. It is necessary that operational actions do not introduce new disturbances to other processes. This must be considered in all levels of control, from basic valve controllers to plant-wide quantity control. At the same time, the control should consider the actual state of the process and optimise plant operations.

The information density in the online measured data of water treatment plants is limited and multiple measurements have to be used to obtain a good view of the actual treatment performance (van Schagen et al., 2006b). By using white or grey models, the process knowledge is no longer stored as historical heuristic rules of thumb or static local control objectives. The local control objectives evolve from applying the new criteria to the existing models in the case of changes to the process, such as boundary conditions, influent properties and desired treated water quality.

The model-based dosing control is part of the new model-based control configuration for the pellet-softening treatment step, consisting of a number of pellet reactors and a bypass. The pellet softening process step at the Weesperkarspel treatment plant is described in the first section. The model-based control configuration is elaborated in the second section. Finally the model-based dosing control scheme is validated in simulation experiments and finally validated in the pilot plant of Weesperkarspel (Amsterdam).

PROCESS DESCRIPTION

In the Netherlands, softening of drinking water in treatment plants is mainly carried out with fluidised pellet reactors. The pellet reactor consists of a cylindrical vessel that is partly filled with seeding material (figure 1). The diameter of the seeding grain is small, between 0.2 and 0.4 mm and consequently the crystallisation surface is large. The water is pumped through the reactor in an upward direction at high velocities, maintaining the seeding material in a fluidised condition. In the bottom of the reactor, chemicals are dosed (caustic soda, soda ash or lime). Calcium carbonate then becomes super-saturated and crystallises on the seeding material, resulting in the

formation of pellets. At regular intervals, pellets at the bottom of the reactor are removed. These pellets can be re-used in industry (van Dijk and Wilms, 1991).

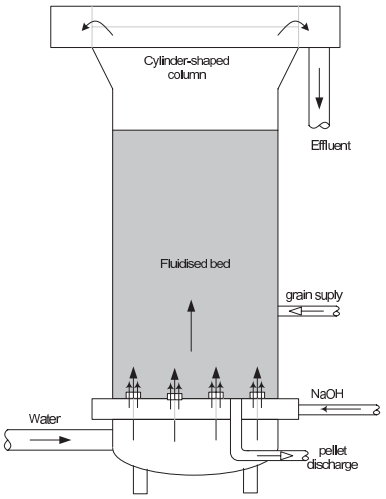


Fig. 1. Fluidised bed reactor for water softening.

Softening in a reactor is normally deeper than the required levels. Therefore, part of the water can be bypassed and mixed with the effluent of the reactors. In general, several identical parallel reactors are installed to increase the reliability of the system and the flexibility in operation. Reactors can be switched on and off in case of flow changes, maintaining water velocities between 60 and 100 m/h.

The mixture of the effluent of the reactors and the bypass water must be chemically stable to avoid crystallisation in the filters after the softening step.

At Weesperkarspel caustic soda (NaOH) is dosed for softening. The seeding material is garnet sand. The dosing of caustic soda in the pellet reactor is adjusted to realise the mixed effluent hardness of 1.5 mmol/l. The pellet removal is based on the hydraulic resistance of the fluidised bed (head loss) and the goal was to keep the hydraulic resistance constant. The garnet sand dosage was a manually set percentage of the mass of discharged pellets. The pH, flow, water temperature and hydraulic resistance were measured every minute, while hardness, calcium, bicarbonate, super saturation, pellet diameter and bed height were measured at longer intervals (Rietveld, 2005).

The characteristics of the softening process at Weespekarspel are given in table 1.

Table 1. Characteristics of softening reactors at Weesperkarspel.

Number of reactors	8	-
Surface area of reactor	5.3	m^2
Maximum bed height	5	m
Typical water velocity	60-100	m/h
Grain size of seeding material	0.25 10^{-3}	m
Density of the seeding material	4114	kg/m^3

CONTROL CONFIGURATION

The aim of the control of the softening process is to achieve a desired calcium concentration and, at the same time, minimise the use of dosage material (caustic soda, seeding grains and acid). The available control inputs are the water flow through the bypass and for each reactor the water flow through the reactor, the grain supply rate, the pellet discharge rate, the caustic soda dosage and the acid dosage.

To control the complete treatment step, a modular control setup is chosen. In this way, the controller complexity is minimised, maximising operator understanding of the control structure. Due to the diverse time constants in the process, these controllers are implemented on different platforms, with appropriate performance for the controllers. Figure 2 shows the control modules that are related to the softening process step. On the vertical axis represents the typical time constant of the controller and the horizontal axis shows the process level of the controller.

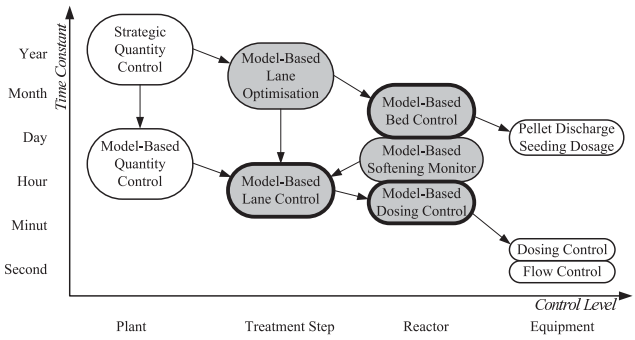


Fig. 2. Control setup for the pellet-softening treatment step. Modular controllers for different time constants and control levels.

The *Strategic Quantity Control* determines the amount of water, which has to be produced at the treatment plant. This is based on yearly consumption patterns, available resources at this plant and, in a multiple plants setup, the other treatment plants. The amount of water to be treated, is then passed to the Model-Based Quantity controller and the Model-Based Lane Optimisation.

The *Model-Based Quantity Control* determines the actual production rate of the entire plant, based on expected daily consumption pattern and the available water in the storage tanks. Restrictions in production rate, due to short-term maintenance, are taken into account and fluctuations of production rate are minimised (DHV, 2008).

The *Model-Based Lane Optimisation* determines the ideal pellet size, bypass ratio and the optimal number of reactors in operation, based on the expected production rate from the Stategic Quantity Control and the expected temperature variations. Changing bed configurations is a long term optimisation, due to the retention time of seeding material in the reactor of approximately 100 days. An extensive description of this optimisation scheme can be found in (van Schagen et al., 2008c).

The *Model-Based Bed Control* achieves the optimal bed composition as found with the Model-Based Lane Optimisation by determining the required pellet discharge and seeding material rates (van Schagen et al., 2008c). It uses the estimation of the current bed composition, determined by the Model-Based Monitor. This can be the model-based monitor of the complete reactor as shown in van Schagen et al. (2006b).

The *Model-Based Monitor* estimates the accuracy of the measurement devices and determines the actual state of the softening process. This monitor is used to verify the measurements that are used by the other controllers. In the case of unexpected differences between measurement and model outcome, operators are notified to take appropriate action. If measurement accuracy is sufficient, the model can be used to estimate unmeasured quality parameters using online measurements and historical laboratory results. Finally the actual state of the process can be estimated, such as the diameters of the pellets in the softening reactor at different heights. An extensive description of this monitoring scheme can be found in van Schagen et al. (2006b).

The *Model-Based Lane Control* determines the current flow and quality setpoints for each lane. It uses the estimated bed composition from the Model-Based Monitor and the actual production rate from the Model-Based Quantity Control. This controller is introduced, since the fluidised bed has limited control possibilities and it is expected that the actual bed composition is different for each reactor. The Model-Based Bed Control strives for the optimal bed composition, while the Model-Based Lane Control adapts to the current bed composition. The Model-Based Lane Control is elaborated in van Schagen et al. (2006a).

The *Model-Based Dosing Control* determines the actual dosing of caustic soda in the reactor to achieve the desired calcium concentration after the reactor, while respecting the constraints of the reactor. The objective of this controller is to follow the setpoint for the Model-Based Lane Control smoothly. The Model-Based Dosing control is shown in this article.

The *Pellet Discharge*, *Seeding Dosage*, *Dosing Control* and *Flow Control* follow the setpoints from the model-based controllers, by adjusting the physical devices such as valves and pumps. These local controllers are implemented in the process automation system of the plant.

MODEL-BASED DOSING CONTROL

The control of water flow and base dosage in the softening reactor is not straightforward. The dosing control and flow control are strongly interrelated. The retention time in the reactor is at least five minutes and response to control actions can only be detected after this time, since water quality can only be measured in the effluent of the reactor. The measurement of the total hardness (the main controlled variable), is a semi-online measurement and has a delay of at least ten minutes. The online pH measurement is inaccurate and has a tendency to drift. Changes in flow and dosing must be gentle, to prevent introduction of process disturbances and fast-changing water quality parameters, which cannot be compensated in consecutive treatment steps. Since the water production rate is predicted, setpoint changes can be predicted as well. Ideally the control should take these predicted changes into account. Finally, the constraints of the reactor, such as maximal height and maximal dosing must never be violated.

Controller Configuration

A model-based multivariable controller is used to meet all requirements. A linear Model Predictive Controller (linear MPC) is used, since in this case calculation time is limited and valid solutions must be guaranteed. The information density in the process is insufficient to use a data-based model. The controller model is therefore obtained through numerical lineralisation of the white nonlinear model described in van Schagen et al. (2008a). The nonlinear model is linearised using the current bed composition found by the Model-Based Lane Control for the given reactor, and the current influent water quality parameters, water flow and caustic soda dosage.

Model predictive control is an online model-based optimal control technique based on the receding horizon principle. An online optimisation algorithm (normally a linear or quadratic programming algorithm) is applied to compute a series of control actions that minimizes a pre-defined cost function or 'performance index', subject to certain constraints. Applying the receding horizon principle means that only the first control sample is implemented and the horizon is shifted one time-step. Then the optimisation starts all over again. Figure 3 shows the principle of receding horizons graphically: $r(k)$, $y(k)$ and $u(k)$ are the reference, output and control (or manipulated) signals, N_m is the 'Minimum cost horizon', N_c is the 'Control horizon' and N the 'Prediction horizon'.

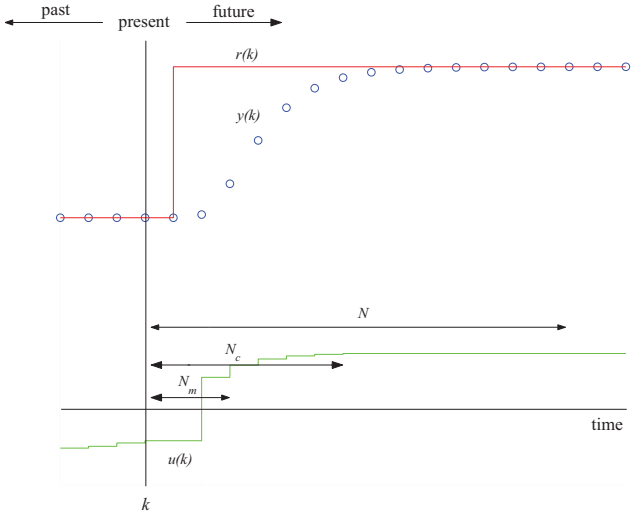


Fig. 3. The principle of linear model predictive control

At time instant k the system output is predicted from time step k until $k+N$ as a function of the control actions. Then the performance index is minimized resulting in an optimal control trajectory $\{u(k|k), ..., u(k+N_c-1|k)\}$.The outputs

from k until $k + N_m - 1$ are left out of the optimisation (to ignore minimum-phase and dead-time behaviour of the system) and the control actions are not allowed to change after time step $k + N_c - 1$.

Many different varieties of model predictive control configurations exist. The one chosen to implement for the pellet reactor controller is the so called 'Standard Predictive Control' (SPC) configuration (van den Boom and Backx, 2001). The advantage of this configuration is its flexibility and its state-space formulation.

The control objectives are to follow the current and future setpoints of the Model-Based Lane Control under smooth variation of the manipulated inputs, as formulated in the following cost function:

$$\begin{aligned} J = & \sum_{j=N_m}^{N} \|\mathbf{y}(k+j|k) - \mathbf{r}_y(k+j)\|_P^2 \\ & + \sum_{j=1}^{N} \|\Delta\mathbf{u}(k+j|k)\|_{Q_{\Delta u}}^2 \\ & + \sum_{j=1}^{N} \|\mathbf{u}(k+j|k) - \mathbf{r}_u(k+j)\|_{Q_u}^2 \end{aligned} \tag{1}$$

where N and N_m are the prediction horizon and the minimum costing horizon, and $\mathbf{r}_u$ and $\mathbf{r}_y$ are the references for the inputs and the outputs. In this way the control can use the setpoint predictions from the Model-Based Lane Control, due to predicted production rate changes.

The inputs are the caustic soda dosage and the water flow through the reactor. The outputs are the fluidised bed height in the reactor and the following water quality parameters in the effluent of the reactor: calcium concentration, pH, M-alkalinity and conductivity.

To meet the physical constraints in the process the linear MPC takes these constraints into account:

$$\begin{aligned} \mathbf{u}_{min} < \mathbf{u}_k < \mathbf{u}_{max} \\ \mathbf{y}_{min} < \mathbf{y}_k < \mathbf{y}_{max} \end{aligned} \tag{2}$$

To introduce extra integration action in the MPC controller, the model is modified to an IIO model. The new state vector consists of the previous output and the difference of the sate vector of the linearised model. The state update equation is now given by:

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_k \\ \mathbf{x}_{k+1} - \mathbf{x}_k \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{C} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{y}_{k-1} \\ \mathbf{x}_k - \mathbf{x}_{k-1} \end{bmatrix} + \\ \begin{bmatrix} D \\ B \end{bmatrix} (\mathbf{u}_k - \mathbf{u}_{k-1}) \end{aligned} \tag{3}$$

with the corresponding output function:

$$\mathbf{y}_k = [\,\mathbf{I}\ \mathbf{C}\,] \begin{bmatrix} \mathbf{y}_{k-1} \\ \mathbf{x}_k - \mathbf{x}_{k-1} \end{bmatrix} + D\,(\mathbf{u}_k - \mathbf{u}_{k-1}) \tag{4}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are the system matrices of the linearised model.

To compensate for plant-model mismatch an observer is used, to estimate the offset in $\hat{\mathbf{y}}_k$. The state update in the MPC controller is therefore given by:

$$\begin{aligned} \begin{bmatrix} \hat{\mathbf{y}}_k \\ \hat{\mathbf{x}}_{k+1} - \hat{\mathbf{x}}_k \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{C} \\ \mathbf{0} & \mathbf{A} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{y}}_{k-1} \\ \hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k-1} \end{bmatrix} + \\ \begin{bmatrix} D \\ B \end{bmatrix} (\mathbf{u}_k - \mathbf{u}_{k-1}) + \\ \begin{bmatrix} \mathbf{L} \\ \mathbf{0} \end{bmatrix} (\mathbf{y}_{k-m} - \hat{\mathbf{y}}_{k-m}) \end{aligned} \tag{5}$$

where $\mathbf{y}_{k-m}$ is the measurement result of m samples ago, due to the measurement delay.

A detailed explanation of the linear MPC algorithm is given in van den Boom and Backx (2001).

Simulation Results

To evaluate the performance of the controller, simulations were performed for the full-scale plant. The sample time for the controller was chosen to be 1 minute. The minimum cost horizon N_m, the control horizon N_c and the prediction horizon N are chosen to be 3,10,20 respectively, since the hydraulic retention time of the reactor is about 3 to 5 minutes. The setpoint for reactor flow and calcium concentration were taken from the lane controller. The simulation is started with a lane flow of 400 m^3/h, increasing the lane flow to 570 m^3/h, due to a production rate change after 1 hour. The reactor flow is kept constant and the bypass flow is increased. As a result from this flow change, the calcium concentration has to change from 50 to 35 mg/l. This is a regular change in calcium setpoint to produce constant water quality in the mixed effluent of reactor and bypass:

$$[\mathrm{Ca}^{2+}]_l = \frac{[\mathrm{Ca}^{2+}]_{in} F_{BP} + [\mathrm{Ca}^{2+}]_r F_{w,r}}{F_{w,l}} \tag{6}$$

Finally, if all lanes are operated at maximum capacity, the lane controller can increase the reactor flow for all reactors that are not yet limited by fluidised bed height. Therefore, in the simulation, the reactor flow is increased to 450 m3/h (the maximum flow for this reactor). The lane flow in this case is 640 m^3/h.

The operating point for the linearised model is the steady-state of the dissolved components in the nonlinear model with current estimated bed composition and the current influent flow and dosage. The states, which describe the bed composition ($\mathbf{m}_g$ and $\mathbf{m}_c$) are kept constant during numerical linearisation. The weighting matrices in equation 1 are diagonal, and the non-zero diagonal elements are given by:

$$\begin{aligned} P(\mathrm{Ca}^{2+}) = 0.1 \\ Q_u(F_w) = 1 \\ Q_{\Delta u}(F_w) = 1 \\ Q_{\Delta u}(F_s) = 0.1 \end{aligned} \tag{7}$$

The non-zero weights in P and Q_u penalise the deviation of the calcium concentration and water flow from their reference values. Change in the manipulated variables are penalised to achieve a smooth transition between operation points. In addition, level constraints are defined for all outputs and inputs, based on their physical ranges. To make the simulation more realistic, noise was added to

the simulated outputs. For the measurements of calcium and M-alkalinity the measurement noise was set at 2%, for bed height, pH and conductivity 1%.

The observer gain was chosen to be diagonal and the same for all measurements, since it is used to estimate model offset. The change in offset is expected to be equal for all measurements.

$$\mathbf{L} = \text{diag}\left([0.2\ 0.2\ 0.2\ 0.2\ 0.2]\right) \tag{8}$$

The simulation results using the nonlinear process model are shown in figures 4 and 5. In figure 4 the dashed-dot line is the setpoint for the calcium concentration, changing from 50 to 35 mg/l, due to a lane flow increase. The solid line is the simulated process values without measurement noise, while the dots are the actual measurement values available for the MPC controller. For calcium, M-alkalinity and conductivity, these measurements are only taken every 10 minutes, with a 10 minute delay. In the graph the measurements are therefore shifted by 10 minutes. The pH measurement and bed height measurements are online measurements and available every minute. The dashed line is output estimation $\hat{\mathbf{y}})_k$ of the MPC controller. In figure 5 the dashed-dot line is the setpoint for the reactor flow from the lane controller and the solid lines are the actual setpoints from the MPC controller.

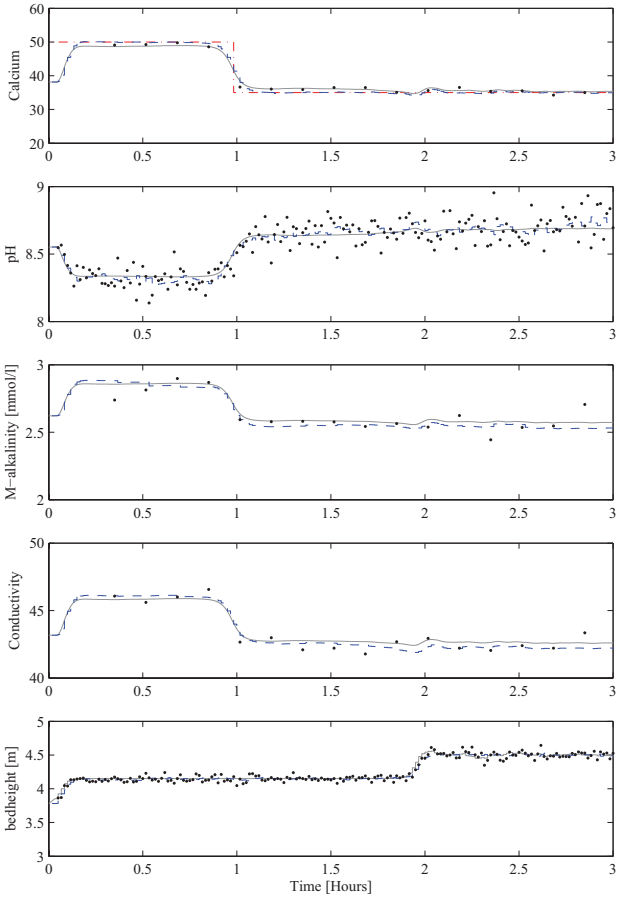


Fig. 4. Simulation results outputs. dashed-dot: Reference, dashed: Estimate, solid: Process, dots: Measurements

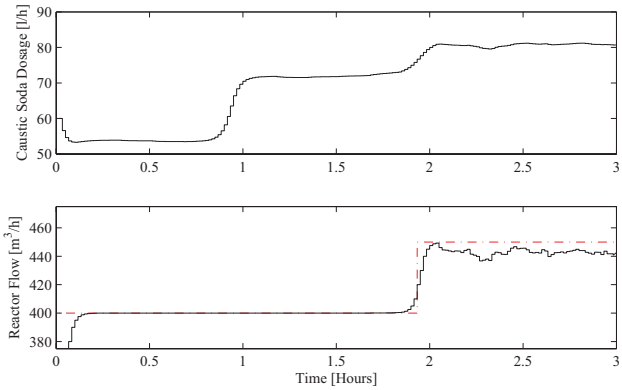


Fig. 5. Simulation results control inputs. dashed-dot: Reference, solid: MPC.

It can be observed, that the tracking of the reference signal is appropriate, including the desired smooth transition. The calcium concentration and the flow change starts before the actual setpoint change, as expected, to get a smooth transition close to the desired setpoint. Another interesting observation is that the water flow through the reactor and the caustic soda dosage are not strictly linked (as opposed to the current heuristic strategy). A flow reference change shows a rapid flow response, but a relatively slow dosage response, which results in a negligible change of the calcium concentration. Finally it can be seen that the MPC controller prevents a flow increase to the setpoint of 450 m^3/h, due to the limitation in bed height.

Pilot plant Results

The MPC controller is also implemented on the pilot plant of Weesperkarspel. The setpoints for the calcium concentration and reactor flow follow a similar pattern as in the full-scale reactor simulation. In this experiment the weighting matrices in equation 1 are diagonal, and the non-zero diagonal elements are given by:

$$\begin{aligned} P(\text{Ca}^{2+}) &= 3 \\ Q_u(F_w) &= 1 \\ Q_{\Delta u}(F_w) &= 0.01 \\ Q_{\Delta u}(F_s) &= 0.01 \end{aligned} \tag{9}$$

The matrices are selected to focus on setpoint achievement and less on smooth transition. The non linear model is the model from a validation experiment. The bed composition in this experiment is determined using the pressure drop measurement with different flows in the reactor. In the pilot-scale plant the pH measurement is not available as online measurement, and is determined semi-online during the M-alkalinity titration. The results from the pilot plant experiments are shown in figures 6 and 7.

The MPC controller in the pilot plant is performing as expected. The relatively small weighting matrix for control variations in equation 10 cause more variation in the caustic soda dosage and flow than for the full-scale simulation experiment.

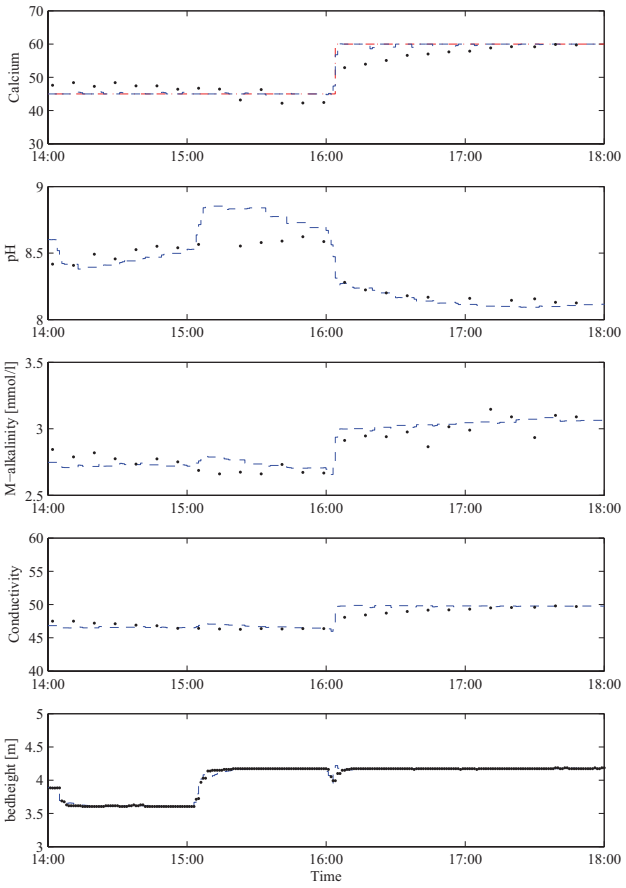


Fig. 6. Pilot plant experiment results outputs. dashed-dot: Reference, dashed: Estimate, dots: Measurements

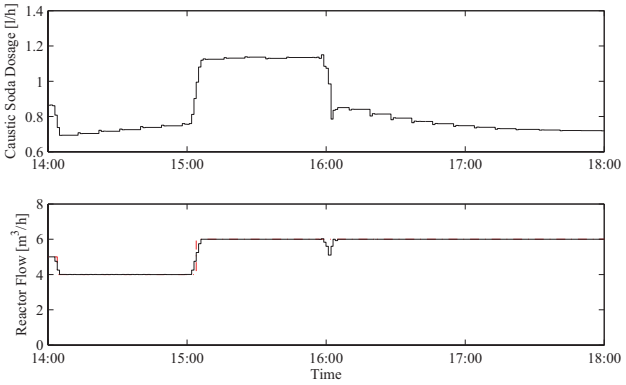


Fig. 7. Pilot plant experiment results control inputs. dashed-dot: Reference, solid: MPC.

CONCLUSIONS

The performance of the softening process step can be improved by applying a model-based control scheme. The control configuration is split in separate controllers for different control levels and time constants. To achieve smooth but quick responses to changing setpoints, a linear MPC is shown to be an effective controller.

A linear MPC controller shows a smooth transition between sudden changes of setpoints, while using a limited number of online and semi-online measurements. The controller is shown to function appropriately in the pilot-scale plant of Weesperkarspel.

REFERENCES

DHV (2008). http://www.aquasuite.com. Internet.

Rietveld, L. (2005). *Improving operation of drinkingwater treatment through modeling*. Ph.D. thesis, Faculty of Civil Engineering and Geosciences, Delft University of Technology.

van den Boom, T.J. and Backx, T.C. (2001). *Model Predictive Control*. TU Delft, Faculty of Information Technology and Systems. Lecture notes.

van Dijk, J. and Wilms, D. (1991). Water treatment without waste material - fundamentals and state of the art of pellet softening. *Journal of Water Supply: Research and Technology-AQUA*, Vol 40(5), 263–280.

van Schagen, K., Rietveld, L., Babuška, R., and E.Baars (2008c). Control of the fluidised bed in pellet softening process. *Chemical Engineering Science*, 63(5), 1390–1400. doi:10.1016/j.ces.2007.07.027.

van Schagen, K., Babuška, R., Rietveld, L., and E.Baars (2006a). Optimal flow distribution over multiple parallel pellet reactors: a model-based approach. *Water Science & Technology*, 53(4-5), 493–501.

van Schagen, K., Bakker, M., Rietveld, L., Veersma, A., and Babuška, R. (2006b). Using on-line quality measurements in drinking water process control. In *AWWA WQTC Conference*. Denver, USA.

van Schagen, K., Rietveld, L., and Babuška, R. (2008a). Dynamic modelling for optimisation of pellet softening. *Journal of Water Supply: Reasearch and Technology-AQUA*, 57(1), 45–56. doi:10.2166/aqua.2008.097.

Repetitive Control and Online Optimization of Catofin Propane Process

Wangyun Won*, Kwang Soon Lee*†, Seokho Lee, and Chansul Jung****

**Department of Chemical and Biomoleulcar Engineering, Sogang University,Seoul, Korea (Tel: +82-2-705-8477; e-mail: kslee@sogang.ac.kr)*
***Samgsung Engineering Co., Ltd., Seoul, Korea (e-mail: sh.yi@samsung.com)*

Abstract: The Catofin propane process is an emerging industrial process for propylene production through dehydrogenation of propane. It is composed of multiple adiabatic fixed-bed reactors which undergo cyclic operations where propane dehydrogenation and catalyst regeneration alternate over roughly ten minute period for each. One of the major concerns in the operation of the Catofin process is maintaining the reactor at an optimum condition while overcoming gradual catalyst deactivation. Addressing this issue, an online optimization of the Catofin process combined with a repetitive control has been investigated. The optimizer computes optimum initial bed temperatures for dehydrogenation and optimum air flow rate for regeneration, and the repetitive controller performs cycle-wise feedback action during regeneration to attain the target bed temperatures at the terminal time of the regeneration period. Numerical studies have shown that the proposed online optimizing control system performs satisfactorily coping with the catalyst deactivation and other disturbances.

Keywords: Catofin process, online optimization, repetitive control, adiabatic fixed-bed reactor

1. INTRODUCTION

Advanced control and online optimization are now accepted as an essential process intensification technology that can create an additional profit in process industries wherever they are applicable. During the past two decades or more, there have been numerous industrial projects for advanced process control alone or integrated with online optimization as reviewed in Qin and Badgwell (2003). Such projects have typically proceeded for continuous processes with linear MPC only or cascaded by online steady state optimization. While the continuous process with steady state operation represents the majority of the chemical processes, non-continuous processes such as batch, semi-batch, and repetitive processes also take an important part. Such processes are run under unceasing dynamics, which renders conventional advanced control and online optimization techniques to show limitations in the performance. In this research, a repetitive process called the Catofin propane process (ABB, 2008) has been addressed and an advanced control technique combined with online optimization that exploits the unique nature of the Catofin process has been investigated.

In this study, an on-line optimizing control system for the Catofin propane process has been proposed and investigated numerically. The optimizing control system is composed of two tiers, a repetitive controller cascaded by an online optimizer. Repetitive control is put into an action during the regeneration (RG) steering the bed temperatures at two axial positions to reach the target values at the terminal time of the RG period. The open-loop operation with only a state estimation is conducted during the dehydrogenation (DH). The optimizer calculates the optimum target values for the bed temperatures and the RG air flow rate under a cyclic steady state. Both repetitive control and online optimization were constructed on the basis of a first principle reactor model reduced to a set of ordinary differential equations (ODE's) using the cubic spline collocation method (CSCM) (Yun and Lee, 2007). For the repetitive control, the model is linearized before the start of each RG cycle around the operating trajectories in the previous cycle. The performance of the proposed optimizing control scheme has been investigated numerically.

2. PROCESS DESCRIPTION

Fig. 1 shows a simplified process flow diagram of the Catofin propane process. It consists of multiple parallel adiabatic fixed-bed reactors that contain Cr_2O_3/Al_2O_3 catalyst, where the DH of propane and RG of catalyst are carried out alternatively over roughly ten minute period each with short periods of purging and evacuation operations in-between.

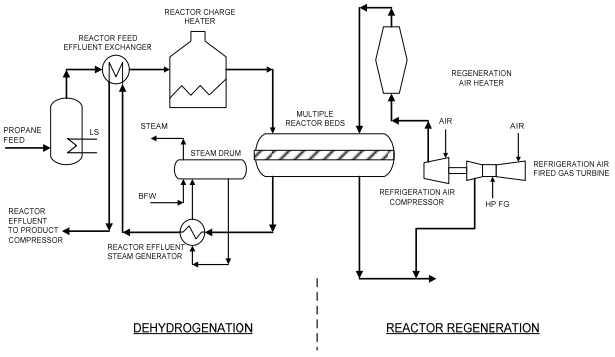


Fig. 1. Process flow diagram of the Catofin propane process.

The DH reaction is endothermic and produces a significant amount of coke. The bed temperatures are decreased and the catalyst loses activity by coke deposit and chromium reduction during this period. The RG reaction is coke burning by hot air and both the bed temperatures and catalyst activity are recovered under the oxidizing condition. The catalyst is known to have two years of life time and gradually loses the activity as the number of active sites is diminished by surface migration and agglomeration of Cr_2O_3 (Nijhuis, Tinnemans, Visser, and Weckhuysen, 2004).

The following apparent reaction kinetics proposed by Kim, Lee, and Song (1980) for the propane DH and Mickley, Nestor, and Gould (1965) and Pena, Monzon, and Santamaria (1993) for the coke combustion were assumed:

$$\begin{cases} C_3H_8 \underset{r_2}{\overset{r_1}{\rightleftharpoons}} C_3H_6 + H_2 \qquad C_3H_6 \xrightarrow{r_3} 3C + 3H_2 \\ C_3H_8 + C_3H_6 \xrightarrow{r_4} 2CH_4 + C_2H_6 + 2C \end{cases} \tag{1}$$

$$C + O_2 \xrightarrow{r_5} CO_2 \tag{2}$$

The rate constants are given in Table 1, which were slightly adjusted from the original values (Kim, Lee, and Song, 1980; Mickley, Nestor, and Gould, 1965; Pena, Monzon, and Santamaria, 1993) to more closely fit the conversion and yield of the real process (ABB, 2008).

Table 1. Parameters and normal operating conditions for the Catofin reactor model

Constants	Bed length = 1.5 (m), Bed diameter = 5.7 (m) C_c = 0.80 (kJ/kg•K), ρ_c=8x10^2 (kg/m^3), C_g = 3.71 for DH, 5.66 (kJ/kg•K) for RG D =1.7 for DH, 0.76 (m^2/min) for RG k_B=1.982 (kJ/min•m•K), R= 8.3462 (kJ/kmol K), DH and RG periods = 9min each
Normal operating condition for DH	Inlet temp=650℃, Propane flow=56.8(kmol/min), P=0.5 (atm)
Reaction rates (kmol/kg-cat.min) for DH	$r_1 = k_1[C_3H_8]RT$, $k_1 = 3.126*10^7 e^{(-47100/RT)}$ $r_2 = k_2[C_3H_8][H_2]R2T2$, $k_2 = 9.70*10^{-3} e^{(-12800/RT)}$ $r_3 = k_3[C_3H_6]R^2T^2$, $k_3 = 8.407*10^9 e^{(-62900/RT)}$ $r_4 = k_4[C_3H_8][C_3H_6]R^2T^2$, $k_4 = 9.498*10^5 e^{(-47800/RT)}$
Normal operating condition for RG	Inlet temp=690℃, Air flow=103.4 (kmol/min), P=2.0 (atm)
Reaction rate for RG	$r_5 = k_5[C][O_2]RT$, $k_5 = 4.129*10^3 e^{(-25575/RT)}$

It is assumed that the bed temperatures are measured at z = 0.2, 0.4, 0.6, 0.8, and 1.0, respectively, and the product gas compositions are available as the time average values over the DH and RG periods each with one cycle of measurement delay. It is also assumed that the RG is conducted under feedback control while the DH is carried out in an open loop state under a constant propane flow rate. The control objective during the RG is to steer the bed temperatures at z=0.2 and 0.4 to the target values provided by the optimizer using the RG air temperature as a manipulating variable (MV). The RG air flow rate was chosen as a decision variable for the optimizer together with the bed temperature target values.

3. REACTOR MODELLING

3.1 Mass and Energy Balances

In an adiabatic fixed-bed reactor, radial distribution of the concentrations and temperatures can be neglected. Under this assumption, the component mass and energy balance equations are written as

$$\begin{aligned} &\frac{\partial C_i}{\partial t} = \left(\frac{D}{L^2}\right)\frac{\partial^2 C_i}{\partial z^2} - \left(\frac{v_g}{L}\right)\frac{\partial C_i}{\partial z} + \left(\frac{1-\varepsilon}{\varepsilon}\right)\rho_c \bar{r}_i \\ &\frac{\partial C_c}{\partial t} = \rho_c \bar{r}_c \\ &\text{I.C. :} \quad C_i = C_i^I(z) \ \text{ at } \ t = 0 \\ &\text{B.C. :} \quad C_i = C_i^0(t) \ \text{ at } \ z = 0, \ \frac{dC_i}{dz} = 0 \ \text{ at } \ z = L_f = 5 \end{aligned} \tag{3}$$

$$\begin{aligned} &\frac{\partial T}{\partial t} = \left(\frac{k_B}{\rho_c c_{pc} L^2}\right)\frac{\partial^2 T}{\partial z^2} - \left(\frac{\varepsilon}{1-\varepsilon}\right)\left(\frac{\rho_g c_{pg} v_g}{\rho_c c_{pc} L}\right)\frac{\partial T}{\partial z} - \left(\frac{1}{c_{pc}}\right)\sum_j \Delta H_j r_j \\ &\text{I.C. :} \quad T = T_o(z) \ \text{ at } \ t = 0 \\ &\text{B.C. :} \quad T = T_o(\text{t}) \ \text{ at } \ z = 0, \ \frac{dT}{dz} = 0 \ \text{ at } \ z = L_f = 5 \end{aligned} \tag{4}$$

where $\bar{r}_i$ and C_i represent the rate of generation (kmol/kg-cat·min) and concentration (kmol/m^3) of component i, which refers to C_3H_8, C_3H_6, H_2, CH_4, C_2H_6 for DH operation, and CO_2, O_2 for RG operation, respectively; $\bar{r}_c$ and C_c represent the rate of generation (kmol/kg-cat·min) and concentration (kmol/m3) of coke, respectively; r_j referes to the rate of the j^{th} reaction; z denotes the normalized axial distance. Note that $\bar{r}_c$ for DH is different from $\bar{r}_c$ for RG. Other parameters and variables in the above model equations are given in Table 1.

In the above, the second boundary condition is specified at $z = 5$ instead of $z = 1$ whereas the spatial domain is $z \in (0, 1]$. The reason for this is to more reasonably represent the true phenomenon, $dT/dz \to 0$ as $z \to \infty$, using a condition at a distant axial position, which was named as the far-side boundary condition (Yun and Lee, 2007).

3.2 ODE Models by Cubic Spline Collocation Method

ODE models for the virtual process and nominal model were derived separately using the CSCM (Yun and Lee, 2007) using ten and five equally spaced collocation points over

(0,1] plus an additional point at $z = 5$, respectively. The resulting ODE models can be concisely written as

$$\frac{d\bar{x}_k^i}{dt} = \bar{f}^i(\bar{x}_k^i, u_k^i), \quad i = DH, RG \tag{5}$$

In the above, the subscript k denotes the cycle number; $\bar{x}^{DH}$ represents the state for the DH model, that consists of bed temperatures, concentrations of C_3H_8, C_3H_6, H_2, CH_4, C_2H_6, and C at the internal collocation points; $\bar{x}^{RG}$ is similar to $\bar{x}^{DH}$ except that the concerned chemical components are C and O_2; u denotes the MV and represents the air temperature T^{air} for *i*=RG and is void for *i*=DH, respectively.

3.3 Process Behavior under a Cyclic Steady State

Fig. 2 shows the bed temperature trajectories of the virtual process under a cyclic steady state at the nominal operating condition. The bed temperatures are initially increased as the higher bed temperatures in the fore part of the respective collocation points recede by the gas flow. After a while, however, bed temperatures are decreased by the endothermic reactions as the propane DH proceeds and restored again by the coke combustion during the RG. The amount for the coke deposit changes during this operation are as shown in Fig. 3. The coke generation is larger at the higher temperature positions and vice versa.

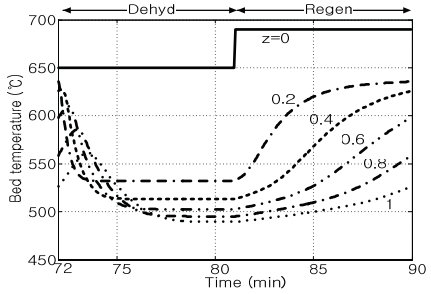


Fig. 2. Bed temperature trajectories at six axial positions under a cyclic steady state.

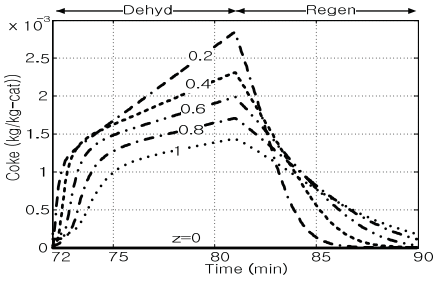
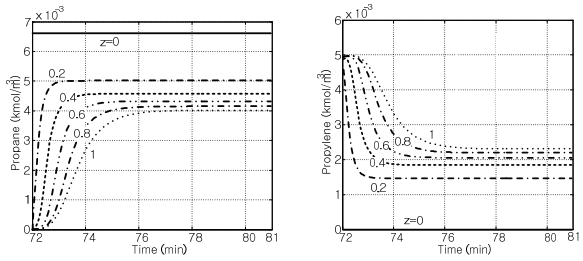


Fig. 3. Trajectories of coke deposit at six axial positions under a cyclic steady state.

The associated propane and propylene concentration trajectories during propane DH are shown in Fig. 4. Over an initial period while the bed temperatures are high, almost complete propane conversion and high propylene yield are obtained at the reactor outlet. As the bed temperatures begin to fall, both the propane conversion and propylene yield decrease. If we scrutinize Fig. 4, it can be seen that the front half of the bed where temperatures are higher than the rear half contribute more than 78.2% of the propylene production. The propane conversion and propylene selectivity averaged over a DH period are 51.5% and 86.2%, respectively.



(a) (b)

Fig. 4. Trajectories of propane and propylene concentrations under a cyclic steady state.

4. OPTIMIZING CONTROL SYSTEM

4.1 Structure

The optimizing control system consists of three major parts: the online cyclic steady state optimizer, the repetitive controller, and the model estimator. Fig. 5 shows the overall structure of the proposed system.

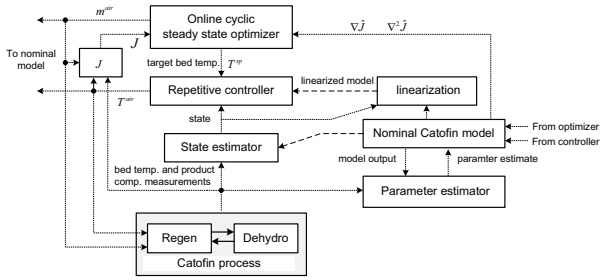


Fig. 5. Structure of the optimizing control system.

Fig. 6 illustrates the information flow through the state estimators along the operational sequence in more detail. The state estimation continues for the DH as well as RG periods based on the measurements of the bed temperatures and average product gas compositions in the previous cycle. Estimates of the coke deposit and bed temperature at the collocation points are transferred from the DH to RG and also from the RG to DH.

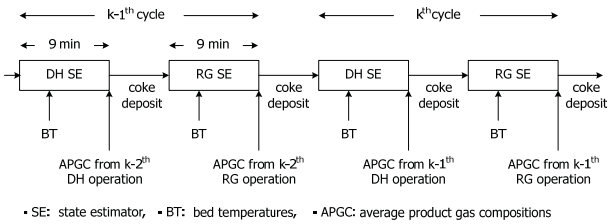


Fig. 6. Information flow along the sequence of operations.

4.2 Repetitive Control

4.2.1 Discrete-time Nominal Model

We first describe how the discrete-time nominal model for the state estimator and controller design is derived. The forward difference approximation applied to (5) results in

$$\bar{x}_k^i(t+1) = \bar{g}^i(\bar{x}_k^i(t), u_k^i(t)), \ \ i = \mathrm{DH}, \mathrm{RG} \tag{6}$$

The output equation can be written as

$$\begin{aligned} \bar{y}_k^i(t) &= V^i \bar{x}_k^i(t), \\ \bar{p}_k^i &= \frac{1}{N}\sum_{n=0}^{N-1} H^i \bar{x}_{k-1}^i(n), \ \ i = \mathrm{DH},\ \mathrm{RG} \end{aligned} \tag{7}$$

where $\bar{y}$ and $\bar{p}$ represent the bed temperatures at z=0.2, 0.4,…,1.0 and the average product gas composition measured at the end of the DH and RG periods with one cycle of measurement delay, respectively; N denotes the total number of sampling instance during the period of DH (or RG). V is a matrix that extracts the bed temperatures from the state and H is defined in a similar way for the compositions at the bed outlet. Hereafter, let us drop the superscript i for notational simplicity wherever there is no confusion.

The composition equation in (7) can be rewritten in the form of a state space equation. For this, let us define

$$\bar{p}_k(t+1) \triangleq \frac{1}{N}\sum_{n=0}^{t} H\,\bar{x}_{k-1}(n)$$

$$w_{1,k}(t) \triangleq \bar{x}_{k-1}(t), \ w_{2,k}(t) \triangleq \bar{x}_{k-1}(t+1), \cdots, w_{N,k}(t) \triangleq \bar{x}_k(t-1) \tag{8}$$

$$\bar{w}_k(t) \triangleq \begin{bmatrix} w_{1,k}(t) \\ w_{2,k}(t) \\ \vdots \\ w_{N-1,k}(t) \\ w_{N,k}(t) \end{bmatrix}, \ M \triangleq \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & I \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}, \ J \triangleq \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ I \end{bmatrix}, \ \bar{J} \triangleq \begin{bmatrix} I & 0 & \cdots & 0 & 0 \end{bmatrix}$$

Then the associated state transition equations are recast to

$$\begin{aligned} \bar{x}_k(t+1) &= \bar{g}(\bar{x}_k(t), u_k(t)) \\ \bar{w}_k(t+1) &= J\,\bar{x}_k(t) + M\bar{w}_k(t) \\ \bar{p}_k(t+1) &= \bar{p}_k(t) + \frac{1}{N} H\bar{J}\,\bar{w}_k(t) \end{aligned} \tag{9}$$

The resulting model equation can be rewritten in the following simplified form:

$$\begin{aligned} x_k(t+1) &= g(x_k(t), u_k(t)) \\ y_k(t) &= c(t)x_k(t) \end{aligned} \tag{10}$$

where

$$\begin{aligned} &x_k(t) \triangleq \begin{bmatrix} \bar{x}_k(t) \\ \bar{w}_k(t) \\ \bar{p}_k(t) \end{bmatrix}, \\ &y_k(t) = \bar{y}_k(t) \ \text{ and } \ c(t) = \begin{bmatrix} V & 0 & 0 \end{bmatrix} \quad \text{for } \ t = 1, \cdots, N-1 \\ &y_k(N) = \begin{bmatrix} \bar{y}_k(N) \\ \bar{p}_k(N) \end{bmatrix} \text{ and } \ c(N) = \begin{bmatrix} V & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \end{aligned} \tag{11}$$

Note that (10) and (11) holds for DH and RG separately.

4.2.2 Control Algorithm for RG Operation

The repetitive control conducts cycle-wise integral control action. To facilitate the construction of the control law, it is convenient to transform (10) to a state space model with $\Delta u_k(t) \triangleq u_k(t) - u_{k-1}(t)$ and $y_k(t)$ as the input and output variables, respectively. Linearization of (10) around the trajectories of the process variables in the k-1$^{\text{th}}$ cycle yields

$$\begin{aligned} \Delta x_k(t+1) &= A_{k-1}(t)\Delta x_k(t) + B_{k-1}(t)\Delta u_k(t) \\ y_k(t) &= y_{k-1}(t) + C_{k-1}(t)\Delta x_k(t) \end{aligned} \tag{12}$$

where $\Delta x_k(t) \triangleq x_k(t) - x_{k-1}(t)$; $A_{k-1}(t)$ represents a shorthand notation of $A_{k-1}(u_{k-1}(t), x_{k-1}(t\,|\,t))$, and similarly for $B_{k-1}(t)$ and $C_{k-1}(t)$. $\Delta u_k(t)$ is allowed to change P times at $t_1(=0)$, $t_2, \cdots, t_P$ during the RG period and determined at each time moment to satisfy the following quadratic prediction objective:

$$\begin{aligned} &\min_{\Delta u_k(\bullet)} \left[\left\| T_k^{sp} - \hat{y}_k(N\,|\,t_m) \right\|_Q^2 + \sum_{n=m}^{P} \left\| \Delta u_k(t_n) \right\|_{R(m)}^2 \right], \ m = 1, \cdots, P \\ &\text{subject to input constraints} \end{aligned} \tag{13}$$

At other occasions than t_m, $m = 1, \cdots, P$ $\Delta u(t) = 0$. In the above, $\hat{y}_k(N\,|\,t_m)$ represents a prediction of $\hat{y}_k(N)$, the bed temperatures at z=0.2 and 0.4, on the basis of the information up to t_m at the k^{th} cycle; T_k^{sp} denotes the target value of $\hat{y}_k(N)$. $\hat{y}_k(N\,|\,t_m)$ is given by the following form:

$$\hat{y}_k(N\,|\,t_m) = \hat{y}_{k-1}(N) + F_{k-1}(t_m)\Delta x_k^{RG}(t_m\,|\,t_m) + \sum_{n=m}^{P} G_{k-1}(t_n)\Delta u_k(t_n) \tag{14}$$

It is straightforward to derive (14) from (12). Note that the state estimate $x_{k-1}(t\,|\,t)$ and $\Delta x_k^{RG}(t_m\,|\,t_m)$ are needed to construct (12) (for linearization) and to solve (13) for $\Delta u_k(\bullet)$ (using (14)), respectively.

4.2.3 State Estimator

The state estimator is constructed separately for DH and RG in the form of the extended Kalman filter (EKF) for (10) and is given as

$$\begin{aligned} x_k(t+1\,|\,t) &= g\left(x_k(t\,|\,t), u_k(t)\right) \\ x_k(t\,|\,t) &= x_k(t\,|\,t-1) - K_k(t)\left(y_k(t) - c_k(t)x_k(t\,|\,t-1)\right) \end{aligned} \tag{15}$$

The observer gain $K_k(t)$ was obtained according to the EKF law using the process and measurement noise covariance matrices as the tuning factors. Using $x_k^{RG}(t_m\,|\,t_m)$ and $x_{k-1}^{RG}(t_m\,|\,t_m)$, $\Delta x_k^{RG}(t_m\,|\,t_m) = x_k^{RG}(t_m\,|\,t_m) - x_{k-1}^{RG}(t_m\,|\,t_m)$ for (14) was estimated. The state estimator acts as a fixed-lag smoother at t=N because the average product gas compositions are measured with one cycle of delay.

4.2.4 Implementation procedure

Over a DH-RG cycle, the following steps take turns in the repetitive control level:

[Step 1] DH period

$x_k^{DH}(t\,|\,t)$ is estimated for $t = 1, \cdots, N$ using (15).

[Step 2] Transition from DH to RG

Initialize $x_k^{RG}(1\,|\,0)$ by carrying over the coke deposit and bed temperature estimates in $x_k^{DH}(N\,|\,N)$ to $x_k^{RG}(1\,|\,0)$. Obtain the linearized model in (12) by linearizing (10) around $x_{k-1}^{RG}(t\,|\,t)$ and $u_{k-1}(t-1)$, $t = 1, \cdots, N$.

[Step 3] RG period

Perform the state estimation using (15). Compute $\Delta x_k^{RG}(t\,|\,t) = x_k^{RG}(t\,|\,t) - x_{k-1}^{RG}(t\,|\,t)$. Determine $\Delta u_k(t_m)$, $m = 1, \cdots, M$ according to (13) and (14). Implement $T_k^{air}(t) = u_k(t) = u_{k-1}(t) + \Delta u_k(t)$ to the process.

[Step 4] Transition from RG to DH

Initialize $x_k^{DH}(1\,|\,0)$ by transferring the coke deposit and bed temperature estimates in $x_k^{RG}(N\,|\,N)$ to $x_k^{DH}(1\,|\,0)$.

4.3 Online Cyclic Steady State Optimizer

The online optimizer determines T^{sp} and m^{air}, the target bed temperatures and the combustion air flow rate, respectively, that minimize the cost function under a cyclic steady state whenever the optimizer is invoked.

$$\begin{aligned}
\min_{T^{sp},\,m^{air}} \; & J = -\alpha_1 m^p Y_{css}^p + \alpha_2 \sum_{t=1}^{N} c_{pa} m^{air} \left(T_{css}^{air}(t) - T_{ref} \right) \\
& + \alpha_3 \left(\max(0, T_{z=0.2}^{sp} - 500) \right)^2 + \|s\|_Q^2, \;\; a_i > 0, \;\; Q > 0 \\
\text{subject to} \;\; & (10) \\
& \begin{bmatrix} 615 \\ 600 \end{bmatrix} \le T^{sp} \le \begin{bmatrix} 720 \\ 690 \end{bmatrix} (^\circ\text{C}) \\
& 10 \le \left| T_{z=0.2}^{sp} - T_{z=0.4}^{sp} \right| \le 30 \, (^\circ\text{C}) \\
& m_{\min}^{air} \le m^{air} \le m_{\max}^{air} \\
& s = x^{DH}(0) - x^{RG}(N)
\end{aligned} \tag{16}$$

where T_{ref}, m^p, and Y^p represent the reference temperature, propane mass flow rate, and average propylene yield over a DH period, respectively and the subscript *css* means the cyclic steady state. The summation is taken over the RG period.

The last term in J is to enforce the cyclic steady state condition, which is slackened by introducing a slack variable *s* defined the last equation in (16).

4.4 Model Parameter Estimator

In this study, the catalyst deactivation was assumed to be the most important process change and the parameter estimator was designed to update the pre-exponent rate constants by minimizing the following quadratic objective on the prediction error:

$$\min_{\theta^i} \; V^i = \frac{1}{k_2 - k_1} \sum_{k=k_1}^{k_2-1} \sum_{t=1}^{N} \left\| y_k^{m,i}(t) - y_k^i(t;\theta^i) \right\|_{\bar{Q}}^2, \; i = \text{DH, RG} \tag{17}$$

subject to $\theta_{\min}^i \le \theta^i \le \theta_{\max}^i$

where $y_k^{m,i}(t)$ and $y_k^i(t;\theta^i)$ represent the measurement and model prediction of $y_k^i(t)$ based on θ^i, respectively.

We devised a three parameter function as in (18), which is to be multiplied to each of the pre-exponent rate constants.

$$\xi(z) = \begin{cases} a + b\left(1 + \sin\left((-\pi/2d)(z+d)\right)\right), & z \in [0,1] \\ a + 2b \quad \text{if } 2d \le z \end{cases} \tag{18}$$

Since there are four rate constants for the DH, $\theta^{DH} \triangleq \left[a_1 \cdots a_4 \; b_1 \cdots b_4 \; d_1 \cdots d_4\right]$. Likewise, $\theta^{RG} \triangleq \left[a_5 \; b_5 \; d_5\right]$.

It is true that (18) is only a rough description of the catalysts activity distribution in the real process. Nonetheless, both the repetitive controller and the optimizer can achieve highly precise tracking as well as the true minimum, respectively, overcoming model uncertainties since the controller performs the cycle-wise integral action and the optimizer searches for the minimum on the basis of the process measurements.

5. SIMULATION CONDITIONS

The sampling period was chosen to be 3 sec resulting in total sampling instants of 360 with $t^{RG} = t^{DH} = 180$ over an entire cycle. The number control moments *P* was chosen as 3 and t_1, t_2, and t_3 were selected as 1, 60, and 120, respectively. The following constraints were imposed on the MV movements for repetitive control:

$$600 \le u(t) = T_{air}(t) \le 750 \, (^\circ\text{C}) \tag{19}$$

In the virtual process, the case of catalyst deactivation is represented by multiplying all k_i's by $1 - 0.5e^{-2.4z}$.

6. RESULTS AND DISCUSSION

The performance of the optimizing control system has been investigated for two cases. In the first case, the reactor was assumed to be initially at an arbitrary open-loop cyclic steady state and the optimizer steers the reactor to an optimum condition. In this case, the model parameter estimator was not invoked. In the second case, the optimum operation condition was assumed to be changed by catalyst deactivation, and the optimizer seeks for a new optimum condition from the previous operating condition determined in the first case. In the second case, the model parameter estimator plays an important role for both the repetitive controller and the online optimizer.

The simulation results for the first case are summarized in Fig. 7. It shows the response of the bed temperatures to their respective target values sent by the optimizer and the decrease of the objective function as the online optimization proceeds. The online optimizer calculates the new optimal target values once a cyclic steady state is reached on the basis of the nominal model and process measurements whereas the repetitive controller maneuvers the air temperature to attain the target values.

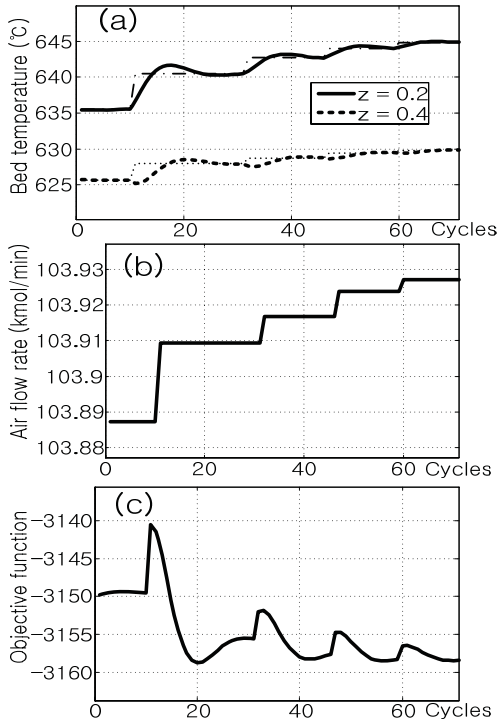
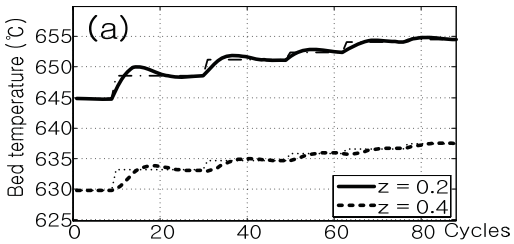


Fig. 7. Results of online optimization starting from an arbitrary open-loop state; (a) bed temperatures and their target values, (b) combustion air flow rate, (c) objective function.

The simulation results for the second case are given in Fig. 8. It can be seen that the overall responses are similar to Fig. 7. Unlike in the previous case, however, θ^{DH} and θ^{RG} were recurrently updated during the optimization. One thing to note is that the bed temperatures are raised even higher from the values determined in the first case to compensate for the catalyst deactivation.



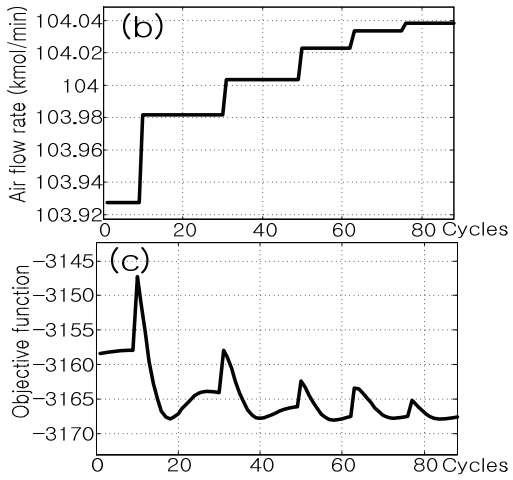


Fig. 8. Results of online optimization after the catalyst deactivation occurs; (a) bed temperatures and their target values, (b) combustion air flow rate, (c) objective function.

REFERENCES

ABB (2008). http://www02.abb.com/GLOBAL/NOOFS/noofs187.nsf/viewunid/494CDF435B2EE22DC12569EE0036977A/$file/CATOFIN.pdf

Bhasin, M. M., McCaina, J. H., Vora, B. V., Imai, T., & Pujad, P. R. (2001). Dehydrogenation and oxydehydrogenation of paraffins to olefins. *Applied Catalysis A: General*, 221, 397–419.

Kim, Y. G., Lee, H. S., & Song, Y. S. (1980). A study on dehydrogenation of propane. *Korean Journal of Chemical Engineering*, 18, 11-19.

Mickley, H. S., Nestor, J. W., & Gould, L. A. (1965). A kinetic study of the regeneration of a dehydrogenation catalyst. *Canadian Journal of Chemical Engineering*, 61-68.

Nijhuis, T. A., Tinnemans, S. J., Visser, T., & Weckhuysen, B. M. (2004). Towards real-time spectroscopic process control for the dehydrogenation of propane over supported chromium oxide catalysts. *Chemical Engineering Science*. 59, 5487-5492.

Pena, J. A., Monzon, A., & Santamaria, J. (1993). Coking kinetics of fresh and thermally aged commercial Cr2O3/Al2O3 catalyst. *Applied Catalysis A*, 101, 185-198.

Qin, S. J. & Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11, 733-764.

Yun, W. & Lee, K. S. (2007). The Use of Cubic Spline and Far-side Boundary Condition for the Collocation Solution of a Transient Convection-Diffusion Problem. *Korean Journal of Chemical Engineering*, 24, 204-208.

Model-based Control Design of a Diesel Oxidation Catalyst

Olivier Lepreux * **Yann Creff** * **Nicolas Petit** **

* *IFP Lyon, Technology, Computer Science and Applied Mathematics Division, BP 3, 69360 Solaize, France (e-mail: olivier.lepreux @ ifp.fr, yann.creff @ ifp.fr)*
** *Centre Automatique et Systèmes, Mathématiques et Systèmes, MINES ParisTech, 60, bd. Saint-Michel, 75272 Paris Cedex 06, France (e-mail: nicolas.petit @ mines-paristech.fr)*

Abstract:
This paper proposes a control strategy for a Diesel Oxidation Catalyst (DOC) which is grounded on a one-dimensional distributed parameter model. This first principles model for the propagation of the temperature variations accounts for spatially distributed heat generation (due to oxidation of reductants). As is discussed, heat generation can be regarded as equivalent inlet temperature variations. This fact is supported by experimental results. By nature, DOC outlet temperature response includes long and time-varying delays. An approximation of the proposed model allows to derive delays analytically, and can be used to schedule control parameters. As a consequence, it is easy to design several standard controllers for the DOC outlet temperature which account for the effects of the inlet temperature (disturbance) and the reductant (control). In this paper, simulation results are presented for a PI, a PID, and a Smith predictor. Interestingly, the three controllers use solely parameters determined from the previous analysis and do not need any extra tuning parameter. The strategies are tested on a standard NEDC driving cycle in simulation. It appears that, among these standard strategies, the DOC partial derivative equations can be efficiently controlled using the presented Smith predictor.

Keywords: Automotive exhaust aftertreatment systems, Diesel oxidation catalyst, Distributed-parameter systems, Boundary control, Control applications

1. INTRODUCTION

1.1 Motivation

On most new Diesel vehicles, increasing requirements regarding particulate matter emissions (Ecopoint Inc., 2008) are satisfied using a Diesel Particulate Filter (DPF). This filter, located in the vehicle exhaust line, stores particulate matter until it is burnt in an active regeneration process (Bisset, 1984). During this phase, DPFs behave like potentially unstable reactors (Achour, 2001), and their inlet temperature must be carefully controlled to prevent filter runaway.

In most current aftertreatment architectures (Koltsakis and Stamatelos, 1997), a Diesel Oxidation Catalyst (DOC) is placed upstream the DPF in the vehicle exhaust line. To increase the DPF inlet temperature, reductant is oxidized in the DOC, which, in turn, increases its outlet temperature. The DOC also conveys, up to some heat losses, its inlet enthalpy flow: in other words, inlet temperature variations propagate through the DOC.

A DOC is a chemical system difficult to control. Classical models are usually composed of a dozen of coupled partial differential equations (PDEs) (Depcik and Assanis, 2005), which complexify the development of model-based control laws. Experimentally, it can be observed that a step change on the inlet temperature propagates to the output of the system with long response times (Oh and Cavendish, 1982). Depending on the engine outlet gas flow rate, these response times significantly vary: they roughly decrease by a factor of 10 from idle speed to full load. Strategies that are commonly used to deal with this problem rely on look-up tables, which, in practice, are difficult and time-consuming to calibrate.

The purpose of this paper is to propose implementable control laws tuned according to a simple control-oriented model. This approach allows faster calibration. To achieve this goal, simplification of the above-mentioned classical models is needed.

After a presentation of a mathematical formulation of the control problem in the second part of this introduction, we show in section 2 how the model proposed in Lepreux et al. (2008), initially using inlet temperature as control variable, can be used to accurately describe actual cases of engineering interest, i.e. cases where the reductant flow is the control variable. Then, we show in section 3 how to approximate the model. Finally, in section 4, this approximation is used to tune several classic controllers. Simulation results serve as comparisons and stress that a Smith predictor tuned using the proposed methodology represents an efficient controller for the DOC.

1.2 Problem Formulation

It has been shown in Lepreux et al. (2008) that, considering only inlet temperature variations and neglecting chemical reactions, a DOC thermal behavior can be accurately described by the following model

$$\begin{cases} \dfrac{\partial T}{\partial t}(z,t) + v\dfrac{\partial T}{\partial z}(z,t) \;=\; -k_1(T(z,t) - T_s(z,t)) \\ \dfrac{\partial T_s}{\partial t}(z,t) \;=\; k_2(T(z,t) - T_s(z,t)) \end{cases} \tag{1}$$

with boundary control

$$T(z=0,t) = T^{in}(t)$$

where T and T_s are respectively gas and solid temperature variations about steady state, v is the channel gas speed which can be derived from mass flow, parameters (k_1,k_2) can be either derived from usual correlation (Osizik, 1977) or identified from experimental data (Lepreux et al., 2008). The output of the system is the outlet gas temperature

$$T^{out}(t) = T(z=L,t)$$

Considering steady-state initial conditions

$$\begin{cases} T(z,0) = 0 \\ T_s(z,0) = 0 \end{cases}$$

system (1) yields the transfer function

$$\hat{T}(z,s) = \hat{T}^{in}(s)\exp\left(-\frac{z}{v}s - \frac{k_1 z}{v} + \frac{m(z)}{s+k_2}\right) \tag{2}$$

where $m(z) = k_1 k_2 z/v$, $\hat{x}$ is the Laplace transform of x, and s is the Laplace variable. We denote Υ the Heaviside function and $\mathrm{I_i}$ the modified Bessel functions of the first kind. The system step response is

$$T(z,t) = \Upsilon(t - \frac{z}{v})\exp\left(-\frac{k_1 z}{v}\right)\times \left[1 + \int_0^{t-z/v} \exp\left(-k_2\tau\right)\sqrt{\frac{m(z)}{\tau}}\mathrm{I}_1(2\sqrt{m(z)\tau})d\tau\right] \tag{3}$$

For an easy evaluation of (3), a formulation using power series expansion is given in Lepreux et al. (2008).

2. REDUCTANT FLOW AS CONTROL VARIABLE

It is shown in Lepreux et al. (2008) that experimentally measured step responses of the system can be identified to model (1) with good quality. However, this representation might seem a bit simplistic in view of real applications since inlet temperature variations are difficult to control and cannot be used directly as control variable. In practice, reductants (hydrocarbons HC) are injected at the inlet of the DOC. They are oxidized on the catalyst and, consequently, increase the DOC temperature. In this section, we compute HC step response and compare it against T^{in} step response.

2.1 Model with Heat Source

During the regeneration process, the DOC is working at high temperatures, which ensures that the rate of conversion of reductants is high. Moreover, large quantity of HC is injected to generate exothermicity. Consequently, the inlet fraction of this reductant is very important, and its effect is dominating over other species'. By construction, a DOC is designed to yield large heat and mass transfer. These transfers are very effective, and the time scales implying the thermal phenomena are much lower than the ones implying chemical reactions. For the experiments presented in § 2.3 gas flows through the DOC approximatively 1000 times faster than the outlet temperature response time. For these reasons, to model the DOC thermal behavior, we propose to encompass all the chemical reactions in a "source term Ψ", leading to the following model

$$\begin{cases} \dfrac{\partial T}{\partial t}(z,t) \;+\; v\dfrac{\partial T}{\partial z}(z,t) = -k_1\left(T(z,t) - T_s(z,t)\right) \\ \dfrac{\partial T_s}{\partial t}(z,t) \;=\; k_2\left(T(z,t) - T_s(z,t)\right) + \Psi(z,t) \end{cases} \tag{4}$$

where $\Psi(z,t)$ is the control variable and $T^{in}(t)$ is regarded as a disturbance. Ψ includes the sum of the enthalpies of the various reactions taking place inside the DOC. We formulate a strong simplifying assumption. Namely, we assume that the rate of reaction is independent of the species concentration. Further, we also assume that it is independent of the temperature. In other words, Ψ is constant over some spatial interval. These assumptions are supported by experimental identification results of § 2.3. Over the whole range of considered operating conditions, the obtained results are quite accurate. We note L_c the length of the portion of the DOC where the enthalpy of reaction is released (see Fig. 1). Formally, we consider the

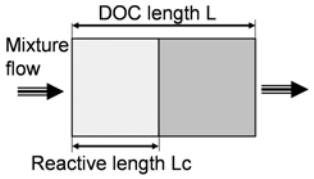


Fig. 1. HC reaction zone

following discontinuous function

$$\begin{cases} \hat{\Psi}(z,s) = \alpha/s, & 0 \le z \le L_c \\ \hat{\Psi}(z,s) = 0, & L_c < z \le L \end{cases} \tag{5}$$

Then, several steps of operational calculus on (4) lead to

$$\hat{T}(L_c,s) = \hat{T}^{in}\exp\left(-\hat{A}(s)L_c\right) + \frac{\hat{B}(s)}{\hat{A}(s)}\left(1 - \exp\left(-\hat{A}(s)L_c\right)\right) \tag{6}$$

with $\hat{A}(s) = \dfrac{1}{v}\left(s + k_1 - \dfrac{k_1 k_2}{s+k_2}\right)$ and $\hat{B}(s) = \dfrac{k_1}{v}\dfrac{\alpha}{s(s+k_2)}$.
In (6), the first term corresponds to the transfer from the inlet temperature $T^{in}(t)$ to the output $T(L_c,t)$, while the second term corresponds to the transfer from the input signal $\hat{\Psi}$ defined in (5) to the output $T(L_c,t)$. The linearity of the two effects will be used to study these phenomena separately in our control strategy. Further, for $z > L_c$, equation (4) gives $\hat{T}(z,s) = \hat{T}(L_c,s)\exp\left(-\hat{A}(z-L_c)\right)$ and, we get

$$\begin{aligned} \hat{T}(z,s) = \hat{T}^{in}\exp\left(-\hat{A}z\right) - \frac{\hat{B}}{\hat{A}}\exp\left(-\hat{A}z\right) \\ + \frac{\hat{B}}{\hat{A}}\exp\left(-\hat{A}(z-L_c)\right) \end{aligned} \tag{7}$$

Eventually, by an inverse Laplace transform of (7) (Abramowitz and Stegun, 1965), one obtains the reductant step response

$$\begin{aligned} T(z,t) = \Upsilon\left(t - z/v\right)\exp\left(-\frac{k_1 z}{v}\right)M(z,t-z/v) \\ - \Upsilon\left(t - z/v\right)\exp\left(-\frac{k_1 z}{v}\right)F(z,t-z/v) \\ + \Upsilon\left(t - z/v\right)\exp\left(-\frac{k_1 z}{v}\right)F(z-L_c,t-z/v) \end{aligned} \tag{8}$$

where

$$\begin{cases} M(z,t) = T^{in}(t) \\ \quad + \int_0^t \exp(-k_2\tau)\sqrt{\frac{m(z)}{\tau}}\mathrm{I}_1(2\sqrt{m(z)\tau})T^{in}(t-\tau)d\tau \\ F(z,t) = g(t) \\ \quad + \int_0^t \exp(-k_2\tau)\sqrt{\frac{m(z)}{\tau}}\mathrm{I}_1(2\sqrt{m(z)\tau})g(t-\tau)d\tau \\ g(t) = \frac{k_1\alpha}{k_1+k_2}t - \frac{k_1\alpha}{(k_1+k_2)^2}\left(1-\exp\left(-(k_1+k_2)t\right)\right). \end{cases}$$

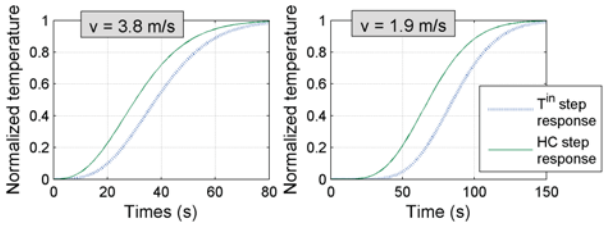
2.2 Fitting the heat source model with an equivalent no-source model

The static gain $G_{T^{in}}$ of the transfer from the inlet temperature T^{in} to the output $T(z,t)$ is equal to 1 (Lepreux et al., 2008). The static gain G_u of the transfer from the input signal $\hat{\Psi}$ to the output $T(z,t)$ can be computed using (7)

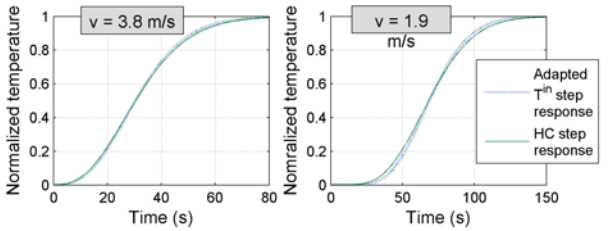
$$G_u = \lim_{t\to\infty} T(t) = \lim_{s\to 0} s\hat{T}(s) = \frac{k_1\alpha L_c}{k_2 v} \tag{9}$$

These last formulas are used during the identification and normalization process in this Section and for controllers design in Section 4. In practice, it is possible to relate α, L_c, the current HC conversion efficiency and the amount of injected reductants (which is itself related to the injector energizing time): for a given (identified) L_c, *α can be regarded as a control variable*.

In Fig. 2a, it is shown that the overall shape of the reductant step response, computed with (8), is very similar to the T^{in} step response (3). This similarity suggests that it is possible



(a) Comparison between HC step response and T^{in} step response. Analytic results obtained respectively from (8) and (3).



(b) Comparison between HC step response (model with source) and adapted T^{in} step response (model with no source)

Fig. 2. HC step response approximation in various operating conditions

to approximate HC step response by T^{in} step response at the expense of an additional identification procedure. We show in Fig. 2b that it suffices to adapt the DOC length, using the model with no source (1), to get results very close to the ones obtained with the model with source (4) [1]. In other words, generating enthalpy by HC is quite equivalent to generating enthalpy by T^{in} with a DOC having a shorter length. Hence, the temperature response of the DOC associated to the T^{in} variations are slower than those associated to HC. From a control point of view, this allows us to reject the T^{in} disturbance.

2.3 Experimental model validation

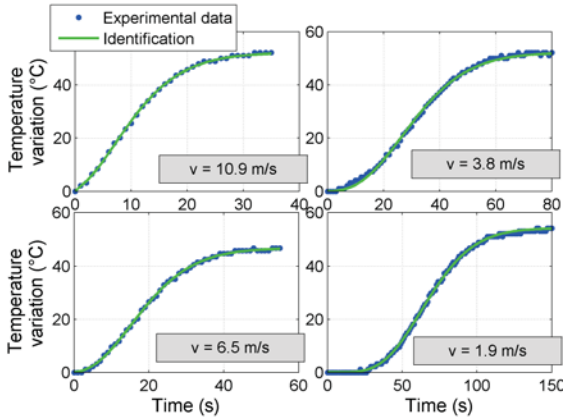


Fig. 3. Experimental HC step response identified to T^{in} model (1) in various operating conditions

As we stressed it in the previous discussion, considering an additional model adaptation of parameters (L_c is a piecewise linear function of v), model (1) and model (4) yield pretty similar results. In Fig. 3, we present experimental HC step responses under various operating conditions. To obtain these data, a 2.2-L 4-cylinder Diesel engine equipped with a 3-inch long 5.66-inch diameter 400-cpsi DOC was tested. These responses are well represented by the equivalent T^{in} step response (3) corresponding to the model with no source (1). It is shown that the model with no source kindly fits experimental data, usually described using a source term.

3. APPROXIMATING DOC EQUATIONS

We wish to simplify the previous model further. The desired representation is a first order plus delay model, which belongs to a class of models relatively easy to design a controller for (Silva et al., 2005).

For small values of $|s|$ (i.e. the range of low frequencies), the DOC transfer function (2) can be approximated in the following way

$$\begin{aligned} &\exp\left(\tfrac{m(z)}{s+k_2}\right) \\ &= \exp\left(-(1-\nu)\tfrac{k_1 z}{k_2 v}s\right)\exp\left(\tfrac{k_1 z}{v}\left((1-\nu)s/k_2 + \tfrac{1}{1+s/k_2}\right)\right) \\ &\approx \exp\left(-(1-\nu)\tfrac{k_1 z}{k_2 v}s\right)\exp\left((1-\nu)s/k_2 + 1 - s/k_2\right)^{\frac{k_1 z}{v}} \\ &\approx \exp\left(-(1-\nu)\tfrac{k_1 z}{k_2 v}s\right)\exp\left(\tfrac{k_1 z}{v}\right)\exp\left(-\nu s/k_2\right)^{\frac{k_1 z}{v}} \\ &\approx \exp\left(-(1-\nu)\tfrac{k_1 z}{k_2 v}s\right)\exp\left(\tfrac{k_1 z}{v}\right)\frac{1}{1+\nu\frac{k_1 z}{k_2 v}s} \end{aligned}$$

where $\nu \in]0,1[$ can be seen as a weighting variable which will be discussed later on. This leads to the following transfer function as an approximation of (2)

$$\exp\left(-\left(\frac{z}{v} + (1-\nu)\frac{k_1 z}{k_2 v}\right)s\right)\frac{1}{1+\nu\frac{k_1 z}{k_2 v}s}$$

[1] further details of this adaptation will be treated in a forthcoming publication

As a result, one obtains a delayed first-order step response $\frac{\exp(-\delta s)}{1+\tau s}$ where $\tau = \nu\frac{k_1 z}{k_2 v}$ and $\delta = \frac{z}{v} + (1-\nu)\frac{k_1 z}{k_2 v}$. The weighting variable ν relates τ and δ. Explicitly, we get $\delta = \frac{z}{v} + \frac{k_1 z}{k_2 v} - \tau$. In an identification standpoint, the next step is to determine a constraint to set the value of τ, which implicitly sets the value of ν.

Inflexion point The second-order derivative of the step response (3) with respect to time is

$$\frac{\partial^2}{\partial t^2}T(z,t+\frac{z}{v}) = -\Upsilon(t)\exp\left(-\frac{k_1 z}{v}\right)\exp(-k_2 t)\times \left[\left(k_2+\frac{1}{t}\right)\sqrt{\frac{m(z)}{t}}\mathrm{I}_1\left(2\sqrt{m(z)t}\right) - \frac{m(z)}{t}\mathrm{I}_0\left(2\sqrt{m(z)t}\right)\right]$$

Using the following asymptotic expansion of Bessel function (Abramowitz and Stegun, 1965)

$$\begin{cases} \mathrm{I}_\nu(z) \approx \frac{\mathrm{e}^z}{\sqrt{2\pi z}}\left(1 - \frac{\mu-1}{8z} + \frac{(\mu-1)(\mu-9)}{2!(8z)^2} - \dots\right) \\ \mu = 4\nu^2 \end{cases} \tag{10}$$

at first-order, we get, for $\frac{3}{16\sqrt{m(z)t}} \ll 1$,

$$\frac{\partial^2}{\partial t^2}T(z,t+\frac{z}{v}) \approx -\Upsilon(t)\exp\left(-\frac{k_1 z}{v}\right)\exp(-k_2 t)\times \left[\left(k_2+\frac{1}{t}\right)\sqrt{\frac{m(z)}{t}}\frac{\mathrm{e}^{2\sqrt{m(z)t}}}{\sqrt{4\pi\sqrt{m(z)t}}} - \frac{m(z)}{t}\frac{\mathrm{e}^{2\sqrt{m(z)t}}}{\sqrt{4\pi\sqrt{m(z)t}}}\right]$$

The equation of the inflexion point, of which t_I is the unknown abscissa, is given by $\frac{\partial^2}{\partial t^2}T(z,t_I) = 0$. With $t_{I'} = t_I - z/v$, this leads to $k_2 + \frac{1}{t_{I'}} - \sqrt{\frac{m(z)}{t_{I'}}} = 0$. Then,

$$t_{I'} \approx \frac{k_1 z}{k_2 v}\left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \frac{4v}{k_1 z}}\right) - \frac{1}{k_2}$$

and assuming $\frac{4v}{k_1 z} \ll 1$, we finally obtain $t_I \approx \frac{k_1 z}{k_2 v} + \frac{z}{v}$.

Parameters τ and δ Note f the delayed first order model step response $f(t) = 1 - \exp\left(-\frac{t-\delta}{\tau}\right)$. Let t_E be the solution of $f(t_E) = T(z,t_I)$. We impose the slope of f at abscissa t_E to equal the slope of T at t_I, i.e.

$$\frac{d}{dt}f(t_E) = \frac{\partial}{\partial t}T(z,t_I) \tag{11}$$

then, we get $\tau = \frac{1-T(z,t_I)}{\exp(-2k_1 z/v)k_2\mathrm{I}_1(2k_1 z/v)}$

To sum up, with the additional requirement (11), it is possible to write explicit values of τ and δ

$$\begin{cases} \tau = \frac{1-T(z,t_I)}{\exp(-2k_1 z/v)k_2\mathrm{I}_1(2k_1 z/v)} \\ \delta = \frac{z}{v} + \frac{k_1 z}{k_2 v} - \frac{1-T(z,t_I)}{\exp(-2k_1 z/v)k_2\mathrm{I}_1(2k_1 z/v)} \end{cases} \tag{12}$$

Typical identification results are presented in Fig. 4. Two different cases that are representative of real DOC parameter values, as motivated by Lepreux et al. (2008), are reported.

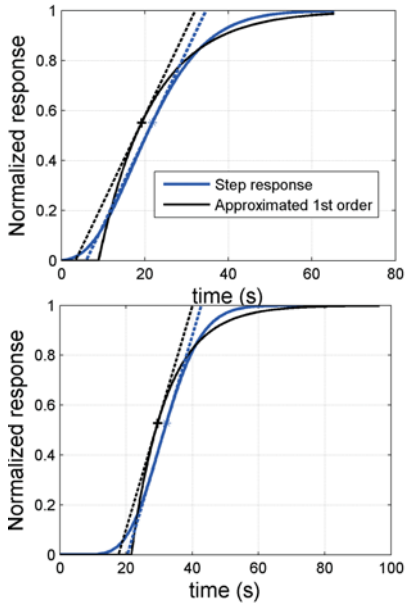


Fig. 4. Matching the DOC response with a first order plus delay model. k_1=400, k_2=0.35, v=4 (up). k_1=1591.09, k_2=0.82, v=4.597 (down).

Further approximation of τ and δ It has been shown that the choice of the constraint (11) leads to good matching of responses results. Further approximation can be made to prevent evaluation of the Bessel function. In experiment of Fig. 4, we get $\frac{3}{8}\frac{1}{2k_1 z/v} \ll 1$ for the two presented cases. Referring to (10), this validates the use an asymptotic expansion of I_1. We can make the approximation $\mathrm{I}_1(2k_1 z/v) \approx \frac{\exp(2k_1 z/v)}{\sqrt{2\pi 2k_1 z/v}}$. Then, we get the following expressions

$$\begin{cases} \tau = \frac{1}{k_2}(1 - T(z,t_I))\cdot 2\cdot\sqrt{\pi}\cdot\sqrt{\frac{k_1 z}{v}} \\ \delta = \frac{z}{v} + \frac{1}{k_2}\frac{k_1 z}{v} - \frac{1}{k_2}(1 - T(z,t_I))\cdot 2\cdot\sqrt{\pi}\cdot\sqrt{\frac{k_1 z}{v}} \end{cases} \tag{13}$$

It is interesting to note that, considering requirement (11), δ does not have an hyperbolic behavior with respect to v. Experimental results for evolution of τ and δ are shown in Fig. 5 (see also Frobert et al. (2009) for more complete results and details about the identification process). Corresponding analytical values are obtained using constant parameters k_1 and k_2, and L_c a function of v as mentioned in section 2.

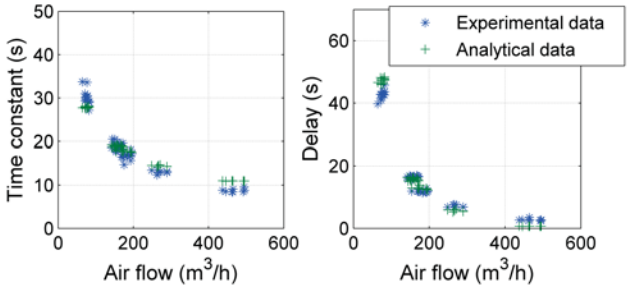


Fig. 5. Experimental evolution of δ and τ versus v

To sum up, the derived model is a combination of two first order plus delay models as shown in Fig. 6. The first one uses T^{in} as

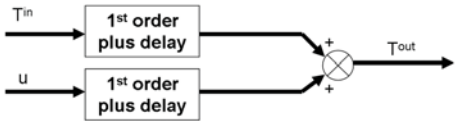


Fig. 6. Scheme of the first order plus delay model

input and $\tau_{T^{in}}$ and $\delta_{T^{in}}$ parameters are evaluated by (12) using the whole DOC length. The second one uses T^{in}-equivalent-to-Ψ as input, and corresponding τ_u and δ_u are evaluated by the same formula (12) using a part of the DOC length as explained in section 2. In both submodels, k_1 and k_2 are constant and equal.

Despite the fact that the proposed method of approximation does not allow to evaluate errors a priori, it is shown in simulations of subsection 4.2 that these methods provide accurate results for real cases of engineering interest such as those of a DOC used in driving conditions.

4. CONTROLLER PRESENTATION

4.1 Control Designs

Based on the results of section 3, we consider three classical control designs and evaluate their performance. The first two designs are simple PI and PID controllers with a feedforward term as presented in Fig. 7. We use respectively Tavakoli and Fleming (2003) and Tavakoli and Tavakoli (2003) parameters tuning for the PI and the PID controllers. Parameters τ and δ are evaluated using (12). The third controller, presented in Fig. 8,

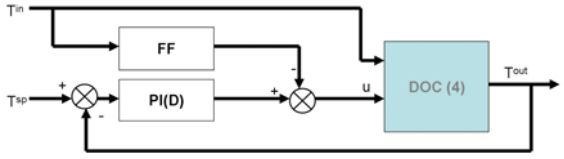


Fig. 7. Control scheme for the PI(D) controller

consists of a Smith controller (Abe and Yamanaka, 2003). For

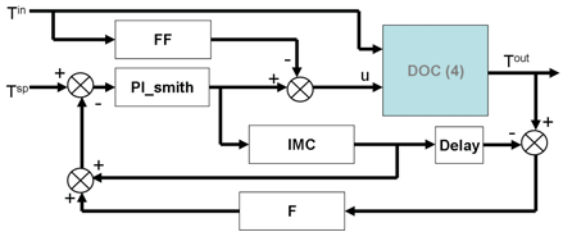


Fig. 8. Control scheme for the Smith controller

the three controllers, the gain G_u is calculated using (9). The FF block is dedicated to treating the T^{in} disturbance. Here, we use a standard feedforward strategy given by

$$FF = \frac{G_{T^{in}}}{G_u} \frac{\tau_u s + 1}{\tau_{T^{in}} s + 1} \exp\left(-\left(\delta_{T^{in}} - \delta_u\right) s\right)$$

Transfer functions for the Smith controller are given as follows

$$PI_{Smith} = \frac{1}{G_u}\left(1 + \frac{1}{\tau_u s}\right)$$

$$IMC = G_u \frac{1}{\tau_u s + 1}$$

The $Delay$ operator applies a delay of δ_u. The robustness filter F is a first order filter which time constant set to 1 s. It is not primordial here because, thanks to the presented detailed analysis of the DOC equations, delays are well approximated.

4.2 Simulation Results

First, we study the influence of a disturbance step variation. Then, we present control performance during a NEDC driving cycle. Simulation results are shown on the model with source (4) with $k_1 = 870$ s^{-1}, $k_2 = 0.45$ s^{-1}, $L_c = 0.0305$ m, $L = 0.0762$ m.

Basic performance Fig. 9 compares performance of the three controllers for a setpoint change. Setpoint is risen from 0 to 50 at $t = 200$. At the end of the rise, the system is disturbed by an important gas speed variation. These variations are directly linked to driver's power request, they are very fast and cannot be avoided. Although both controllers show similar tracking performance, the Smith controller shows much better disturbance rejection. Similar results are presented in Fig. 10

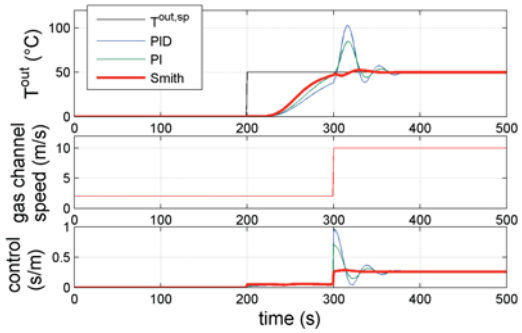


Fig. 9. Step setpoint transition and step v variation for PI, PID, and Smith controllers. $T^{in} = 0$.

with a 20% error on the k_1 parameter, implying important delay misestimation. All three proposed controllers are quite robust regarding this error.

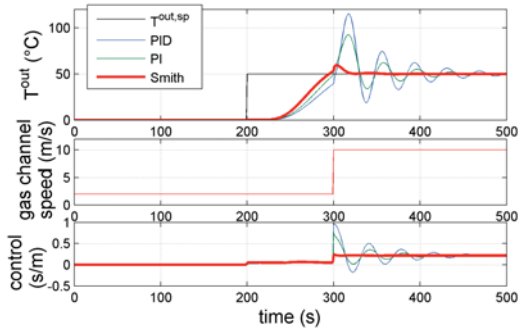


Fig. 10. Step setpoint transition and step v variation for PI, PID, and Smith controllers with a 20%-error on k_1. $T^{in} = 0$.

Performance on the NEDC cycle The three controllers are now tested on a simulated NEDC cycle. Results are presented in Fig. 11. In this case of a constantly-varying air flow rate, the differences between controllers are very small. Similar results are presented in Fig. 12 with a 20% error on the k_1 parameter. Once again, the presented controllers show good results regarding robustness on this fundamental parameter.

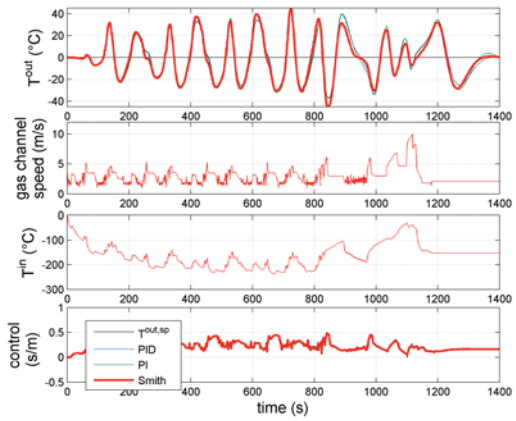


Fig. 11. PI, PID, and Smith controllers on NEDC cycle

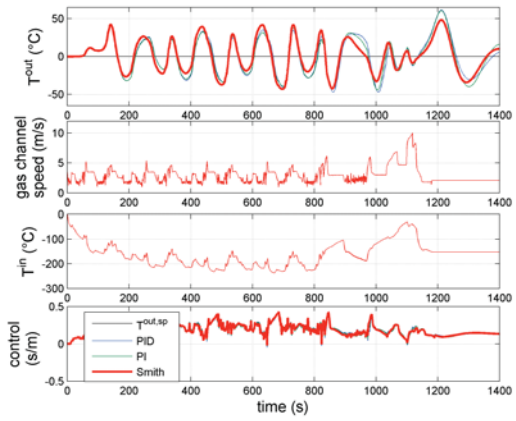


Fig. 12. PI, PID, and Smith controllers on NEDC cycle with a 20%-error on k_1

Conclusion The presented Smith controller requires more computational effort than the presented PI(D) controllers and shows some advantages in specific cases (setpoint transition, large air flow rate variation). It should be discussed based on further experimentations if it is necessary to use it or not for every specific case of application. It should be noted that its major drawback (lack of robustness toward a misestimation of the delay) has been circumvented thanks to the presented detailed analysis of the DOC equations.

5. CONCLUSIONS AND FUTURE WORKS

Control-oriented DOC modeling has been validated with experimental data in a former work (Lepreux et al., 2008). Grounding the current work on these equations and using several steps of approximation lead us, in this paper, to present a simple delayed first-order control-oriented model to approximate the DOC thermal behavior. This model is used as a starting point for control design. Interestingly, presented controllers require no particular tuning effort. They are tested in simulation on NEDC driving cycle. First, a simple PI or PID in which parameters are scheduled using developed approximations, reveal to be overreactive in certain specific cases. Alternatively, a Smith design shows good results and turns out to be fairly robust.

The presented experimental results allow a good level of confidence in our model. However, the next step is to present experimental results on the driving cycle using the proposed controllers.

A significant part of performance achieved by controllers presented in this paper, is due to the feedforward treatment of disturbances. To simplify the analysis, the presented controllers use classical feedforward control laws. However, going deeper into analysis, allows to achieve much better results. This feedforward control will be fully detailed in a forthcoming publication.

Being a part of an integrated system in the vehicle, the DOC is subject to whole class of reductants oxidizing onto it. It is worth properly estimating their flow, resulting from in-cylinder combustion, because it represents important disturbances for the DOC outlet temperature controller. This estimator is the subject of future works.

REFERENCES

Abe, N. and Yamanaka, K. (2003). Smith predictor control and internal model control - a tutorial. In *SICE 2003 Annual Conference*, volume 2, 1383–1387.

Abramowitz, M. and Stegun, I.A. (1965). *Handbook of mathematical functions*. Dover Publications.

Achour, L. (2001). *Dynamique et contrôle de la régénération d'un filtre à particules Diesel*. Ph.D. thesis, École des Mines de Paris.

Bisset, E. (1984). Mathematical model of the thermal regeneration of a wall-flow monolith diesel particulate filter. *Chem. Eng. Sci.*, 39, 1232–1244.

Depcik, C. and Assanis, D. (2005). One-dimensional automotive catalyst modeling. *Prog. Energy Combust. Sci.*, 31, 308–369.

Ecopoint Inc. (2008). DieselNet. Available online: http://www.dieselnet.com/.

Frobert, A., Creff, Y., Lepreux, O., Schmidt, L., and Raux, S. (2009). Generating Thermal Conditions to Regenerate a DPF: Impact of the Reductant on the Performances of Diesel Oxidation Catalysts. In *SAE International*. Detroit, USA.

Koltsakis, G. and Stamatelos, A. (1997). Catalytic automotive exhaust aftertreatment. *Prog. Energy Combust. Sci.*, 23, 1–39.

Lepreux, O., Creff, Y., and Petit, N. (2008). Motion planning for a Diesel Oxidation Catalyst. In *Proc. of American Control Conference 2008*. Seattle, USA.

Oh, S. and Cavendish, J. (1982). Transients of monolithic catalytic converters: response to step changes in feedstream temperature as related to controlling automobile emissions. *Ind. Eng. Chem.*, 21, 29–37.

Osizik, M. (1977). *Basic heat transfer*. McGraw-Hill.

Silva, G.J., Datta, A., and Bhattacharyya, S.P. (2005). *PID controllers for time-delay systems*. Birkhäuser.

Tavakoli, S. and Fleming, P. (2003). Optimal tuning of PI controllers for first order plus dead time/long dead time models using dimensional analysis. In *Proc. of the 7th European Control Conference*.

Tavakoli, S. and Tavakoli, M. (2003). Optimal tuning of PID controllers for first order plus time delay models using dimensional analysis. In *Proc. of the 4th International Conference on Decision and Automation (ICCA'03)*. Montreal, Canada.

Controller Design in a Fuel-Cell Powered Automobile

Srinivas Palanki[a 1] **John C. Telotte**[b]

[a] *Department of Chemical and Biomolecular Engineering, University of South Alabama, Mobile, AL 36688, USA*
[b] *Department of Chemical and Biomedical Engineering, Florida State University, Tallahassee, FL 32310-6046, USA*

Abstract: In this paper, the control problems that arise during dynamic operation of a fuel-cell powered automobile, are analyzed. In particular, it is shown that there are three distinct control problems that need to be solved when the power demand fluctuates. A logic-based switching controller is proposed that switches to the battery backup when the fuel cell is unable to provide the necessary power to the motor. An adaptive controller is developed based on a linear model that adjusts the hydrogen flow into the fuel cell in response to changing power demand. Finally, a thermal controller is developed based on a nonlinear model that regulates the temperature of the fuel cell. Interaction between these controllers is analyzed via simulations under realistic road conditions.

Keywords: controller design, PEM fuel cell, automotive power generation system

1. INTRODUCTION

Fuel cell power systems for automotive applications have received increased attention in recent years because of their potential for high fuel efficiency and lower emissions [Zalc and Loffler, 2002]. While there have been significant advances in fuel cell technology, one reason this technology has not seen wide-spread applications in the automotive industry has been the lack of an efficient hydrogen distribution center and the difficulties associated with storing hydrogen onboard an automobile. One option to alleviate these problems is to develop a system that utilizes a commonly available carbon-based hydrogenous fuel such as gasoline or methane to generate the necessary hydrogen *in situ* on an "as needed" basis. In this paper, we identify three separate control problems that need to be solved in a fuel-cell powered automobile.

2. SYSTEM DESIGN CONSIDERATIONS

A schematic of the fuel cell system under consideration is shown in Fig. 1. The two main components of the overall system are (1) the fuel processing subsystem and (2) the power generation subsystem. Methane enters the fuel processing subsystem and is converted to hydrogen. Hydrogen enters the fuel cell where it mixes with oxygen to generate electrical power which drives an electric motor. In addition to the fuel cell, there is a battery backup that the electric motor switches to when the hydrogen delivered to the fuel cell is insufficient to meet the *instantaneous* power demands of the electric motor. This battery backup is essential because significant load transitions occur frequently as a result of sudden acceleration on highway ramps as well as terrain changes [Zalc and Loffler, 2002].

[1] Corresponding author: palanki@usouthal.edu

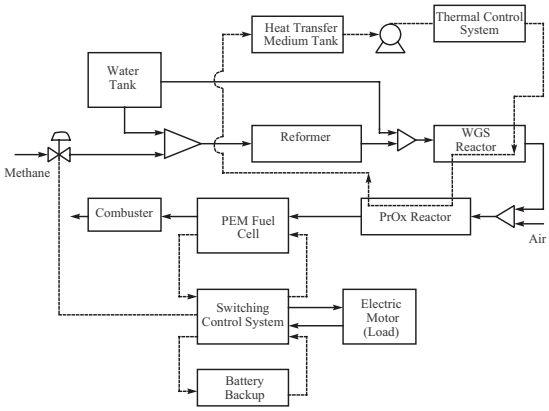


Fig. 1. Schematic of Fuel Cell System

In an earlier paper [Kolavennu *et al.*, 2006], the primary components of a fuel cell power system, that utilizes methane to generate hydrogen, were analyzed. In particular, basic chemical engineering principles were utilized to design a reactor train that converts methane to hydrogen of the desired purity. The relation between power produced by a PEM fuel cell and methane entering the reactor train *at steady state* was calculated. However, a typical automobile does not operate at steady state. The power demand for an automobile motor undergoes significant variations due to acceleration, changes in road surface and traffic conditions.

In this paper, we analyze the control problems that arise during dynamic operation of a fuel-cell powered automobile in the face of fluctuating power demand. In particular, it is shown that there are three distinct control problems that need to be solved when the power demand fluctuates. When power demand goes down, the excess hydrogen can

be diverted from the fuel cell. A sudden *increase* in power demand requires an instantaneous increase in hydrogen flow rate into the fuel cell. However, the conversion of methane to hydrogen takes several seconds which leads to an unacceptable lag between power demand of the motor and the power supplied by the fuel cell. For this reason, a backup battery is required that takes over this power load during the time it takes for the fuel cell to generate the necessary power and a suitable controller is required that switches between the fuel cell and the battery backup. If sufficient hydrogen is being produced by the fuel processor, a separate controller is required to adjust the hydrogen flow into the fuel cell in response to changing power demand. Finally, the fuel cell is subject to temperature changes and a thermal controller is required to regulate the temperature to the desired setting. In the paper, we propose the following controllers:

- A logic-based switching controller that switches to the battery backup when the fuel cell is unable to provide the necessary power to the motor.
- An adaptive controller based on a linear model that adjusts the hydrogen flow into the fuel cell in response to changing power demand.
- A thermal controller based on a nonlinear model that regulates the temperature of the fuel cell.

The design of these controllers is described in the sections below.

3. SWITCHING CONTROLLER DESIGN

There has been considerable research effort in modeling fuel cells [Nguyen and White, 1993]. In electric and fuel cell vehicles the battery is charged and discharged continuously and so knowledge of the transient behavior of the batteries is very important. Dynamic models developed from electrochemical principles like the cell sandwich model give spatial distribution of potentials and chemical compositions inside the cell as well as the transient behavior of cell potential and temperature. However for control oriented studies we require models which can be simulated quickly. Equivalent electric circuit models have been developed in the literature which give an accurate prediction of state-of-charge (SOC) of the battery [He and Hodgson, 2002].

He and Hodgson [2002] have observed that while discharging a battery over a period of time there exists a cutoff or critical voltage beyond which the battery performance deteriorates rapidly as the voltage begins to fall rapidly. To avoid operation near the critical voltage the state of charge is set to zero at the cutoff voltage and is defined as

$$SOC = 1 - \frac{Voc_{cutoff}}{Voc_{full}} \tag{1}$$

where Voc_{full} is the voltage of the battery at full capacity and Voc_{cutoff} is the battery terminal voltage at the critical point. Practically, it is difficult to measure the open circuit voltage at each instant hence utilizing the relationship between the SOC and the available battery capacity SOC can be redefined as

$$SOC = 1 - \frac{UsedCapacity}{TotalCapacity} \tag{2}$$

The total current drawn from the battery can be used as an indicator for the used capacity and is given by.

$$CAP_{used} = \int_0^t I.dt \tag{3}$$

So now the SOC is one when the battery is fully charged and zero when discharged to the critical voltage. It is desirable to maintain the SOC around 0.5-0.7.

A battery model which requires experimentally obtained open-circuit voltage and battery resistance data and predicts the battery terminal voltage, current, and SOC as a dynamic function of operator imposed power demand has been developed based on the model by He and Hodgson [2002]. The model consists of the battery as an ideal voltage source with an internal resistance. This battery model is characterized by the idealized open circuit voltage, V_{oc}, and the internal battery resistance, R_b. The terminal voltage can be expressed in terms of V_{oc} and R_b as

$$V_{term} = V_{oc} - I * R_b \tag{4}$$

The terminal voltage of a battery during discharge is lower than the instantaneous open circuit voltage because of the internal resistance inside the battery. Hence current I is given a positive sign when the cell is discharging. Similarly when the cell is charging we need to apply a voltage greater than the V_{oc} to overcome the internal resistance inside the cell so the current in this case is chosen to be negative.

The open circuit voltage and the internal resistance of the battery are both functions of SOC and temperature. For a battery operating at constant temperature the relationship between V_{oc}, Rb and the SOC can be determined experimentally.

The power available at the terminals of the battery is given by the product of voltage and current and substituting the expression for voltage from equation 4. we have

$$Pwr_{term} = V_{term}I = IV_{oc} - I^2 * R_b \tag{5}$$

For a particular power demand we can calculate the current by solving equation 5 which is a quadratic equation in I.

$$I = \frac{V_{oc} - \sqrt{(V_{oc}^2 - 4.R_b.Pwr)}}{2R_b} \tag{6}$$

where Voc and Rb are both functions of the SOC. The same sign convention as was used for the current is used i.e. the power is positive during discharge and negative during charge. The current calculated from eq. 6 is used to calculate the used capacity from eq. 3, which in turn is used to calculate the SOC by eq. 2. The V_{oc} and R_b are obtained for the new SOC from the experimental data. Using the new values of V_{oc} and R_b the current is estimated using eq. 6.

The switching controller is a logic based on-off controller that switches back and forth between the fuel cell and the battery to meet the power demand. As discussed earlier there is a time lag between the methane entering the reformer and the hydrogen coming out of the fuel processor. If the power demand remains constant the power produced by the fuel cell is sufficient to meet the power demand. The actual power demand curve is not a straight line and has a lot of fluctuations. To meet this fluctuating power demand, the fuel cell may switch to the battery. The switching controller has to address the following scenarios:

- *Increase in Power Demand*: Whenever there is an increase in power demand the fuel cell cannot produce the required power (P_r) because of the time delay (τ) in producing power and hence any deficit in power demand is handled by switching to the battery until the fuel cell can produce sufficient power. During this time delay the power produced by the battery is

$$P_{bat} = P_r - P_{fc} \quad for \quad t < \tau \tag{7}$$

- *Decrease In Power Demand*: During deceleration or decrease in power demand the fuel cell continues to produce the power requested until the time delay has elapsed. This excess power produced by the fuel cell during decrease in power demand should be routed to the battery, so that the battery can be charged. The same equation used in the scenario above can be used here. Since here the power requested is less than the power produced by the fuel cell the P_{bat} is negative which indicates that the battery is being charged.
- *State of Charge*: The state of charge of the battery should be always maintained above a specified target(SOC_{target}). But during sudden increase in power demand the battery might be discharged rapidly and the SOC might fall below the specified target and also the initial SOC itself might be less than the SOC_{target}. When the SOC of the battery falls below SOC_{target} the controller should direct the fuel cell to produce power to charge the battery in addition to the power demand.

$$P_{fc} = P_r + P_{bat} \quad if \quad SOC < SOC_{target} \tag{8}$$

- *Total Power Demand*: Since the fuel processor and the fuel cell system were designed for a maximum power output of 50 kW, the switching controller should make sure that the power demand from the fuel cell is not greater than 50 kW.

The fuel processor, fuel cell system and battery model along with the switching controller were simulated in MATLAB for different power demands. For a simple case where the power demand is a step increase followed by a step decrease the power profiles are given in Figure 2. Notice that the fuel cell supplies the power with a time delay of 4 seconds in the meantime the battery supplies the requested power demand. Once the fuel cell is able to meet the power demand the battery is turned off until 15 seconds at which time the battery again is used to supply the necessary power demand. At 30 seconds when there is a decrease in power demand the deficit power is sent to the battery to charge it until the fuel cell reaches the level of the new power demand.

To get a more realistic power vs time profile we obtained the power profile for a small car from an existing speed vs time profile using ADVISOR software package [NREL, 2002]. The Urban Dynamometer Driving schedule (UDDS) which is designed for light duty vehicle testing in city driving conditions has been used. The speed versus time profile is shown in Figure 3. The profiles of power requested, fuel cell power and battery power versus time are plotted in Figure 4.

The power supplied by the battery also depends on the initial SOC of the battery. For the same cycle the system was simulated for different initial SOC as shown in Figure 5. The controller was designed to maintain the SOC above

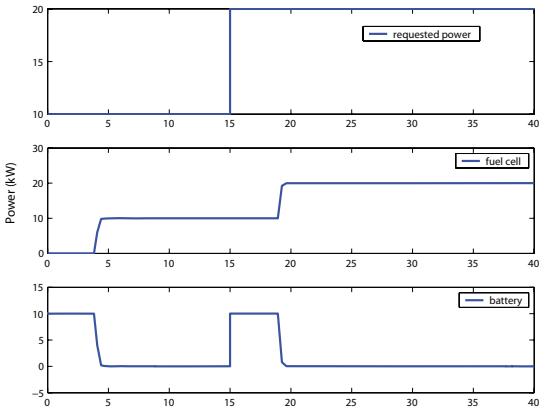


Fig. 2. Power profile

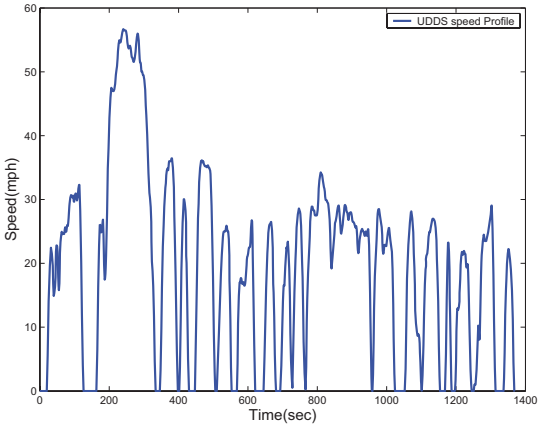


Fig. 3. Speed Vs time profile for UDDS

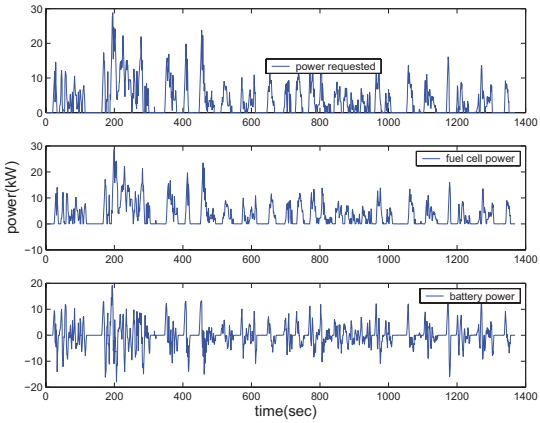


Fig. 4. Power profiles for the UDDS

0.5. For the initial conditions where the battery is almost charged (SOC=0.9) and semi charged (SOC=0.64) the profiles look similar. But for the case where the initial SOC is less than 0.5 the controller is activated and brings the SOC level to above 0.5.

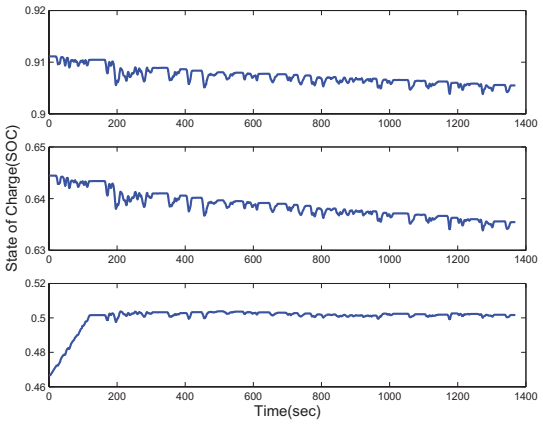


Fig. 5. State of charge for different initial conditions

Current battery technology in hybrid vehicles involves the use of nickel metal hydride (NiMH) battery packs. For instance, the Toyota Prius consists of 38 prismatic modules of a new generation NiMH design with a total pack nominal voltage of 273.6 V and a total energy capacity of 1.8 kWh [Kelly *et al.*, 2001]. The dimensions of this battery pack are 19.6 mm x 106 mm x 275 mm (volume of 57 l). The battery considered in this paper is of the same order of magnitude with a peak voltage of 300 V. If we assume that the energy capacity of the battery is 1.8 kWh (same as the Prius battery) and the battery has to have a state of charge of at least 50%, this battery would deliver 50 kW for 1 minute starting from a fully charged state before depleting to 50%. Thus, from a cold start, the reformer would have to be operational within 1 minute so that the car can switch from the battery to the fuel cell. It was shown in the previous section that once the fuel cell is operational, under realistic city driving conditions, the charge of the battery never goes under 50%. Newer gas-electric and fuel cell-electric hybrid vehicles use lithium ion battery technology. These batteries have superior power density versus energy density characteristics when compared to either NiMH batteries or supercapacitors. The second generation Honda Clarity fuel cell-electric hybrid is equipped with such a battery module. It is rated for 283 volts, and replaces the super capacitor energy storage system of the first generation Clarity. Improvement in lithium ion battery technology is ongoing. Current research indicates that energy capacity of 6-18 kWh are achievable with a calendar life of 15 years and 2500-5000 charge depleting cycles [Axsen *et al.*, 2008]. Preliminary research on magnesium ion battery systems suggest another order of magnitude in performance improvent is achievable [Axsen *et al.*, 2008].

4. ADAPTIVE CONTROLLER DESIGN FOR POWER GENERATION SUBSYSTEM

Pukrushpan [2003] developed and experimentally verified a dynamic model for a PEM fuel cell stack system similar to the one shown in Fig. 1. The model incorporates transient behavior that is important for controller design and analysis. In particular, a time-scale analysis of the various components was conducted and dynamic balances were developed for those operations that relate to automobile operations. Slower dynamics associated with temperature regulation and heat dissipation were ignored. Inertia dynamics along with nonlinear curve fitting of the compressor characteristic map were used to model the compressor. The manifold dynamics were based on lumped-volume filling dynamics. Static models of the air cooler and air humidifier were developed from thermodynamic relations. The fuel cell stack model was composed of four interacting submodels, namely stack voltage, cathode flow, anode flow and membrane hydration. The dynamic equations at the cathode and anode were developed using mass conservation principles and thermodynamic and psychometric properties of air. All gases were assumed to behave like an ideal gas. Spatial variations in temperature and concentration were ignored. It was assumed that the anode inlet flow rate could be instantaneously adjusted by a valve to maintain the minimum pressure difference between the cathode and the anode. Mass transport of water across the fuel cell membrane was calculated in membrane hydration model. Both water content and mass flow were assumed to be uniform over the surface of the membrane. However, this model developed by Pukrushpan [2003] consists of a large number of coupled, nonlinear differential and algebraic equations (DAE) and adaptive control theory cannot be applied directly to this system.After suitable substitution of variables, we obtained a reduced model of the fuel cell system that is a set of nine ordinary differential equations and is suitable for controller design and analysis [Kolavennu *et al.*, 2008]. In this model, it is assumed that all the cells in the stack perform similarly, i.e., by analyzing the polarization curve of a single cell, the stack performance can be estimated. The power from the fuel cell, which is a function of the current and voltage, is given by:

$$P = V_{st} I = (N_c V_c)(i A_c) \tag{9}$$

where P is the power produced by the fuel cell, V_{st} is the voltage of the stack which is the product of the number of cells N_c and the individual cell voltage V_c, I is the current drawn from the cell and is the same for each cell and depends on the area of cross section A_c, i is the current density.

The reversible standard potential E^o for the above cell reaction is 1.23 V at 25 ^oC as determined from the change in the Gibb's free energy. The actual voltage depends upon the concentration of the species and temperature at which the fuel cell is operating. The concentration dependence is given by the Nernst equation [Pukrushpan, 2003] as shown below:

$$\begin{aligned} E \;=\; & 1.229 - 8.5 \times 10^4 (T_{fc} - 298.15) + \\ & 4.3085 \times 10^{-5} T_{fc} \left[ln(P_{H_2}) - \frac{1}{2} ln(P_{O_2}) \right] \end{aligned} \tag{10}$$

where E is the open circuit voltage, the fuel cell temperature T_{fc} is in K, and reactant partial pressures P_{H_2} and P_{O_2} are expressed in atm. The actual cell voltage at any given current density is obtained by subtracting the activation, ohmic and concentration losses from the reversible potential as expressed below.

$$\nu_{fc} = E - \nu_{act} - \nu_{ohm} - \nu_{conc} \tag{11}$$

where ν_{act}, ν_{ohm} and ν_{conc} are activation, ohmic and concentration overvoltages. These losses are a function of the current density, pressure, membrane humidity and

also on the type of membrane and are represented by the empirical equations given below

$$\nu_{act} = \nu_0 + \nu_a(1 - e^{10i}) \tag{12}$$

$$\nu_{ohm} = i.R_{ohm} \tag{13}$$

$$\nu_{conc} = i\left(c_2 \frac{i}{i_{max}}\right)^2 \tag{14}$$

where ν_0, ν_a and c_2 are functions of temperature, pressure and membrane humidity of the cell. Using this model we can calculate the power produced by the fuel cell based on the voltage current characteristics. For a given current demand the voltage is calculated using Eq. 11 and thereby the power output of the fuel cell.

For the fuel cell systems to operate at levels comparable to existing internal combustion engines, the key issue that should be addressed is the *transient* behavior of fuel cell systems. Automobiles are subjected to significant load transitions during operation and the fuel cell system should be able to produce power which can follow this varying load profile. Power produced by the fuel cell is dependent on the voltage current characteristics. The transient response data from the nonlinear model presented in [Kolavennu *et al.*, 2008] was generated by subjecting the nonlinear system to a series of step inputs in the current around the 100 Amperes operating point. Utilizing this input output data from the nonlinear model system identification techniques were employed to derive a linear second order model was fit between the current demand and the voltage produced by the fuel cell stack. The transfer function G_p is given below

$$G_p = \frac{-390.78}{s^2 + 27.291s + 2068.8} \tag{15}$$

This transfer function is used in this paper to design an adaptive controller to regulate the power output of the fuel cell to the power demand. This adaptive controller is then implemented on the *nonlinear* model described in [Kolavennu *et al.*, 2008]. The control problem is to track the power demand of the motor using current as the manipulated variable.

To get a more realistic power vs time profile we obtained the power profile for a small car from an existing speed vs time profile using ADVISOR software package [NREL, 2002]. The Urban Dynamometer Driving schedule(UDDS) which is designed for light duty vehicle testing in city driving conditions was used.

Model reference adaptive control (MRAC) is derived from the model reference control (MRC) problem. The objective of MRC is to find the feedback control law that changes the structure and dynamics of the plant so that its I/O properties are exactly the same as those of a reference model. The structure of an MRC scheme for a LTI, SISO plant is shown in Fig. 6 [Ioannou and Sun, 1996]. Here, $W_m(s)$ is the transfer function of the reference model, $r(t)$ a given reference input signal, $y_m(t)$ the output of the reference model and $y(t)$ is the plant output. The feedback controller, denoted by $C(\Theta_c)$, is designed so that all signals are bounded and the closed-loop plant transfer function from r to y is equal to $W_m(s)$. This transfer function matching guarantees that for any given reference input $r(t)$, the tracking error $e = y - y_m$, which represents the

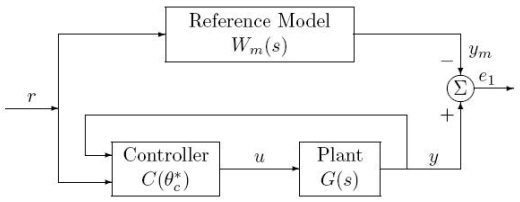


Fig. 6. Model Reference Adaptive Control

Table 1. Performance of MRAC on different road profiles

Profile or Cycle	ITAE error
UDDS	40.5
Federal Test Procedure	42.76
US06	55.13
Highway Fuel Economy Test	11.09
Extra Urban Driving Cycle	8.20
Indian Highway Profile	10.20

deviation of the plant output from the desired trajectory ym, converges to zero with time.

The model reference is chosen to be:

$$W_m = \frac{1}{s + 0.023} \tag{16}$$

The performance of the adaptive controller can be improved by adding some derivative action, i.e., using a PD controller in conjunction with the adaptive controller. This essentially makes the linearized plant represented by eq. 15 of unity relative degree which is the same as that of the reference model eq. 16.

The Environmental Protection Agency (EPA) reviews and revises as necessary the regulations governing the Federal Test Procedures (FTP) to insure that vehicles are tested under circumstances which reflect the actual current driving conditions under which motor vehicles are used, including conditions relating to fuel, temperature, acceleration, and altitude. The adaptive controller was tested on a variety of profiles. The controller was designed for the UDDS profile using the linearized model represented by eq. 15 and the same settings were employed for the remaining profiles. The resulting adaptive controller was implemented on the *nonlinear* model given in [Kolavennu *et al.*, 2008]. The Integrated Time Averaged Error (ITAE) was computed for each power profile. The results are shown in Table 1. It is observed that the adaptive controller with derivative action is able to track power profiles resulting from a wide variety of road conditions. A PID tuning procedure for the UDDS profile resulted in an ITAE error of 91.46. However, this controller when implemented on the US HWY profile resulted in loss of stability. On the other hand, the adaptive controller that was designed for the UDDS profile was able to successfully track the Federal Test Procedure profile, the US06 profile, the Highway Fuel Economy Test profile, the Extra Urban Driving Cycle profile, and the Indian Highway profile with no off-line tuning.

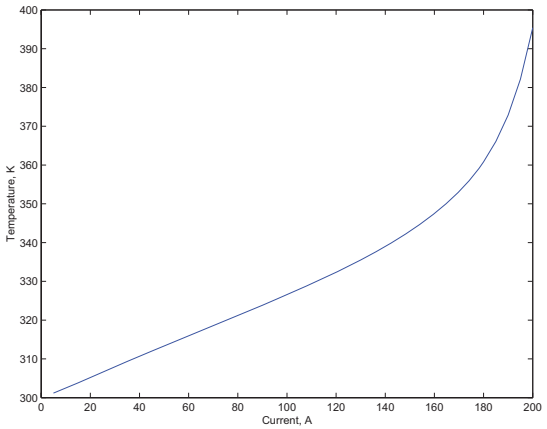


Fig. 7. Steady State Temperature versus Current in Fuel Cell

5. THERMAL CONTROLLER FOR FUEL CELL TEMPERATURE REGULATION

In this section, a dynamic model is developed that accounts for temperature changes in a PEM fuel cell. The dynamic model is obtained by extending a static current-voltage description to include temperature difference and by dynamically modeling the stack temperature. The following chemical reaction occurs in the fuel cell:

$$2H_2 + O_2 \rightarrow 2H_2O \tag{17}$$

The accumulation term for each species is negligible in the fuel cell compared to the mass of the fuel cell stack. Thus, a steady state model can be assumed for the gaseous and liquid species as follows:

$$\dot{N}_{i,out} = \dot{N}_{i,in} + \nu_i \dot{\xi} \tag{18}$$

where $\dot{N}_i$ is the molar flow rate of species i, ν_i is the stoichiometric coefficient and $\dot{\xi}$ is the reaction rate. It can be shown that this steady state assumption leads to the following dynamic balance for the temperature of the fuel cell stack:

$$m_{fc} C_{p_{fc}} \frac{dT_{fc}}{dt} = -C_p^* I(T_{fc} - T_\infty) - \Delta\tilde{H} I - hA(T_{fc} - T_\infty) - VI \tag{19}$$

where m_{fc} is the mass of the fuel cell stack, $C_{p_{fc}}$ is the specific heat of the fuel cell stack, C_p^* is the mole average specific heat of the reacting species, and T_∞ is the ambient temperature.

This provides a dynamic relation between the stack temperature, current and voltage. Fig. 7 shows the steady state relationship between temperature and current at a humidity of 50% in the cell. It is observed that the cell temperature increases nearly linearly with current until the design point (167 A), which is at optimum power output, and then increases rapidly due to cell inefficiency. The above equation was integrated numerically with realistic operating conditions of the fuel cell system and it was observed that the temperature dynamics are very slow compared to the dynamics of the fuel cell. In particular, it takes about 50 minutes to go from a cold start to the steady state temperature when the current is 150 A.

We are currently developing a nonlinear controller that utilizes the above model to regulate temperature to the desired set-point in the face of fluctuating power demand.

REFERENCES

Axsen, J., Burke, A., and Kurani, K., "Batteries for Plug-in Hybrid Electric Vehicles (PHEVs): Goals and the State of Technology circa 2008," *UCD-ITS-RR-08-14*, 2008

He, X. and Hodgson, J.W., " Modeling and Simulation for Hybrid Electric Vehicles: Modeling." *Intelligent Transportation Systems*, volume 3, 235-243, (2002)

Ioannou P. and J. Sun, *Robust Adaptive Control*, Prentice Hall, Inc., 1996

Kelly, K.J., Mihalic, M., and Zolot, M., "Battery Usage and Thermal Performance of the Toyota Prius and Honda Insight during Chassis Dynamometer Testing," *NREL/CP-540-31306*, 2001

Kolavennu, P.K., Telotte, J.C., and Palanki, S., "Design of a Fuel Cell Power System for Automotive Applications," *Int. J. Chemical Reactor Engng*, Vol 4: A19 (2006)

Kolavennu, P.K., Palanki, S., Cartes, D.A. and Telotte, J.C., "Adaptive Controller for Tracking Power Profile in a Fuel Cell Powered Automobile," *J. Process Control*, 18, 558-567 (2008)

Nguyen, T.V., and White, R.E., "A Water and Heat Management Model for Proton-exchange-membrane Fuel Cells," *J. Electrochem. Soc.*, 140, 8, 2178-2186 (1993)

National Renewable Energy Laboratory, "ADVISOR (Advanced Vehicle Simulator)", 2002

Pukrushpan, J. T., *Modeling and Control of Fuel Cell Systems and Fuel Processors*, PhD thesis, The University of Michigan, Ann Arbor, 2003

Zalc, J.M., and Loffler, D.G., "Fuel processing for PEM fuel cells: transport and kinetic issues of system design," *J. Power Sources*, 111, 58-64 (2002)

Fault Detection and Diagnosis

Oral Session

Fault Detection in Process Systems using Hidden Markov Disturbance Models

Wee Chin Wong * **Jay H. Lee** *

* *School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA*
(e-mail: {weechin.wong, jay.lee}@chbe.gatech.edu).

Abstract: Fault detection and diagnosis is critical for maintaining the health of process systems. Common fault signals include process and disturbance parameter changes, as well as sensor and actuator malfunctions (such as persistent drifts and biases). These may be characterized by the existence of latent 'fault' states. This work examines the effectiveness of a Hidden Markov Model framework for modeling such fault regimes. The proposed methodology may be interpreted as a generalization of a commonly-employed Mixture-of-Gaussians (Kesavan and Lee (1997)) approach and is demonstrated through a shell-and-tube heat exchanger problem. Furthermore, the flexibility of the method is shown in the context of detecting valve stiction. This is a significant problem in process industries where a valve's output suffers from excessive friction and is unable to track its input leading to degradation in closed-loop performance.

Keywords: Fault detection, Markov model, Sensor failure, Control valves.

1. INTRODUCTION

Tracking the closed-loop performance and health of process systems, although intuitively important, is oftentimes overlooked during the design of control solutions. Maintenance, required to mitigate the effects of system faults, typically necessitates expert personnel not found within normal plant situations (Kesavan and Lee (1997)). For this reason, multiple process monitoring algorithms have been developed so that such faults may be automatically detected, diagnosed and eventually removed.

Process monitoring methods may be further classified as i) data-driven ii) analytical and/ or iii) knowledge-based (Chiang et al. (2001)). The first involves statistical treatment of large quantities of process data and are typified by data-mining and machine learning techniques (such as principal and independent component analysis), statistical control charts and so on. Knowledge-based methods employ qualitative reasoning and are oftentimes rules-based with a strong logic underpinning. A thorough overview of all three classes is presented by Chiang et al. (2001) and the references therein.

This work, relying on dynamical models of the process for fault detection, is a particular type of analytical approach. Consequently, a necessary standing assumption is the availability of a mathematical model derived from first principles or otherwise. Given the wide-spread popularity of model-based control (such as Model Predictive Control), the controller's model can be readily ported over for the purpose of fault-detection. A model structure, such as in (1), (2), is therefore relevant in subsequent developments.

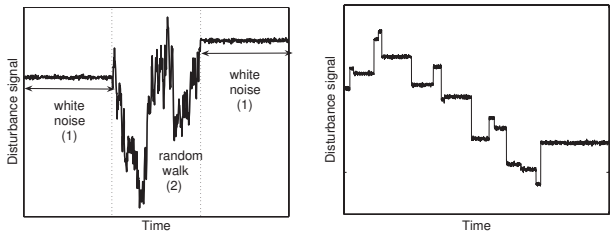
$$\begin{aligned} x_t &= f(x_{t-1}, \theta_{t-1}, u_{t-1}, \omega_t) \\ y_t &= g(x_t, \theta_t, v_t) \end{aligned} \tag{1}$$

$$\begin{aligned} \gamma_t &= \mathbb{A}_t \gamma_{t-1} + \mathbb{B}_t \varphi_t \\ \theta_t &= \mathbb{C}_t \gamma_t + e_t \end{aligned} \tag{2}$$

Here, $x_t \in \mathbb{R}^{n_x}$ represents the state at discrete time index t, $u_t \in \mathbb{R}^{n_u}$, the control input, and $y_t \in \mathbb{R}^{n_y}$, a noise-corrupted measurement signal. $\theta_t \in \mathbb{R}^{n_\theta}$ represents a fault vector with potentially time-varying dynamics governed by matrices $(\mathbb{A}_t, \mathbb{B}_t, \mathbb{C}_t)$ and noise vectors (φ_t, e_t) [1]. ω_t and v_t are process and measurement noise signals respectively. $f(\cdot)$, which may represent an integration of the continuous-time model over a unit sample-time, is the state transition map. Similarly, $g(\cdot)$ represents the state-to-output map.

Faults are typically manifested (Kesavan and Lee (1997)) as i) process parameter changes, and/ or ii) disturbance parameter changes, as well as iii) actuator and sensor problems – all captured by θ. Depending on circumstances, these may be sudden jumps (*e.g.* due to an abrupt introduction of significant sensor bias), or slow drifts or random walk-type changes (*e.g.* as a result of catalyst fouling) or even a mixture of both (Fig. 1). Such failure modes, which cannot be directly observed, and need to be estimated, are conveniently incorporated into the fault model (2) by adding the notion of latent states (denoted by r), each of which modifies the fault model (see (2)) differently. This work explores the use of a Hidden Markov chain, used previously to model realistic disturbances in the context of process control (Wong and Lee (2007)), to describe the temporal, probabilistic transitions between the latent states. Furthermore, this work can be interpreted as a generalization of the popular approach of assuming statistical independence, from one time period to the next, between hidden states. For example, at each time step t, Willsky (1976) and Kesavan and Lee (1997) allowed the statistics

[1] In practice, the user would model θ according to disturbance scenarios of interest.



(a) Intermittent drifts: white noise probabilistically interspersed with integrated white noise.

(b) Abrupt jumps.

Fig. 1. Possible disturbance signals (θ).

of φ_t and e_t to be described by a Mixture-Of-Gaussians (MOG)[2]. This captures the situation where faults that do occur happen infrequently but with significantly larger magnitudes. Persistent faults like drifts, which are easily described by the proposed Hidden Markov Model (HMM) approach, are captured in the MOG context by introducing additional states or non-linearities in the model.

The main contribution is to show that the aforementioned faults (abrupt jumps/ biases and drifts) can be better modeled and detected by the proposed method. Another novel application is in the context of detecting valve stiction, where it is demonstrated that the output of the valve (which is not normally measured) can be effectively tracked using the same proposed framework.

Section 2 provides the details behind an HMM, its subsequent use for fault detection and relevance to prior work. Section 3 demonstrates the effectiveness of the proposed method in the context of a heat exchanger. Section 4 explores the valve stiction issue before concluding remarks regarding future research are presented in Section 5.

2. FAULT MODELING USING HIDDEN MARKOV MODELS

HMMs represent a useful class of statistical models where a latent state, taking values from an alphabet $\mathcal{J} \in \{1, 2, \ldots, J \in \mathbb{Z}^+\}$ of cardinality J, transitions probabilistically in a Markovian[3] fashion from one sampling time to the next. Mathematically, a finite-state Markov chain is a sequence of random variables $(r_0, r_1, ..., r_t, ...)$, where the transition probability matrix $\Pi = (\pi_{ij}) = (pr(r_t = j|r_{t-1} = i), i, j \in J) : \sum_{j=1}^{J} \pi_{ij} = 1, \forall i \in \mathcal{J}$, governs the probabilistic temporal transitions. The term 'Hidden' signifies that the actual regime label is usually not known with complete certainty and must be inferred from available noisy measurements of itself or other related states. In the simplest case, each latent state has a probability distribution over a finite set of possible output symbols. All Markov chains under consideration are ergodic. For simplicity, the Markov chain is assumed to be at steady state, satisfying $\pi = \Pi'\pi$, where π is a column vector containing the unconditional and initial probabilities of each regime. HMMs have found widespread applications in science and engineering - ranging from speech recognition (Rabiner (1989)) to bioinformatics and diverse fields such as econometrics.

HMMs and their generalizations have been used in fault detection, with significant differences to our proposed approach. Smyth (1994), for example, did not consider an explicit fault model (*i.e.,* (2)). Instead, the process parameters are continuously estimated (in batch mode) and treated as output of an underlying Markov chain. This necessitates linking the process parameter vector to fault modes, which is not always possible. A recursive maximum a posteriori filter is then used for fault-mode detection. Huang (2008) suggested a similar (see Section 2.1) HMM approach to sensor problem diagnosis but limited considerations to faults in the output channels and input signals taking values from a finite, discrete set. Almeida and Park (2008) learned an HMM corresponding to each operating condition and, unlike the approach proposed in this work, does not make use of the process model. There, fault detection is achieved by a classification scheme that chooses the HMM that maximizes the probability of a given sequence of observations.

2.1 Proposed Fault Model: Intermittent Drifts & Abrupt Jumps

Following the successes in other fields, a generalization of (2) is considered by allowing the statistics of (φ_t, e_t) (and potentially the fault model parameters $(\mathbb{A}, \mathbb{B}, \mathbb{C})$) to vary according to a hidden Markov chain.

Intermittent Drifts. In the case of one-dimensional intermittent drifts (Fig. (1a)), one has:

$$\begin{aligned} \gamma_{t+1} &= \gamma_t + \varphi_{r_{t+1}} \\ \theta_t &= \gamma_t + e_t \\ r_t &\in 1, 2 \\ \pi_{11} &\approx 1,\ \pi_{11} < 1 \\ \pi_{22} &\approx 1,\ \pi_{22} < 1 \end{aligned} \tag{3}$$

Here, φ_{r_t} and e_t are uncorrelated, zero-mean Gaussian signals with covariances (that may depend on r_t) of $Q^{\varphi}_{r_t}$ and Q^{e}_{t}. The abuse of notation on the subscript of φ emphasizes the dependence of the covariance of the noise signal on the underlying Markov chain. When $r_t = 1$ (*i.e.,*, the white-noise regime), $Q^{\varphi}_{r_t=1} \approx 0$. Random-walk type behavior occurs when the hidden state switches to $r_t = 2$, where $Q^{\varphi}_{r_t=2} >> 0$; Q^{e}_{t} is invariant to the hidden regime and of appropriate magnitude. Since it is common that there is low probability of switching once the system enters a particular regime, a diagonally-dominant Π is employed.

Abrupt Jumps. In the case of modeling abrupt jumps, (3) is adjusted such that $\pi_{11} = \pi_{12} = p \approx 1$, $p < 1$, so that $\Pi = [p,\ 1-p; p,\ 1-p]$. This ensures that the jump state (the second one, in this case) is infrequently accessed and when it is, a significant step-change occurs.

In this latter case, since it is assumed that the Markov chain is at steady state, this form of the transition matrix implies that the probability of entering a particular regime is independent of the current mode. It is thus clear that the HMM framework subsumes an MOG description.

[2] *i.e.,* at each time step, a member from a set of Gaussians, from which the noise signal is to be sampled, is selected with some time-invariant probability.

[3] transitions depend only upon the immediate past.

Fault detection and diagnosis is performed via state estimation (in particular to track θ) without the knowledge of the latent state trajectory. Hence, a brief mention of state estimation, based on a model resulting from the concatenation of (1), and (2) is necessary.

2.2 Fault Detection via State Estimation of Jump Markov Systems

Equations (1) and (2) can be merged to yield:

$$\begin{aligned}\begin{bmatrix} x_{t+1} \\ \gamma_{t+1} \end{bmatrix} &= \mathcal{F}_{r_{t+1}}\Big(\begin{bmatrix} x_t \\ \gamma_t \end{bmatrix}, u_t, \xi_{r_{t+1}} \Big) \\ y_t &= \mathcal{G}_{r_t}\Big(\begin{bmatrix} x_t \\ \gamma_t \end{bmatrix}, n_{r_t} \Big) \\ pr(r_t = j | r_{t-1} = i) &= \pi_{ij} \end{aligned} \tag{4}$$

Here, $\mathcal{F}$ is implicitly understood to include model structures and parameters from $\{f, \mathbb{A}, \mathbb{B}, \mathbb{C}\}$ and the hidden Markov chain. A similar remark is extended to $\mathcal{G}$. Besides $\mathcal{F}$ and $\mathcal{G}$, the statistics of the noise ξ (a concatenation of (ω, φ, e)) and n (a concatenation of (v, e)) can depend on r. The system represented by (4) is also termed a Markov jump system. Without knowledge of the sequence $(r_0, \ldots, r_t)$, the optimal filter involves averaging over an exponentially growing number of linear filters. The number of filters scales as J^t, where J is the cardinality of the set containing all possible realizations of r.

The following paragraphs outline the Generalized Pseudo Bayesian estimation algorithm of order 2 (GPB2), a popular sub-optimal method, developed by Bar-Shalom and Li (1993). The main idea to have trajectories whose last 2 terms differ be merged (via moment-matching) into a single Gaussian. Using the law of total probability and Bayes' Rule, it can be shown that:

$$\begin{aligned} x_{t+1|t+1} &= \sum_{r_{t+1}} p(r_{t+1}|t+1) x_{t+1|(t+1,r_{t+1})} \\ x_{t+1|(t+1,r_{t+1})} &\triangleq \sum_{r_t} x_{t+1|(t+1,r_{t+1},r_t)} p(r_t|r_{t+1}, t+1) \\ P_{t+1|t+1} &= \sum_{r_{t+1}} \{(x_{t+1|t+1} - x_{t+1|(t+1,r_{t+1})})(\cdot)' \\ &\quad + P_{t+1|t+1,r_{t+1}}\} p(r_{t+1}|t+1) \\ P_{t+1|t+1,r_{t+1}} &= \sum_{r_t} \{(x_{t+1|t+1,r_{t+1}} - x_{t+1|(t+1,r_{t+1},r_t)})(\cdot)' \\ &\quad + P_{t+1|t+1,r_{t+1},r_t}\} p(r_t|r_{t+1}, t+1) \\ p(r_t|r_{t+1}, t+1) &= \frac{1}{c_1} p(y_{t+1}|t, r_{t+1}, r_t) p(r_t|r_t) p(r_t|t) \\ p(r_{t+1}|t+1) &= \frac{1}{c_2} \sum_{r_t} p(y_{t+1}|t, r_{t+1}, r_t) p(r_{t+1}|r_t) p(r_t|t) \end{aligned}$$

The term $p(y_{t+1}|t, r_{t+1}, r_t)$ refers to the probability density of the corresponding one-step ahead output prediction. $x_{t+1|(t+1,r_{t+1})}$ refers to the estimate of x_{t+1} given output measurements $\{y_0, \ldots, y_{t+1}\}$ and a certain realization of r_{t+1}; $P_{t+1|(t+1,r_{t+1})}$ denotes the corresponding error covariance matrix. The pair $(x_{t+1|(t+1,r_{t+1},r_t)}, P_{t+1|(t+1,r_{t+1},r_t)})$ are similarly defined. It is noted that starting from $(x_{t|(t,r_t)}, P_{t|(t,r_t)})$, a single application of the time and measurement update steps of the (extended) Kalman filter yields these latter quantities. c_1 and c_2 are normalizing constants such that the merging probabilities $p(r_t|r_{t+1}, t+1)$ and $p(r_{t+1}|, t+1)$ sum to unity.

2.3 A-posteriori Regime Estimation

If required, a prediction and/ or filtered estimate of the hidden regime can be obtained viz:

$$\begin{aligned} \hat{r}_{t+1|t} &= \arg\max_{r_{t+1}} \left\{ p(r_{t+1}|t) \triangleq \sum_{r_t} pr(r_{t+1}|r_t) \cdot pr(r_t|t) \right\} \\ \hat{r}_{t|t} &= \arg\max_{r_t} \{p(r_t|t)\} \end{aligned} \tag{5}$$

3. EXAMPLE 1: FAULT TRACKING IN A SHELL & TUBE HEAT EXCHANGER

In this example, the usefulness of the proposed method in detecting faults is studied in the context of a shell and tube heat exchanger (6) also considered by Kesavan and Lee (1997). In particular, we contrast the proposed HMM approach against an MOG method (Kesavan and Lee (1997)) in modeling the latent states that govern the fault signals (see Section 3.1 for simulation details). The main difference is that the latter framework assumes that each latent state occurs with a (time-invariant) probability that is independent of the previous realization. The governing non-linear ordinary differential equations used for simulation but not estimator design, are:

$$\begin{aligned} \frac{dT_c}{dt} &= \frac{q_c}{V_c}(T_{ci} - T_c) + \frac{\alpha_c}{V_c}(T_h - T_c) \\ \frac{dT_h}{dt} &= \frac{q_h}{V_h}(T_{hi} - T_h) - \frac{\alpha_h}{V_h}(T_h - T_c) \\ y &= \begin{pmatrix} T_c \\ T_h \end{pmatrix} + \mu_v + v \end{aligned} \tag{6}$$

Here, the measured state variables are the temperatures of the hot and cold streams respectively: $[T_c; T_h]$. $[T_{ci}; T_{hi}]$ are the temperatures of the incoming cold and hot streams respectively. $[\alpha_c; \alpha_h]$ are system parameters reflecting the heat transfer coefficient, heat transfer area, density, specific heat capacity of the cold and hot streams respectively. Similarly, $[q_h; q_c]$ are the flow rates of the hot and cold streams and represent the degrees of freedom available to a controller. $[V_c; V_h]$ are the volumes of the cold and hot sides. Steady-state values are reported in Table 1. v refers to zero-mean measurement noise of covariance $R \triangleq \mathbb{E}[vv']$. μ_v is nominally a null vector but might be subject to changes due to disturbances.

3.1 Simulation Conditions

Although a variety of fault types may be considered (*e.g.* those affecting the various input and output channels and/ or changes in parameters (α_c, α_h), as discussed in Section 1), for clarity of exposition, only two different fault types are assumed. Furthermore, these affect only the cold side. Given initially quiescent conditions (see Table 1), one considers:

(1) An abrupt step that is normally distributed with zero mean [L/min] and variance q_u^{hi} [L^2/min^2] affecting the input channel on the cold side (q_c) at some unknown time t_u. This may be thought of as a sudden bias developing in the input channel:

$$q_{ct} = q_{ct-1} + \varphi_t^u \cdot \delta(t, t_u),\ \varphi_t^u \sim \mathcal{N}(0, q_u^{hi}) \quad (7)$$

$\delta(\cdot,\cdot)$ is the Dirac delta function. q_u^{hi} has a value of 2 in the following experiments.

(2) A sudden drift (see Fig. 1a) affecting the sensor relaying T_c (*i.e.,* y_1) measurements between an unknown time span: $\mathcal{T} \triangleq [t_{y,1}, t_{y,2}]$. Namely, one has:

$$\mu_{v,1_t} = \mu_{v,1_{t-1}} + \varphi_t^y \quad (8)$$

where $\mathbb{E}[\varphi_t \varphi_t'] = q_y^{hi} = 0.5$ if $t \in \mathcal{T}$ and $\mathbb{E}[\varphi_t \varphi_t'] = q_y^{lo} = 10^{-10} \approx 0$, for other time periods. $\mu_{v,2}$ remains at the origin for all time.

The above non-linear model is not available for state estimation. Instead, a version linearized about the nominal operating conditions is available. With a sampling time of 0.5 min, A = [0.91, 0.03; 0.03, 0.91], B = [-0.12, 0.002; -0.002, 0.12], $C = diag([1, 1])$. Measurement covariance, R, is set to $diag([0.5, 0.5])$ and known. Since estimation is the focus of this example, the system is run in the absence of feedback control.

3.2 Proposed HMM Method to Handle Abrupt Jumps & Intermittent Drifts

The following Markov jump linear model, a specialization of (4), is employed:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + b\theta_t^u + \omega_{t+1} \\ \theta_{t+1}^u &= \theta_t^u + \varphi_{r_{t+1}}^u \\ \theta_{t+1}^y &= \theta_t^y + \varphi_{r_{t+1}}^y \\ y_t &= Cx_t + \theta_t^y + v_t \end{aligned} \quad (9)$$

where x_t, the state variable at discrete time index t are deviations from $[T_c^*; T_h^*]$. Similarly, the vector $u_t \in \mathbb{R}^2$ represents deviations from $[q_c^*; q_h^*]$. b represents the first column of matrix B, consistent with the fact that disturbances enter the q_c channel. $[\theta^u; \theta^y]$ are input and output disturbance state variables respectively. Both θ^u and θ^y are modeled as integrators but distinguished by the effects of the hidden Markov regime on the second moments of φ^u and φ^y. Consistent with the assumption of an abrupt jump, the covariance of φ^u is assumed to be large with a small probability, and vice versa. θ^y is naturally modeled as an intermittent drift (see (3)). Details are given in the following paragraphs.

Table 1. Nominal steady state operating conditions

Variable	Value	Units
$q_c^* = q_h^*$	10	L/min
T_{ci}^*	25	^{o}C
T_{hi}^*	100	^{o}C
T_c^*	43.75	^{o}C
T_h^*	81.25	^{o}C
α_c^*	5	m^3/min
α_h^*	5	m^3/min
$V_c^* = V_h^*$	75	L

A four-regime Markov chain is considered. These regimes represent the following scenarios:

(1) No disturbance in input channel, No disturbance in output channel ('LO-LO')
(2) No disturbance in input channel, Drifting disturbance in output channel ('LO-HI')
(3) Abrupt disturbance in input channel, No disturbance in output channel ('HI-LO')
(4) Abrupt disturbance in input channel, Drifting disturbance in output channel ('HI-HI')

Accordingly, a simple method for determining the values of the transition probability matrix (Π) is proposed. Per the earlier discussion (Section 2.1), two (sub) transition probability matrices are appropriate for the input (Π^u) and output channels (Π^y) respectively, the first state being the 'normal' regime in both cases.

$$\Pi^u = \begin{pmatrix} 0.99 & 0.01 \\ 0.99 & 0.01 \end{pmatrix};\ \Pi^y = \begin{pmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{pmatrix} \quad (10)$$

An overall transition probability matrix (Π) accounting for the four scenarios can be obtained by assuming statistical independence between the input and output channels. For example in computing π_{23}, one has transitions between the 'normal' to 'abnormal' state for the input channel and the opposite transitions for the output channels so that

$$\pi_{23} = \pi_{21}^u \pi_{12}^y \quad (11)$$

The overall Π [4] is:

$$\begin{pmatrix} 0.98 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 & 0.01 \\ 0.98 & 0.01 & 0.01 & 0.01 \\ 0.01 & 0.98 & 0.01 & 0.01 \end{pmatrix}$$

In accordance to the noise statistics of the possible fault scenarios, the covariance of the overall noise vector $\xi_t \triangleq [\omega_t, \varphi_t^u, \varphi_t^y]$ for the 4 regimes are:

(1) 'LO-LO': $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, 10^{-10}, q_y^{lo}])$
(2) 'LO-HI': $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, 10^{-10}, q_y^{hi}]$
(3) 'HI-LO': $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, q_u^{hi}, q_y^{lo}]$
(4) 'HI-HI': $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, q_u^{hi}, q_y^{hi}]$

Process noise ω is negligible compared to θ^u and will be assumed to be absent for simplicity.

3.3 Alternative MOG Description

If one were to be restricted to an MOG description of the latent regime, then an additional state (θ^β) is required:

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t + b\theta_t^u + \omega_{t+1} \\ \theta_{t+1}^u &= \theta_t^u + \varphi_{t+1}^u \\ \theta_{t+1}^\beta &= \theta_t^\beta + \varphi_{t+1}^\beta \\ \theta_{t+1}^y &= \theta_t^y + \theta_t^\beta \\ y_t &= Cx_t + \theta_t^y + v_t \end{aligned} \quad (12)$$

Similar to (9), θ^u refers to the input channel disturbance and is modeled as an abrupt jump. However, the output

[4] the rows do not sum to unity due to rounding errors

disturbance (θ^y) is now modeled as a double integrator (driven by θ^β). θ^β itself may be interpreted as a velocity term and is driven by φ^β which is set to have a small covariance (10^{-10}) with large probability and a large covariance (of q_y^{hi}) with small probability. This captures the (rare) event of a velocity change when the output disturbance transitions from the white-noise regime to the random-walk mode and vice versa (see Fig. 1(a)). In this case, the sub transition matrices for the input and output channels are:

$$\Pi^u = \Pi^y = \begin{pmatrix} 0.99 & 0.01 \\ 0.99 & 0.01 \end{pmatrix}$$

The overall transition matrix may be obtained as before, per (11). The covariance of the overall noise vector $\xi_t \triangleq [\omega_t, \varphi_t^u, \varphi^\beta, \varphi_t^y]$ for the 4 regimes are:

(1) $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, 10^{-10}, q_y^{lo}, 10^{-10}])$
(2) $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, 10^{-10}, q_y^{hi}, 10^{-10}]$
(3) $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, q_u^{hi}, q_y^{lo}, 10^{-10}]$
(4) $\mathbb{E}[\xi_t \xi_t'] = diag([10^{-10}, 10^{-10}, q_u^{hi}, q_y^{hi}, 10^{-10}]$

3.4 Example 1: Results

Table 2 presents a summary (average over 100 realizations) of the state-estimation error for both the input and output channel. A typical realization is depicted in Fig. 2.

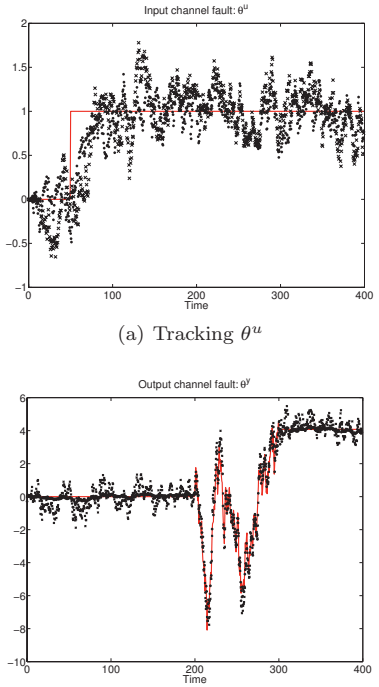
Table 2. 2-norm of state-estimation error (Average of 100 realizations)

Channel	Proposed see (9)	MOG approach see (12)
Input	11.4	12.9
Output	13.3	19.7

Due to the similarities in modeling the abrupt jump in the output channel, it can be seen from Fig. 2(a) and the first line of Table 2 that the performance of the state estimator corresponding to both approaches yield similar performances. However, the MOG approach fares significantly worse than the proposed HMM approach in tracking the fault signal (which is an intermittent drift) corresponding to the output channel (see Fig. 2(b) and the second row of Table 2).

4. EXAMPLE 2: VALVE STICTION

Valve stiction is a common problem in control valves, the latter being widely used in process industries (Choudhury et al. (2005)). Due to the effects of friction, the output (u^x) of the control valve does not track its input (u^c) (*i.e.,* the control signal prescribed by the controller) instantaneously. Instead, u^x has been observed to demonstrate a delayed and sluggish response to u^c, where the valve 'sticks' to its current position if changes in the control signal (and/ or the absolute magnitude itself) are insufficiently large to overcome friction effects. This is usually to the detriment of closed-loop performance. It is assumed that the plant is linear and therefore parameterized by matrices (A, B, C), where A is the state-transition map, B, the input-to-state map and C, the state-to-output map. Technical definitions, first-principles and empirical models of stiction can be found in the articles by Choudhury et al. (2005, 2008) and the references therein. For simplicity,



(a) Tracking θ^u

(b) Tracking θ^y

Fig. 2. Tracking θ^u and θ^y. Comparing the proposed HMM vs. MOG approaches. Legend: solid line - actual fault signal; Dots (·) - HMM; Crosses (x) - MOG

an efficient single-parameter model employed by Stenman et al. (2003) and Srinivasan and Rengaswamy (2005) for stiction detection is used for simulations in the sequel:

$$u_t^x = \begin{cases} u_{t-1}^x, \text{ if } |u_t^c - u_{t-1}^x| \leq d \\ u_t^c, \text{ otherwise} \end{cases} \tag{13}$$

where d represents the valve stiction band. The larger the value of d, the more severe the stiction problem.

The detection, diagnosis and compensation-for valve stiction has received much attention in academia and industry. Based on (13), Stenman et al. (2003) proposed a suitable model for detecting stiction:

$$u_t^x = \tilde{\delta}_t \cdot u_{t-1}^x + (1 - \tilde{\delta}_t) \cdot u_t^c$$

where $\tilde{\delta}_t$ is a binary (0/1) mode parameter occurring with a certain (i.i.d) probability.

For the same purpose of stiction detection and estimating the typically unmeasured u_t^x, we allow $\tilde{\delta}_t$ to have statistics governed by an underlying Markov chain so that observations reflecting persistent 'stickiness' can be more effectively modeled. Also, instead of identifying the segmentation sequence $\{\tilde{\delta}_1, \ldots, \tilde{\delta}_t\}$ that maximizes the posterior quantity $pr(\tilde{\delta}_1, \ldots, \tilde{\delta}_t | y_1, \ldots, y_t)$ through dynamic programming, we propose a novel Markov jump linear description that is consistent with (13) to be used by a GPB2 state-estimator:

$$\begin{pmatrix} x_t \\ u^x_{t-1} \end{pmatrix} = \begin{pmatrix} A & B^x_{r_{t-1}} \\ 0 & \tilde{\delta}_{r_{t-1}} \end{pmatrix} \begin{pmatrix} x_{t-1} \\ u^x_{t-2} \end{pmatrix} + \begin{pmatrix} B^c_{r_{t-1}} \\ 1 - \tilde{\delta}_{r_{t-1}} \end{pmatrix} u^c_{t-1}$$

$$y_t = \begin{pmatrix} C & 0 \end{pmatrix} \begin{pmatrix} x_t \\ u^x_{t-1} \end{pmatrix} + v_t \tag{14}$$

When $r = 1$, stiction is absent, $\tilde{\delta} = 0$, $B^x = 0$, $B^c = B$. Conversely, when $r = 2$, stiction is present, $\tilde{\delta} = 1$, $B^x = B$, $B^c = 0$.

4.1 Simulation Studies: Mixing Tank

For simulation studies, we consider a simple isothermal mixing-tank (of cross-sectional area $\mathcal{A}$) with an outlet stream whose flow-rate is controlled by a valve (with resistance $\mathcal{R}$):

$$\frac{dm}{dt} = \frac{1}{\mathcal{A}}(q_1 + q_2 - \frac{m}{\mathcal{R}}) \tag{15}$$

The controlled (and also measured) variable is the liquid level (m). The flow-rate of the first stream, q_1, is a measured disturbance whereas that of the other stream (q_2) represents the manipulated variable. A PI controller (with gain K_c, and integral time constant τ_I) is given by:

$$u^c_t = u^c_{t-1} + K_c[e_t - e_{t-1} + \frac{h}{\tau_I} e_t],\ e_t \triangleq l - y_t$$

Here l is the set-point, nominally calibrated to a value of 6. For ease, $\mathcal{A}$, $\mathcal{R}$, K_c, τ_I and the measured disturbance signal q_1, are all set to nominal values of 1. A relatively large value for the stiction band is employed: $d = 0.5$. A sampling time of $h = 0.05$ is employed, resulting in the following parametrization to be used by the state estimator: $A = 0.951$, $B = 0.0488$ and $C = 1$. Measurement noise is set to have a known covariance of $R \triangleq \mathbb{E}[v_t v_t'] = 10^{-4}$. To reflect the high degree of stiction, the transition probability matrix Π is:

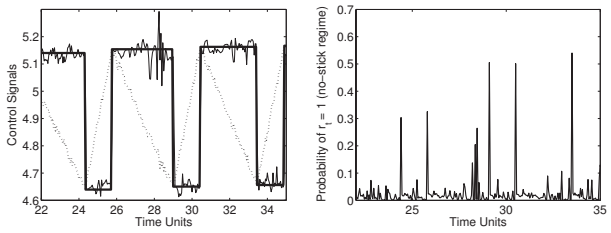
$$\begin{pmatrix} 0.01 & 0.99 \\ 0.01 & 0.99 \end{pmatrix} \tag{16}$$

4.2 Results: Estimating Valve Output & Detecting Stiction

Tracking results for a typical closed-loop realization are shown in Fig. 3. The existence of the cycles in u^c and u^x (Fig. 3(a)) is due to the presence of integral action as well as the valve stiction phenomenon. From Fig. 3(a), it can be seen that the proposed methodology is able to estimate u^x. Observing the (a-posteriori) probability (see (5) and Fig. 3(b)) of the first mode (or equivalently the second) via reveals the time instances where a switch occurs (by means of the probability peaks). Doing so represents an effective way for detecting stiction.

5. CONCLUSIONS & FUTURE WORK

The main contribution of this work is to show that the common faults (abrupt jumps/ biases and drifts) can be better modeled and detected by the proposed HMM-based method. Another novel application is in the context of detecting valve stiction, where it is demonstrated that the output of the valve (which is not normally measured) can be effectively estimated. Future work involves extending the problem to large scale systems (*e.g.* a network of unit operations) of industrial interest.



(a) Time series plot of valve input and output and estimate of valve output. Legend: dotted: u^c; bold line: u^x; line: $\hat{u}^x$

(b) Probability of mode 1

Fig. 3. Tracking unmeasured valve output in mixing-tank example.

REFERENCES

Almeida, G.M. and Park, S.W. (2008). Fault detection and diagnosis in the damadaics benchmark actuator systems - a hidden markov model approach. In *17th IFAC World Congress*, 12419. Seoul, Korea.

Bar-Shalom, Y. and Li, X.R. (1993). *Estimation and Tracking: Principles, Techniques, and Software*. Artech House.

Chiang, L., Russell, E., and Braatz, R. (2001). *Fault detection and diagnosis in industrial systems*. Advanced textbooks in control and signal processing. Springer.

Choudhury, M.S., Jain, M., and Shah, S. (2008). Stiction - definition, modeling, detection and quantification. *Journal of Process Control*, 18, 232–243.

Choudhury, M.S., Thornhill, N.F., and Shah, S. (2005). Modeling valve stiction. *Control Engineering Practice*, 13, 641–658.

Huang, B. (2008). Bayesian methods for control loop monitoring and diagnosis. *Journal of process control*, 18, 829–838.

Kesavan, P. and Lee, J.H. (1997). Diagnostic tools for multivariable model-based control. *Industrial & Engineering Chemistry Research*, 36, 2725–2738.

Rabiner, L.R. (1989). A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.

Smyth, P. (1994). Hidden markov models for fault detection in dynamic systems. *Pattern recognition*, 27, 149–164.

Srinivasan, R. and Rengaswamy, R. (2005). Control loop performance assessment: 1. a qualitative approach for stiction diagnosis. *Industrial & Engineering Chemistry Research*, 44, 6708–6718.

Stenman, A., Gustafsson, G., and Forsman, K. (2003). A segementation-based method for detection of stiction in control valves. *International journal of adaptive control and signal processing*, 17, 625–634.

Willsky, A.S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, 12, 601–611.

Wong, W.C. and Lee, J.H. (2007). Disturbance modeling for process control via hidden markov models. In *8th International Symposium on Dynamics and Control of Process Systems*. Cancun, Mexico.

Root Cause Diagnosis of Plantwide Disturbance Using Harmonic Analysis

M. A. A. Shoukat Choudhury * **Shubharthi Barua** * **Mir Abdul Karim** *
Nahid Sanzida *

* *Department of Chemical Engineering, Bangladesh University of Engineering and Technology (BUET), Dhaka - 1000, Bangladesh (e-mail: shoukat@che.buet.ac.bd)*

Abstract: Disturbances in the form of oscillations are usually originated in process plants due to various faults such as sensor faults, valve faults, process faults and controller tuning faults. Many of these faults can be represented as nonlinearities. Faults in the form of nonlinearities may produce oscillations with a fundamental frequency and its harmonics. This study presents a novel method based on the estimated frequencies, amplitudes and phases of the fundamental oscillation and its harmonics to troubleshoot or isolate the root-cause of plantwide or unit-wide disturbances. Once the root cause is known, the oscillations can be eliminated, and the process can be operated more economically and profitably. The successful application of the method has been demonstrated both on simulated and industrial data sets.

Keywords: Plantwide oscillations, nonlinearity, harmonic, control performance, stiction

1. INTRODUCTION

Modern process plants are designed based on the concept of energy and material integration in order to minimize the energy requirements and pollution levels. Large process plants, such as oil refineries, power plants and pulp mills, are complex integrated systems, containing thousands of measurements, hundreds of controllers and tens of recycle streams. The integration of energy and material flow, required for efficiency, results in the spread of fluctuations throughout a plant. The fluctuations force the plant to be operated further from the economic optimum that would otherwise be possible, and thus cause decreased efficiency, lost production and in some cases increased risk. Because of the scale of operation of process plants, a small percentage decrease in productivity has large financial consequences. It can be extremely difficult to pinpoint the cause of these fluctuations. In the most difficult case, the fluctuations are in the form of oscillations. The problem is that oscillations have no defined beginning and end, and so the cause cannot be isolated by standard techniques. Finding the cause of oscillations is a tedious, labor-intensive, often fruitless task. Once the cause is understood, removal of the oscillations is usually straightforward. Therefore, it is important to detect and diagnose the causes of oscillations in a chemical process.

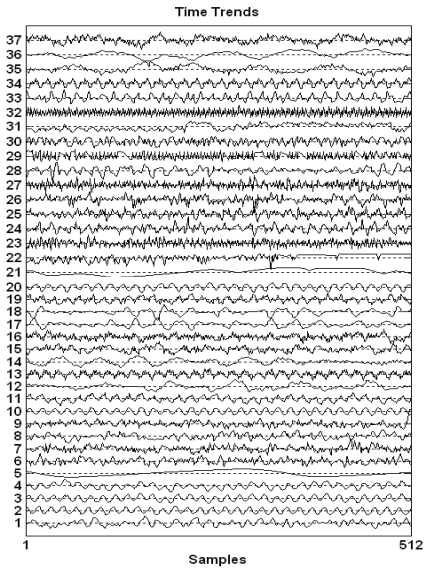
Most of the available techniques for oscillation detection focus on a loop by loop analysis (Hagglund, 1995). Thornhill and co-workers have presented some detection tools that consider the plant-wide nature of oscillations (Thornhill *et al.*, 2003). To detect oscillations in process measurements and identify signals with common oscillatory behavior, use of spectral principal component analysis (Thornhill *et al.*, 2002) or autocorrelation functions (acf) (Thornhill *et al.*, 2003) is suggested. Xia and Howell (2003) have proposed a technique that takes into account the interactions between control loops. Thornhill and Horch (2007) provided an overview of the advances and new direction for solving plantwide oscillation problems. A recent book (Choudhury *et al.*, 2008) provides two chapters on the state of the art technologies for plantwide oscillation detection and diagnosis. This paper demonstrates a method for detecting plantwide oscillations and isolating the root causes of such oscillations.

2. WHAT ARE PLANTWIDE OSCILLATIONS?

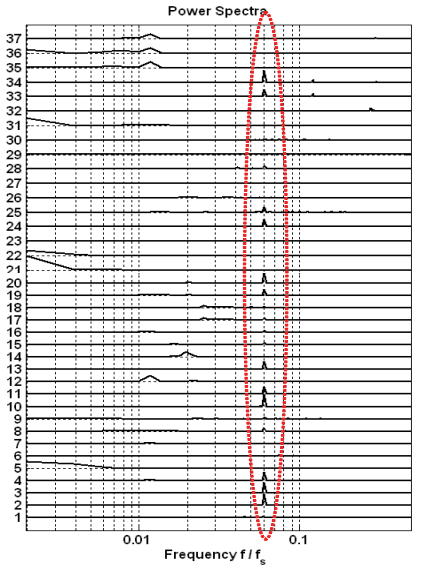
When one or more oscillations is generated somewhere in the plant and propagates throughout a whole plant or some units of the plant, such oscillations are termed as plantwide or unitwide oscillations. Oscillation may propagate to many units of the process plants because of the tight heat and mass integration in the plant as well as the presence of recycle streams in the plant. Figure 1 shows an example of a plant-wide oscillation problem. The top panel shows the time trends of 37 variables representing a plant-wide oscillation problem in a refinery (courtesy of South-East Asia Refinery). The bottom panel shows the power spectra of these variables. A common peak in the power spectra plot indicates the presence of a common disturbance or oscillation at a frequency of 0.06 or approximately 17 samples/cycle in many of these variables. The presence of such plant-wide oscillations takes a huge toll from the overall plant economy.

3. DETECTION OF PLANTWIDE OSCILLATIONS

Detection of plantwide oscillation is relatively an easy problem. Often times the plant operators notice some oscillations in the plant, which leads to a deeper investigation of the problem and may cause the invention of a plantwide oscillation of a larger nature. Over the last few years, some studies were carried out to detect plantwide oscillations (Tangirala *et al.*, 2005; Jiang *et al.*, 2006) and to group the similar oscillations together. The following are the brief description of some of these techniques that can be used for detecting plant-wide oscillations.



(a) Time Trends



(b) Power Spectra

Fig. 1. Time trends and their power spectra for the South-East Asia Refinery data set

3.1 High Density Plot - An Excellent Visualisation Tool

This plot describes time series data and their spectra in a nice compact form in one plot. From this plot, one can easily visualize the nature of the data and the presence of common oscillation(s) in the data. However, this method is not automated and cannot provide a list of the commonly oscillating variables. Figure 1 is an example of a high density plot.

3.2 Power Spectral Correlation Map (PSCMAP)

The power spectral correlation index (PSCI) is defined as the correlation between the power spectra of two different measurements. It is a measure of the similarity of spectral shapes, i.e., measure of the commonness of frequencies of oscillations. The PSCI for any two spectra $|X_i(\omega)|^2$ and $|X_j(\omega)|^2$ is calculated as

$$PSCI = correlation(|X_i(\omega)|^2, |X_j(\omega)|^2) = \frac{\sum_{\omega_k} |X_i(\omega_k)|^2 |X_j(\omega_k)|^2}{\sqrt{|X_i(\omega_k)|^4 |X_j(\omega_k)|^4}} \quad (1)$$

The PSCI always lies between 0 and 1. In the detection of plantwide oscillations, the objective is to collect variables with similar oscillatory behaviour.

For multivariate processes, the PSCI is a matrix of size $m \times m$, where m is the number of measured variables. In order to provide an effective interpretation of the PSCI, the matrix is plotted as a colour map, which is termed as the power spectral correlation map. An important aspect of this colour map is its ability to automatically re-arrange and group variables together with similar shapes, i.e., variables, which oscillate at a common frequency and have therefore similar values of PSCI. For a detailed discussion on this method, refer to (Tangirala *et al.*, 2005).

3.3 Spectral Envelope Method

In (Jiang *et al.*, 2007), the spectral envelope method has been used to troubleshoot plantwide oscillations.

Let $\mathbf{X}$ is a data matrix of dimension $n \times m$, where n is the number of samples and m is the number of variables. If the covariance matrix of $\mathbf{X}$ is $\mathbf{V_X}$ and the power spectral density (PSD) matrix of $\mathbf{X}$ is $\mathbf{P_X}(\omega)$, then the spectral envelope of $\mathbf{X}$ is defined as:

$$\lambda(\omega) \triangleq \sup_{\beta \neq 0} \{ \frac{\beta^* \mathbf{P_X}(\omega) \beta}{\beta^* \mathbf{V_X} \beta} \} \quad (2)$$

where ω represents frequency and is measured in cycles per unit time, for $-1/2 < \omega \leq 1/2$, the $\lambda(\omega)$ is the spectral envelope at the frequency ω, $\beta(\omega)$ is the optimal scaling vector that maximizes the power (or variance) at the frequency ω, the '*' represents conjugate transpose. The quantity $\lambda(\omega)$ represents the largest portion of the power (or variance) that can be obtained at frequency ω from a scaled series. Jiang et. al (2007) provided a detailed description of this method.

4. DIAGNOSIS TECHNIQUES FOR PLANTWIDE OSCILLATIONS

In a control loop, oscillations arises due to the following primary reasons:

(1) Presence of a poorly tuned controller
(2) An oscillatory external disturbance
(3) Presence of a faulty valve, e.g., a sticky valve or saturated valve.
(4) A highly nonlinear process
(5) Model-plant mismatch for an active MPC controller.

As described, the detection of plant-wide oscillation is relatively an easy problem compared to the diagnosis of its root-cause. Recently a number of papers appeared in the literature describing a few techniques to perform root-cause diagnosis of plant-wide oscillation (Thornhill *et al.*, 2001; Thornhill and Horch, 2007; Choudhury *et al.*, 2007; Jiang *et al.*, 2007; Zang and Howell, 2007; Choudhury *et al.*, 2008).

Oscillations originated in process plants due to various faults such as sensor faults and valve faults may be represented as nonlinearities. Faults in the form of nonlinearity produce oscillations with a fundamental frequency and its harmonics. It is well known that the chemical processes are low-pass filters in nature. Therefore, when a fault propagates away from its origin or source, the higher order harmonics get filtered out.

4.1 Oscillation and Harmonics

Sinusoidal fidelity states that if a sinusoidal input passes through a linear system, the output of the linear system is a sinusoid with the same frequency, but with a different magnitude and phase. A linear system does not produce any new frequency. On the other hand, when a sinusoidal signal with a certain frequency passes through various types of nonlinear systems or functions such as a square function, an exponential function, a logarithmic function and a square-root function, nonlinear systems may generate harmonics in addition to the original fundamental frequency of the input sinusoid. Therefore, nonlinearity induced oscillatory signals generally contain a fundamental frequency and its harmonics. Harmonics are oscillations whose frequencies are integer multiples of the fundamental frequency.

4.2 Fourier Series and Harmonics

Fourier series states that any signal can be represented as a summation of sinusoids. Therefore, any time series, $y(t)$, where, $t \in \Re$ can be represented as

$$y(t) = \sum_{i=0}^{\infty} A_i \, cos(\lambda_i \, t + \phi_i) \tag{3}$$

For a signal containing harmonics, Equation 3 can be rewritten as:

$$y(t) = \sum_{i=0}^{M} A_i \, cos(i * \lambda \, t + \phi_i) + \varepsilon(t) \tag{4}$$

where λ is the fundamental frequency. Each term of equation 4 contains three unknowns namely, amplitude, frequency and phase. The basic idea is to estimate the amplitudes, frequencies and phases for each term of equation 4 for any time series and then examine the relationships among the frequencies to find whether they are harmonically related.

From the experience of the author, for useful application of the harmonic analysis of chemical process data, it suffices to use $M = 5$.

4.3 Total Harmonic Content (THC)

A new index called Total Harmonic Content (THC) can be defined as:

$$THC = n * WHM \tag{5}$$

where n is the number of harmonics found and WHM is the Weighted Harmonic Mean. WHM is defined as

$$WHM = \frac{\sum_{i=1}^{M} w_i}{\sum_{i=1}^{M} \frac{w_i}{A_i}} \tag{6}$$

where w_i is weights and is defined as $w_i = i / \sum_{i=1}^{M} i$ so that the summation of the weights are equal to 1 and the weights for the higher harmonics are large. More weights are given to the higher harmonics because due to the low-pass filtering effect of the chemical processes the higher harmonics get filtered out gradually as the signal propagates away from the source or the root cause.

For plant-wide oscillations, the amplitudes, frequencies and phases of first five term of Equation 4 are estimated. For all tags or variables which have the same fundamental frequency are identified and the Total Harmonic Contents (THC) are calculated using Equation 5. After calculating the THCs, the variables are ranked according to the descending order of THC. The variable with the highest THC is likely to be the root cause. Plant information such as Piping and Instrumentation (P&I) diagrams, Process Flow Diagrams (PFD) and operators' knowledge should be utilised in conjunction with the information provided by THC to confirm the root cause. The chance of being right first time is high. However, if the variables with the maximum value of THC is not the root cause, the variable with the second highest value of THC should be investigated as a root cause. Thus maintenance effort should be started from variable with the maximum value of THC to the variables in the descending order of THC.

Thornhill *et al.* (2001) described a similar method using a distortion factor, which was defined as the ratio of the total power of the signal except the power at the fundamental frequency to the power of the fundamental frequency. They used power spectrum to estimate the distortion factor. The method was successful to a limited extent because the power spectrum is heavily affected by the signal noise. On the other hand, the method described here uses only the amplitudes of the harmonics and the fundamental frequency, therefore the THC is not influenced by the signal noise except some small contamination occurs during the frequency and amplitude estimation.

5. SIMULATION EXAMPLE

This simulation example describes a hypothetical process where a nonlinear function, a square function, followed by some linear filters are present. The simulink block diagram is shown in figure 2. The process was excited by a sinusoid

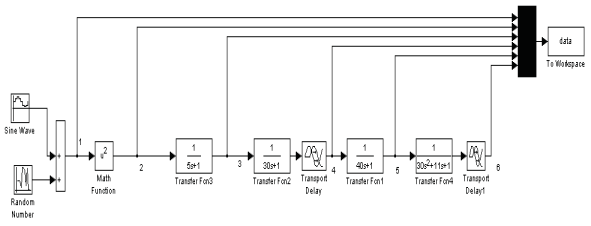


Fig. 2. Simulink block diagram for simple oscillation propagation

with frequency 0.25 rad/sec. Random noise with variance 0.05 was added to the sinusoid. The simulated time series data with their power spectra are shown in Figure 3. From the power spectra, it is hard to see the harmonics generated by the square function because the fundamental frequency has high power. It is interesting to note that for tags 4, 5 and 6, a low frequency oscillation has been developed due to the low pass filtering of the random noise by the process. The fundamental oscillation and its harmonic are gradually filtered out as the signal propagates through the system.

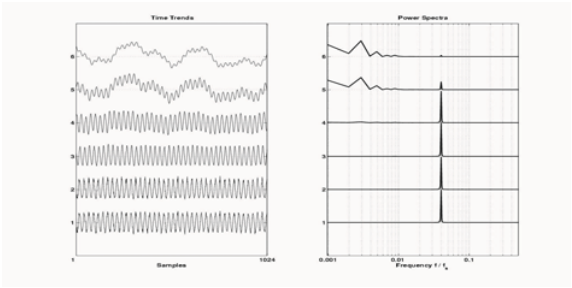


Fig. 3. Simulated data and their power spectra

Table I shows the harmonic analysis of the simulated data. The algorithm correctly identifies the presence of sinusoids in the signal. Five sinusoids are estimated for each signal. For the first signal (tag 1), the magnitude of the first sinusoid is much larger (more than 50 times) than the other sinusoids. The other sinusoids came into play due to the addition of random noise which has power in all frequencies. Research is undergoing to formulate a statistical hypothesis test to detect the presence of true sinusoids. The current algorithm correctly estimates the frequency of the main sinusoid as 0.25 rad/sec. Two dominant sinusoids with frequencies 0.25 and 0.5 rad/sec are estimated for tag 2. For tag 3, the sinusoid with frequency 0.5 rad/sec is present but its power has been decreased because of its attenuation by the first order filter. For tag 4, 5 and 6, the fundamental frequency sinusoid (0.25 rad/sec) has become gradually weak and has been masked with the noise, as evident from the estimated magnitudes shown in the table. The Total Harmonic Content (THC) was calculated for each tag where oscillation with fundamental frequency and its harmonic are found. The maximum THC corresponds to tag 2 indicating the source or root-cause of the propagated oscillation.

6. CASE STUDIES

6.1 Simulation Example - A Non-Linear Dynamic Vinyl Acetate Process

This example describes a simulation case study for root-cause diagnosis of plantwide oscillations using a non-linear dynamic model of a Vinyl Acetate process. The nonlinear dynamic model of the Vinyl Acetate process is published by (Chen *et al.*, 2003) and is freely available from the authors' website. Figure 4 shows a simplified schematic of the Vinyl Acetate Process. The process model contains 246 state variables, 26 manipulated variables and 43 measurements. The process takes approximately 300 minutes time to reach steady state. For details, refer to (Chen *et al.*, 2003).

After the process reached steady state, a 5% stiction ($S = 5$, $J = 2$) in the manipulated variable corresponding to the cooling water flow rate for the separator jacket temperature cooling valve was introduced using the stiction model developed in (Choudhury *et al.*, 2005). Simulation data set consisted of 1000 minutes of data with a sampling time of 15 seconds containing a total of 4000 observations for each variable. The last 1024 data points were used in this analysis in order to avoid transient behaviour due to the sudden introduction of stiction. Figure 5 shows the time trends and power spectra of the manipulated variables of the Vinyl Acetate process. The power spectra show that the variables 1, 2, 4, 5, 6, 7, 8, 9, 11, 12, 14, 19, 21, 22 and 23 are oscillating with a common oscillation at

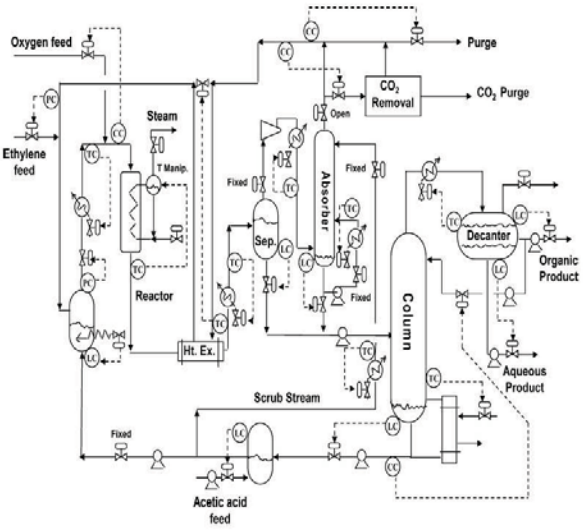


Fig. 4. Schematic of the Vinyl Acetate Process

a normalized frequency of 0.0505. Total Harmonic Content (THC) was calculated for these variables. Figure 6 shows the calculated THC values against the variable or tag number. The maximum THC corresponds to the tag 9 correctly indicating the root-cause of the plantwide oscillation because stiction was introduced in this variable during simulation.

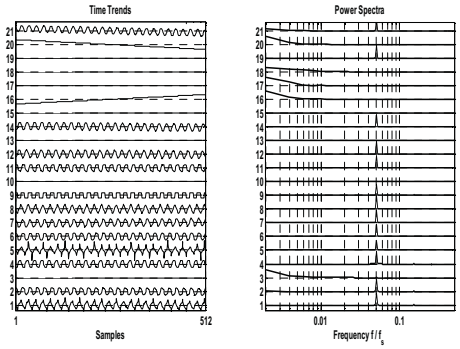


Fig. 5. Time trends and power spectra for the Vinyl Acetate Process Variables

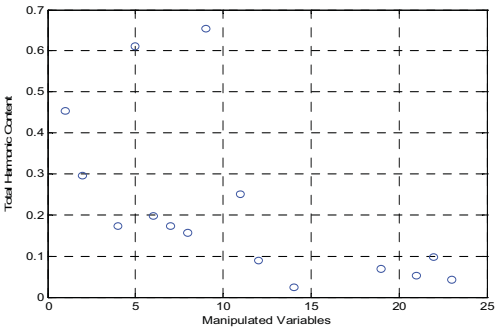


Fig. 6. THC values for the Vinyl Acetate Process Variables

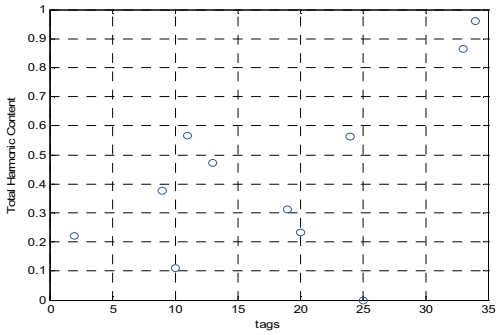


Fig. 7. Total Harmonic Contents (THC) Results for SEA data sets

6.2 An Industrial Example - Application to a Refinery Data Set

The proposed method was applied to a benchmark industrial data set for plantwide oscillations study appeared in the literature such as (Tangirala *et al.*, 2007; Tangirala *et al.*, 2005; Thornhill *et al.*, 2001). The data set, courtesy of a SE Asian Refinery, consists of 512 samples of 37 measurements sampled at 1 min interval. It comprises measurements of temperature, flow, pressure and level loop along with some composition measurements. The time trends of the controller errors are shown in Figure 1(a) and the corresponding power spectra are shown in Figure 1(b). From these figures or using the technique of power spectral correlation map (PSCMAP) described in (Tangirala *et al.*, 2005), it can be found that the tags 2, 3, 4, 8, 9, 10, 11, 13, 15, 16, 17, 19, 20, 24, 25, 28, 33 and 34 are oscillating together with a common frequency of 0.0605 or 17 samples/cycle approximately. All data corresponding to the variables with the common frequency were first normalized so that they had zero-mean and unit variance. Then the amplitudes, frequencies and phases for first five sinusoids were estimated and THC were calculated for these variables. The calculated THC values are plotted against the tag number in Figure 7. The highest THC value corresponds to the tag no. 34, which is the first candidate for the possible root-cause of this plantwide oscillation. In real plant investigation if this tag is not found to be the root cause, then the tag corresponding to next highest value of THC should be investigated. For this case, earlier studies (Thornhill *et al.*, 2001; Tangirala *et al.*, 2005; Tangirala *et al.*, 2007) found tag 34 as the root-cause. Therefore, the proposed THC index correctly detected the root-cause of this plantwide oscillations.

7. CONCLUSIONS AND FUTURE WORKS

This study describes a method to troubleshoot plantwide oscillation using harmonic information present in the signal. The amplitudes, frequencies and phases of the fundamental signal component and its harmonics are estimated and used for the diagnosis of the root-cause of plantwide oscillation. A new index called Total Harmonic Contents (THC) has been defined and used for isolating the root-cause. The method can be automated to facilitate troubleshooting of plantwide oscillation.

8. ACKNOWLEDGEMENTS

All assistance including laboratory, computing and financial supports from Bangladesh University of Engineering and Technology (BUET), Dhaka, Bangladesh are gratefully acknowledged. Authors are also grateful to South-East Asia Refinery for permitting to publish its plant data.

REFERENCES

Chen, Rong, Kedar Dave, Thomas J. McAvoy and Michael Luyben (2003). A nonlinear dynamic model of a vinyl acetate process. *Ind. Eng. Chem. Res.* **42**, 4478–4487.

Choudhury, M. A. A. S., N. F. Thornhill and S. L. Shah (2005). Modelling valve stiction. *Control Engineering Practice* **13**, 641–658.

Choudhury, M. A. A. S., S. L. Shah and N. F. Thornhill (2008). *Diagnosis of Process Nonlinearities and Valve Stiction - Data Driven Approaches*. Springer-Verlag. Germany.

Choudhury, M.A.A.S., V. Kariwala, N.F. Thornhill, H. Douke, S.L. Shah, H. Takadac and J.F. Forbes (2007). Detection and diagnosis of plant-wide oscillations. *Canadian Journal of Chemical Engineering* **85**, 208–219.

Hagglund, T. (1995). A control loop performance monitor. *Control Engineering Practice* **3**(11), 1543–1551.

Jiang, H., M. A. A. S. Choudhury and S. L. Shah (2007). Detection and diagnosis of plantwide oscillations from industrial data using the spectral envelope method. *Journal of Process Control* **17**, 143–155.

Jiang, H., M.A.A.S. Choudhury, S.L. Shah, J. W. Cox and M. A. Paulonis (2006). Detection and diagnosis of plant-wide oscillations via the spectral envelope method. In: *the proceedings of ADCHEM 2006*. Gramado, Brazil. pp. 1139–1144.

Tangirala, A. K., J. Kanodia and S. L. Shah (2007). Non-negative matrix factorization for detection of plant-wide oscillations. *Ind. Eng. Chem. Res.* **46**, 801–817.

Tangirala, A.K., S.L. Shah and N.F. Thornhill (2005). PSCMAP: A new tool for plantwide oscillation detection. *Journal of Process Control* **15**, 931–941.

Thornhill, N. F. and A. Horch (2007). Advances and new directions in plant-wide disturbance detection and diagnosis. *Control Engineering Practice* **15**, 1196–1206.

Thornhill, N. F., B. Huang and H. Zhang (2003). Detection of multiple oscillations in control loops. *Journal of Process Control* **13**, 91–100.

Thornhill, N. F., S. L. Shah and B. Huang (2001). Detection of distributed oscillations and root-cause diagnosis. In: *Preprints of CHEMFAS-4 IFAC*. Korea. pp. 167–172.

Thornhill, N. F., S. L. Shah, B. Huang and A. Vishnubhotla (2002). Spectral principal component analysis of dynamic process data. *Control Engineering Practice* **10**, 833–846.

Xia, C. and J. Howell (2003). Loop status monitoring and fault localisation. *Journal of Process Control* **13**, 679–691.

Zang, X. and J. Howell (2007). Isolating the source of whole-plant oscillations through bi-amplitude ratio analysis. *Control Engineering Practice* **15**, 69–76.

Table 1. Harmonic analysis results for simple oscillation propagation example

Tags	λ_1	λ_2	λ_3	λ_4	λ_5	A_1	A_2	A_3	A_4	A_5	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	RESS	λ_1/λ_1	λ_2/λ_1	λ_3/λ_1	λ_4/λ_1	λ_5/λ_1	THC
1	**0.25**	**1.25**	2.88	1.95	1.22	**1.394**	**0.028**	0.027	0.023	0.023	2.06	0.39	2.17	-2.85	-0.74	24.40	1.0	5.0	11.5	8.0	4.9	3.91
2	**0.25**	**0.50**	2.88	**1.25**	1.99	**1.387**	**0.130**	0.028	**0.028**	0.023	2.06	-2.13	2.27	0.33	-3.01	26.67	1.0	2.0	11.5	5.0	8.0	8.96
3	**0.25**	**0.50**	0.02	0.06	0.03	**1.409**	**0.079**	0.028	0.026	0.024	1.15	2.96	2.46	2.79	0.98	6.08	1.0	2.0	0.1	0.2	0.1	1.91
4	**0.25**	0.02	0.03	0.00	0.05	**1.367**	0.206	0.132	0.110	0.100	-0.5	3 2.17	0.82	-0.95	1.98	21.30	1.0	0.1	0.1	0.0	0.2	0.73
5	0.02	0.25	0.00	0.03	0.02	0.864	0.688	0.545	0.407	0.301	1.66	-1.97	-1.20	0.04	-1.06	83.98	1.0	15.2	0.3	1.8	1.4	
6	0.02	0.00	0.03	0.02	0.25	0.953	0.682	0.454	0.336	0.264	1.30	-1.17	-0.33	-1.31	1.13	88.34	1.0	0.2	1.7	1.4	15.1	

Systematic Development of Automata Generated Languages for Fault Diagnosis in Continuous Chemical Processes

Chuei-Tin Chang. Jung-Yang Chen

National Cheng Kung University, Tainan, Taiwan 70101, ROC (Tel:886-6-275-7575 ext. 62663; e-mail: ctchang@mail.ncku.edu.tw).

Abstract: A SDG-based simulation procedure is presented in this study to qualitatively predict all possible effects of one or more fault propagating in a given process system. All possible state evolution behaviors are characterized with an automaton model. By selecting a set of on-line sensors, the corresponding diagnoser can be constructed and the diagnosability of every fault origin can be determined accordingly. Furthermore, it is also possible to construct a formal diagnostic language on the basis of this diagnoser. Every string (word) in the language is then encoded into an IF-THEN rule and, consequently, a comprehensive fuzzy inference system can be synthesized for on-line diagnosis. The feasibility of this approach is demonstrated with a simple example in this paper.

Keywords: fault diagnosis, automata, signed directed graph, formal language, fuzzy logic.

1. INTRODUCTION

The fault diagnosis methods have been widely recognized as indispensable tools for enhancing process safety. Generally speaking, they could be classified into three distinct groups, i.e., the model based approaches, the knowledge based approaches, and the data-analysis based approaches (Venkatasubramanian et al., 2003a, b). However, in order to carry out these strategies on-line, it is usually necessary to first analyze the historical data and/or operational experiences obtained during *every* serious accident. This requirement cannot always be satisfied in practice.

To circumvent the above drawbacks, a qualitative cause-and-effect model, i.e., the signed directed graph (SDG), is used in the present study to characterize fault propagation mechanisms. The advantage of this modelling approach is mainly due to the fact that the causal relations in process systems can always be established according to generic engineering principles without any quantitative knowledge. On the other hand, it should be noted that such causal models are basically static in nature. Many SDG-based fault identification techniques were therefore implemented on the basis of the *steady-state* symptoms only, e.g., Maurya et al. (2006). Since the effects of fault(s) and/or failure(s) usually propagate throughout the entire system dynamically in sequence, a series of intermediate events may occur before the inception of catastrophic consequences. Thus, the performance of a qualitative diagnosis scheme should be evaluated not only in terms of its correctness but also its timeliness.

To enhance diagnostic efficiency, it is obviously necessary to consider the precedence order (in time) of various fault propagation effects derived from the qualitative models. Extensive studies have already been carried out to develop effective diagnosis strategies by incorporating both the eventual symptoms and also their *occurrence order* into a fuzzy inference system (FIS). This approach has been applied successfully to a number of loop-free processes (Chang et al., 2002) and also to systems with feedback and/or feed forward control loops (Chang and Chang, 2003; Chen and Chang, 2006; 2007).

Despite the fact that diagnostic performance can be significantly improved with the aforementioned technique, the representation, analysis and synthesis of inference systems are still very cumbersome. In particular, many different versions of the symptom occurrence orders can often be deduced from a single fault origin on the basis of SDG model. Manual enumeration of all such scenarios for all origins may become intractable even for a moderately complex system. Furthermore, the diagnosability issues concerning the resulting FIS have never been systematically addressed in the past. Thus, there is a definite need to develop a unified theoretical framework to extract the intrinsic features of dynamic fault propagation mechanisms. Our concern here is primarily with the sequence of system states visited after the occurrence of fault origin(s) and also the associated events causing the state transitions. A systematic procedure is proposed in this paper to construct automata and language models for the purpose of representing these sequences accurately and succinctly. As a result, additional insights can be revealed and, also, more compact inference rules can be produced accordingly. A simple example is provided at the end of this paper to demonstrate the feasibility and effectiveness of the proposed procedures for FIS synthesis and for fault diagnosis.

2. AUTOMATA CONSTRUCTON

2.1 Qualitative Simulation Procedure

Although other qualitative models may be equally acceptable, the SDG is adopted in the present study to simulate (or predict) the effects of faults and failures. This is due to the fact that the needed implementation procedure is conceptually straightforward. Notice first that the fault origins can usually be associated with the primal nodes, i.e., the nodes without inputs. A set of five values, i.e., {-10, -1, 0, +1, +10}, may be assigned to every node in the digraph to represent deviation from the normal value of corresponding variable. The value 0 represents the normal steady state. The negative values are used to denote the lower-than-normal states and the positive values signify the opposite. The magnitudes of non-zero deviations, i.e., 1 or 10, can be interpreted qualitatively as "small" and "large" respectively. The causal relation between two variables can be characterized with a directed arc and the corresponding gain. Each gain may also assume one of the five qualitative values mentioned above. The output value of every arc in digraph can be computed with the gain and its input value according to the following equation:

$$v_{out} = \begin{cases} g \times v_{in} & \text{if } -10 \le g \times v_{in} \le +10 \\ +10 & \text{if } g \times v_{in} > +10 \\ -10 & \text{if } g \times v_{in} < -10 \end{cases} \quad (1)$$

where g, v_{in} and v_{out} denote respectively the gain, input and output values. It is obvious that the deviation values of all variables affected by one or more fault origin can always be computed with this formula, but the time at which each deviation occurs is indeterminable. Without the reference of time in the SDG-based simulation results, it can nonetheless be safely assumed that *the change in an input variable should always occur earlier than those in its outputs*. In essence, this is the most basic assumption adopted in this study. Notice that, if the precedence order of various fault propagation effects is to be considered in fault diagnosis, a large number of different versions of qualitative simulation results may be generated accordingly. All such scenarios can be captured with the automaton model described in the sequel.

2.2 System Automata

A formal definition of a deterministic automaton $\mathcal{A}$ can be found in Cassandras and Lafortune (1999). Specifically, it is a six-tuple

$$\mathcal{A} = (\mathbb{X}, \mathbb{E}, f, \Gamma, x_0, \mathbb{X}_m) \quad (2)$$

where, $\mathbb{X}$ is the set of system states; $\mathbb{E}$ is the finite set of events associated with the transitions in automaton; $f : \mathbb{X} \times \mathbb{E} \to \mathbb{X}$ is the transition function; $\Gamma : \mathbb{X} \to 2^{\mathbb{E}}$ is the active event function; x_0 is the initial system state; $\mathbb{X}_m \subseteq \mathbb{X}$ is the set of marked states. In the present application, each system state $x \in \mathbb{X}$ is either a collection of node values at a particular instance after an initiating failure occurs or the initial state itself. Every event $e \in \mathbb{E}$ represents a previously nonexistent fault effect. Notice that the precedence order of these events must be consistent with the basic assumption mentioned above. The active event function $\Gamma(x)$ is used to specify the events which could change the system state x, while the transition function $f(x, e)$ is used for stipulating the resulting state caused by $e \in \Gamma(x)$. Finally, it should be noted that the initial state x_0 in this study is always associated with the *normal* condition and the set $\mathbb{X}_m$ contains the final steady states reached in all possible fault propagation scenarios.

To facilitate illustration of the automaton construction steps, let us consider the most fundamental digraph configuration, i.e., tree. More specifically, let us use the fictitious SDG model in Figure 1 as an example and also assume that a positive deviation in the upstream variable d, i.e., $d(+1)$, is the only possible fault origin in this case. Notice that, although the precedence order of any two effects along the same branch path in this digraph can be uniquely identified with the proposed qualitative simulation procedure, the order of two distinct events located on *separate* branches should be considered as indeterminable. The corresponding automaton can thus be described with the state transition diagram presented in Figure 2. Every system state here is characterized with a collection of the qualitative values of all variables in the digraph and all of them are listed in Table 1. Three equally possible event sequences between the initial and final system states can be identified from this automaton model, i.e.,

1. $d(+1)x(+1)y(+1)z(-1)u(+1)$,
2. $d(+1)x(+1)y(+1)u(+1)z(-1)$,
3. $d(+1)x(+1)u(+1)\ y(+1)\ z(-1)$.

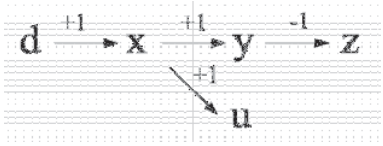


Fig. 1. A tree-shaped SDG model.

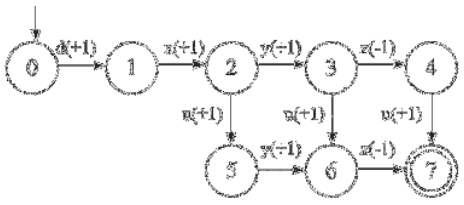


Fig. 2. The state transition diagram of automaton derived from Figure 1.

The automaton resulting from a "large" disturbance can be obtained by following a similar procedure. An auxiliary assumption is introduced in this work to facilitate an accurate description of the fault propagation mechanism, i.e., *the smaller deviation of a process variable must occur before*

reaching a larger one of the same variable. Thus, the automaton in Figure 2 can be revised to incorporate this requirement (see Figure 3).

Table 1. System states in Figure 2.

State	*d*	*x*	*y*	*z*	*u*
0	0	0	0	0	0
1	+1	0	0	0	0
2	+1	+1	0	0	0
3	+1	+1	+1	0	0
4	+1	+1	+1	-1	0
5	+1	+1	0	0	+1
6	+1	+1	+1	0	+1
7	+1	+1	+1	-1	+1

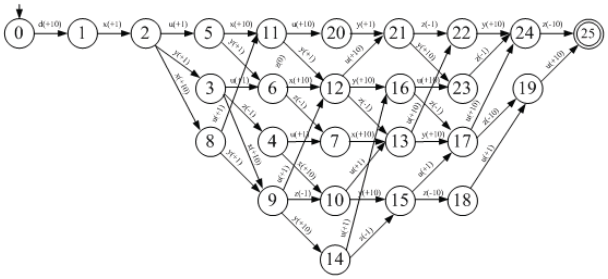


Fig. 3. The automaton resulted from $d(+10)$ in Figure 1

2.3 Diagnoser and Diagnisability

In realistic applications, the fault origins (i.e., failures or upsets) and some of the process variables cannot be monitored on-line. Thus, the event set of an automaton model can be further divided into the observable and unobservable event subsets, i.e., $\mathbb{E} = \mathbb{E}_o \bigcup \mathbb{E}_{uo}$. To check diagnosability of each fault origin and also facilitate diagnostic inference with the available sensors, the system automaton $\mathcal{A}$ should be converted to a *diagnoser* $\mathcal{A}_{diag}$, which is in essence a transformed automaton with the observable subset $\mathbb{E}_o$ as its event set. Although a systematic construction procedure has already been developed by Sampath et al. (1996) for the discrete event systems in general, the diagnosers for the present applications are built with an intuitive but more convenient alternative approach. Specifically, if a state is reached immediately after an unobservable event, then this state is merged with its predecessor(s) in the original automaton model. For example, let us assume that $d(+1)$ is the fault origin and $y(+1)$ is not observable in Figure 2. The corresponding diagnoser can be easily obtained by applying the aforementioned principle (see Figure 4). The numerical node labels here are the same as those in Figure 2, while the subscript of each label is used to reflect whether or not the fault origin has occurred at the corresponding state.

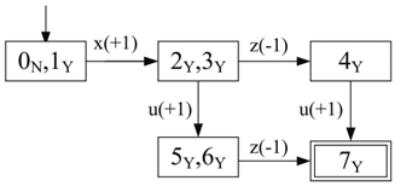


Fig. 4. The diagnoser obtained by assuming $y(+1)$ in Figure 2 is unobservable.

It should be noted that this construction method is applicable even when multiple scenarios are possible. For example, let us consider the SDG model in Figure 5 and assume that there are two measured variables, i.e., y and z , and four potential fault origins, i.e., (1) $d_x(+1)$, (2) $d_y(+1)$, (3) $d_z(+1)$ and (4) $d_u(+1)$.

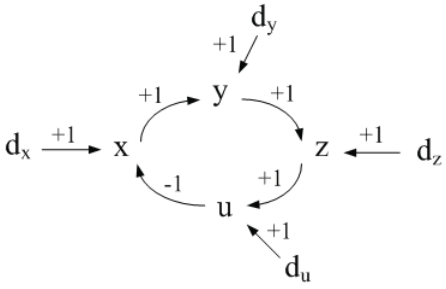


Fig. 5. A SDG model with negative feedback loop

The automaton model of this system and the corresponding diagnoser can be found in Figures 6(A) and 6(B) respectively. Obviously, the issue of diagnosability becomes important in this situation. Although the formal necessary and sufficient conditions of system diagnosability has also been derived and proven rigorously by Sampath et al. (1995), the identifiability of each fault origin in our studies can be determined simply by inspecting the diagnoser. In particular, the diagnosability of a fault origin can be established if it is the unique cause of at least one diagnoser state. Otherwise, the corresponding on-line symptoms should be indistinguishable from those of one or more scenarios caused by other origins. It can be determined on the basis of this criterion that fault origins (3) and (4) are both diagnosable, while the observable event sequences in scenarios (1) and (2) are identical and thus cannot be differentiated from one another.

The feasibility of this simple checking procedure is attributed mainly to the fact that the automata used in the present applications form a special subclass of those for modelling the discrete event systems. More specifically, since the *continuous* chemical processes are considered in this work, the corresponding automata can be characterized with the following unique features:

1. The initial automaton state is always associated with the normal system condition.
2. Every initial state transition is triggered by

failure event(s).

3. Recurrence of system state is not possible, i.e., the automaton is free of any feedback loop. Notice that this feature is due to our assumption that a final steady state is reachable in every possible scenario.

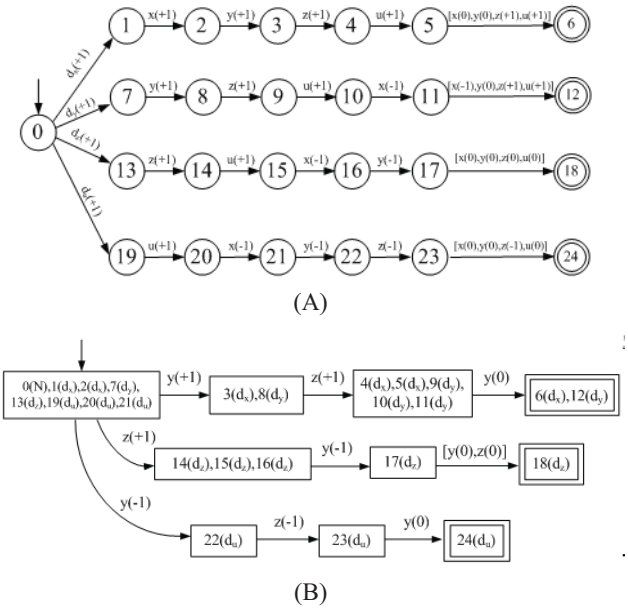


Fig 6. The automaton (A) and diagnoser (B) constructed according to the SDG in Figure 5

3. LANUGAE GENERATION

A *language* $\mathfrak{L}$ is regarded in this work as a collection of finite-length event sequences. These sequences are referred to as *strings* or *words*. The set of all possible events (alphabets) is the set $\mathbb{E}$ defined in equation (2). An additional set $\mathbb{E}^*$ is also utilized here to include all possible strings (including the empty string ε) constructed over $\mathbb{E}$. Thus, it is obvious that $\mathfrak{L} \subseteq \mathbb{E}^*$.

Since fault diagnosis can only be performed according to the on-line symptoms, the automaton $\mathcal{A}_{diag}$ (not $\mathcal{A}$) is used to generate a diagnostic language for the purpose of enumerating all *observable* event sequences caused by a given fault origin. Specifically,

$$\mathfrak{L}\left(\mathcal{A}_{diag}\right)=\left\{s\in \mathbb{E}^* \mid f\left(x_0,s\right) \text{ is defined by } \mathcal{A}_{diag}\right\} \quad (3)$$

The transition function $f\left(x_0,s\right)$ here can be evaluated recursively according to the following rules:

$$\begin{aligned} f\left(x,\varepsilon\right)&=x \\ f\left(x,se\right)&=f\left(f\left(x,s\right),e\right) \end{aligned} \quad (4)$$

where, $s\in \mathbb{E}^*$ and $e\in \mathbb{E}$. In addition, the *marked* language of automaton $\mathcal{A}_{diag}$ can be defined as

$$\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}\right)=\left\{s\in \mathfrak{L}\left(\mathcal{A}_{diag}\right) \mid f\left(x_0,s\right)\in \mathbb{X}_m\right\} \quad (5)$$

Notice that an automaton-based language can be synthesized by first identifying the longest strings and then obtaining all their *prefixes*. Since the marked states in the present application are always terminal, $\mathfrak{L}\left(\mathcal{A}_{diag}\right)$ can be produced by taking the prefix closure of $\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}\right)$ (Cassandras and Lafortune, 1999), i.e.

$$\mathfrak{L}\left(\mathcal{A}_{diag}\right)=\overline{\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}\right)} \quad (6)$$

where, $\overline{\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}\right)}$ denotes the set of all prefixes of the strings in $\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}\right)$. From equation (6), it can be shown that every diagnoser considered in this study must be *nonblocking*, i.e., any string $s\in \mathfrak{L}\left(A_{diag}\right)$ can be always extended by another string t such that $st\in \mathfrak{L}_{\mathrm{m}}\left(A_{diag}\right)$.

Let us use the diagnoser in Figure 4 as an example to illustrate the proposed approach. The two languages marked and generated respectively by $\mathcal{A}_{diag}$ in this case should be

$$\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}\right)=\left\{x(+1)z(-1)u(+1),\ x(+1)u(+1)z(-1)\right\} \quad (7)$$

$$\mathfrak{L}\left(\mathcal{A}_{diag}\right)=\begin{Bmatrix}\varepsilon, x(+1),\ x(+1)z(-1),\ x(+1)u(+1),\\ x(+1)z(-1)u(+1),\ x(+1)u(+1)z(-1)\end{Bmatrix} \quad (8)$$

If the possibilities of multiple fault origins are incorporated in a diagnoser, then it is necessary to further generate a sublanguage specific to every fault origin, i.e.

$$\mathfrak{L}\left(\mathcal{A}_{diag}\right)=\bigcup_i \mathfrak{L}\left(\mathcal{A}_{diag}^{F_i}\right) \quad (9)$$

where, $\mathcal{A}_{diag}^{F_i}$ is an automaton obtained by removing all the abnormal states in A_{diag} which are *not* caused by the i th fault origin F_i. The marked sublanguages of the fault origins in Figure 6(B) can be easily produced with this method, i.e.,

$$\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{d_x(+1)}\right)=\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{d_y(+1)}\right)=\left\{y(+1)z(+1)y(0)\right\} \quad (10)$$

$$\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{d_z(+1)}\right)=\left\{z(+1)y(-1)\left[y(0),z(0)\right]\right\} \quad (11)$$

$$\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{d_u(+1)}\right)=\left\{y(-1)z(-1)y(0)\right\} \quad (12)$$

4. FUZZY INFERENCE SYSTEM

Every string in $\mathfrak{L}\left(\mathcal{A}_{diag}\right)$ is encoded with an IF-THEN rule in this work. These rules can be incorporated in a fuzzy inference system to evaluate the *existence potential* of the corresponding fault origin. In particular, if at least one event sequence in the marked sublanguage $\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{F_i}\right)$ can be confirmed, then it is highly possible that they are caused by the corresponding fault origin F_i. To assert such a belief, the fuzzy conclusion "cs_i is OCR" is adopted in the inference rule, where OCR is the linguistic value of the *occurrence index* cs_i reflecting the highest confidence level in confirming the existence of F_i. More specifically, this rule can be written as

$$\text{IF } s_o \in \mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{F_i}\right) \text{ THEN } cs_i = \text{OCR}$$

where s_o denotes the observed event string.

On the other hand, it is certainly reasonable to disregard the possibility of a fault if none of the corresponding event strings in $\mathfrak{L}\left(\mathcal{A}_{diag}^{F_i}\right)$ can be observed. Thus, the diagnosis for this scenario should be "cs_i is NOC", where NOC is the linguistic value representing the lowest level of confidence. In other words,

$$\text{IF } s_o \notin \mathfrak{L}\left(\mathcal{A}_{diag}^{F_i}\right) \text{ THEN } cs_i = \text{NOC}$$

The diagnostic conclusion for each of the remaining strings should be UCT_ℓ, i.e., uncertain with confidence level ℓ. In particular, this rule can be written as

$$\text{IF } s_o \in \mathfrak{L}\left(\mathcal{A}_{diag}^{F_i}\right) \setminus \mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{F_i}\right) \text{ THEN } cs_i = \text{UCT}_\ell$$

In this study, the confidence level ℓ in confirming the existence of the root cause(s) is assumed to be proportional to the string length. The highest possible confidence level is of course assigned to the strings in $\mathfrak{L}_{\mathrm{m}}\left(\mathcal{A}_{diag}^{F_i}\right)$.

Finally, it should be noted that the aforementioned IF-THEN rules can be implemented with the two-layer fuzzy inference framework developed by Chen and Chang (2006).

5. CASE STUDY

Let us consider the level control system presented in Figure 7 and the corresponding SDG model in Figure 8. All on-line signals, i.e., s_5 - s_8, are assumed to be available for fault diagnosis in this example. For illustration convenience, only two possible scenarios are considered here, i.e., (1) a moderate (controllable) increase in the flow rate of stream 3 while control valve CV-01 sticks and (2) an uncontrollable increase in the flow rate of stream 3.

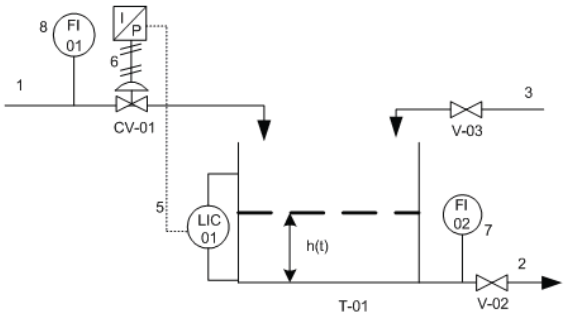


Fig. 7. A level-control system.

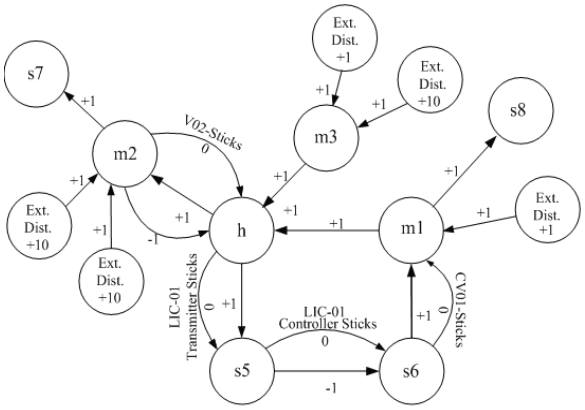
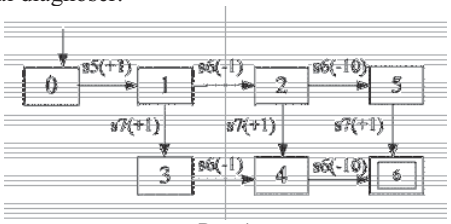
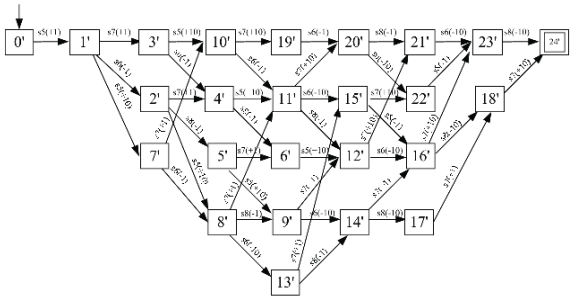


Fig. 8. The SDG model of level-control system.

The diagnoser for these two fault origins can be found in Figure 9. Notice that this automaton is presented in two parts for clarity. States 0 and $0'$ are used to represent the combined states of the normal condition and the system conditions reached immediately after the occurrence of fault origin in scenario 1 and scenario 2 respectively. These two states, i.e., 0 and $0'$, should be lumped into a single one in the actual diagnoser.



Part 1



Part 2

Fig. 9. The diagnoser used for level-control system

To verify the effectiveness of the proposed fault diagnosis approach, extensive numerical simulation studies have been carried out in this work. The on-line measurement data of all fault propagation scenarios were generated with SIMULINK. These data were then used in Sugeno's inference procedure with the fuzzy-logic module of MATLAB toolbox. As an example, let us first examine the occurrence index of the event $m3(+10)$ in scenario 2. It can be observed from Figure 10(A) that the diagnosis is clearly swift and quite accurate. Specifically, the existence of fault origin is detected almost immediately and fully confirmed at about 500 second after its introduction. On the other hand, the occurrence index of the incorrectly assumed fault origin in scenario 2, i.e., $m3(+10)$, is presented in Figure 10(B). Notice that the nonzero occurrence index in the period between 1000 and 2600 sec can be attributed to the fact that the observed event strings caused by the two fault origins can be matched partially during the *initial stage*. More specifically, the set of matched strings is

$$\left\{\overline{s5(+1)s6(-1)s7(+1)}, \overline{s5(+1)s7(+1)s6(-1)}\right\}$$

As the on-line symptoms developed further, none of the longer strings generated by the first part of automaton in Figure 9 can be used to characterize the measurement data obtained after 2600 sec and thus the occurrence possibility of the second fault origin was rejected with the proposed inference mechanisms (Chen and Chang, 2006)..

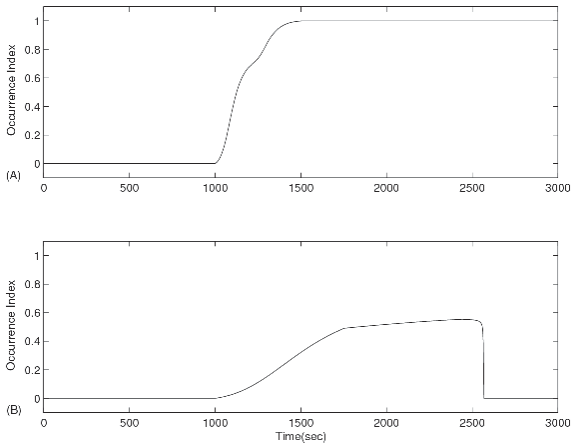


Fig. 10. Diagnosis results of two different scenarios in the level control system. (A) Occurrence index of the second fault origin using simulation data obtained by introducing the same event; (B) Occurrence index of the second fault origin using simulation data obtained by introducing the basic events in the first scenario.

6. CONCLUSIONS

In this study, a SDG-based reasoning procedure is proposed to qualitatively predict all possible symptom patterns and also their progression sequences. These intrinsic features of symptom evolution patterns are captured with automata and language models. The resulting IF-THEN rules can be incorporated in a fuzzy inference system and this system can be installed on-line to identify not only the locations of fault origins but also their magnitude levels with relatively high resolution.

REFERENCES

Cassandras, C.G., and Lafortune S. (1999). *Introduction to Discrete Event Systems*, Kluwer Academic Publisher. Boston.

Chang, S.Y., and Chang, C.T. (2003). A fuzzy-logic based fault diagnosis strategy for process control loops. *Chem. Eng. Sci.,* 58, 3395.

Chang, S.Y., Lin, C.R., and Chang, C.T. (2002). A fuzzy diagnosis approach using dynamic fault trees. *Chem. Eng. Sci.,* 57, 2971.

Chen, J.Y., and Chang, C.T. (2006). Fuzzy diagnosis method for control systems with coupled feed forward and feedback loops. *Chem. Eng. Sci.,* 61, 3105.

Chen, J.Y., and Chang, C.T. (2007). Systematic enumeration of fuzzy diagnosis rules for identifying multiple faults in chemical processes. *Ind. Eng. Chem. Res.,* 46, 3635.

Ju, S.N., Chen, C.L., and Chang, C.T. (2004). Constructing fault trees for advanced process control systems - application to cascade control loops. *IEEE Trans on Reliab*., 53, 43.

Maurya, M.R., Rengaswamy, R., and Venkatasubramanian, V. (2004). Application of signed digraph-based analysis for fault diagnosis chemical process flowsheets. *Engineering Applications of Artificial Intelligence.,* 17, 501.

Maurya, M.R., Rengaswamy, R., and Venkatasubramanian, V. (2006). A signed directed graph-based systematic framework for steady state malfunction diagnosis control loops. *Chem. Eng. Sci.*, 61, 1790.

Oyeleye, O.O., and Kramer, M.A. (1988). Qualitative simulation of chemical process systems: steady-state analysis. *AIChE Journal.*, 34, 1441.

Sampath, M., Sengupta, R., Lafortune, S., Sinnamohideen, K., and Teneketzis, D. (1996). Failure Diagnosis using discrete-event models. *IEEE Trans. Cont. Syst. Tech.,* 4, 105.

Sampath, M., Sengupta, R., Lafortune, S., Sinnamohideen, K., and Teneketzis, D. (1995). Diagnosability of discrete-event systems. *IEEE Trans. Auto. Cont*., 40, 1555.

Venkatasubramanian. V., Rengaswamy, R., Yin, K., and Kavuri, S. (2003a). A review of process fault detection and diagnosis, part 1: quantitative model based methods. *Comput & Chem Engng.*, 27, 293.

Venkatasubramanian, V., Rengaswamy, R., and Kavuri, S. (2003b). A review of process fault detection and diagnosis, part 2: qualitative model and search strategies. *Comput. & Chem. Engng.*, 27, 313.

Sensor Location for Effective Fault Diagnosis in Micro Chemical Processes

Osamu Tonomura*. Satoshi Nagahara. Jun-ichi Kano. Manabu Kano. Shinji Hasebe

Department of Chemical Engineering, Kyoto University, Kyoto, Japan, (*e-mail: tonomura@cheme.kyoto-u.ac.jp).

Abstract: It is clear from worldwide research that micro chemical processes (MCPs) offer a unique approach to the spatial and temporal control of chemical reactions. The well-known advantages of MCPs are often counterbalanced by serious faults such as channel blockage and catalyst deterioration. To realize stable long-term operation of MCPs, it is necessary to develop a monitoring system that can detect and diagnose these faults. In this work, a physical model-based process monitoring system for a tubular microreactor is developed. A state space model is derived by using the orthogonal collocation method, and the extended Kalman filter is used as an observer. The optimal sensor locations are determined so that unknown parameters such as catalyst effectiveness can be estimated most accurately. In addition, the validity of the conventional observability measures in solving the sensor location problems of MCPs is assessed.

Keywords: Microreactor, Sensor location, Process monitoring, Parameter estimation, Fault diagnosis.

1. INTRODUCTION

In microspaces, viscous force, surface tension, conduction heat transfer, and molecular diffusion become dominant. In addition, the contact time and interfacial area between fluids are precisely controlled. These features achievable in microspaces enable us to handle highly exothermic and rapid reactions and to produce fine particles with narrow size distribution (Hessel et al., 2005). However, the above features are often counterbalanced by serious faults such as channel blockage and catalyst deterioration. To realize stable long-term operation of micro chemical processes (MCPs), it is necessary to develop a monitoring and control system suitable for MCPs. Such a system is usually based on the measurements available from installed sensors. However, the existing miniaturized sensors are too expensive in terms of the initial as well as the maintenance costs. In addition, the sensors connected to microreactors in series are not allowed to observe the internal states of microreactors, because they generate dead volume and affect the flow conditions. Therefore, it is important to develop a monitoring system that can estimate unmeasured variables and unknown parameters from a few indirect on-line measurements and quickly detect and diagnose faults. Thus, our technical imperatives are to develop MCPs-oriented sensing devices, to develop a system that can estimate the internal states of MCPs, to propose an approach for effective fault detection and diagnosis in MCPs, etc. So far, there are only few papers about fault detection and diagnosis of MCPs (Kano et al., 2007). In this work, optimal sensor locations for effective fault diagnosis of a tubular microreactor (TMR) are investigated. In addition, the validity of the conventional observability measures in solving the sensor location problems of the TMR is assessed. Finally, operation policies and control structures for MCPs with an external numbering-up structure are investigated. Two types of operation policies, total flow control and pressure drop control, are compared from the viewpoint of flow uniformity when blockage occurs.

2. TUBULAR MICROREACTOR (TMR)

Applications of TMRs can be found in nitration of aromatic compounds, radical polymerization reactions, etc.

2.1 Concept of Fault Detection and Diagnosis

The following method to detect and diagnose faults in TMRs is proposed. A limited number of temperature sensors are embedded in walls of TMRs. Wall temperatures are used to estimate unknown parameters such as catalyst effectiveness. At the same time, the optimal sensor location problems have to be solved so that unknown parameters can be estimated most accurately. Previous similar researches on conventional reactors often neglect heat conduction inside walls when constructing their process models. In case of TMRs, it is crucial to rigorously model the wall heat conduction due to high volume ratio of walls to channels. In addition, there are two methods to formulate process models: empirical model-based method and physical model-based method. In this work, the latter method is adopted.

2.2 Physical Model

Figure 1 shows a schematic diagram of a TMR. Premixed reactants, A and B, are fed into the inner tube, and a coolant is fed into the outer tube. Each flow is assumed to be plug

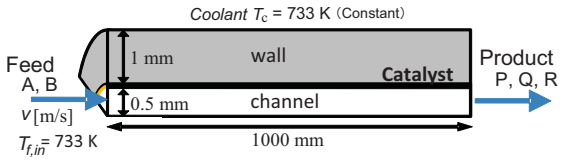


Fig. 1. Schematic diagram of TMR.

flow, and the inner wall surface is coated with a catalyst. On the catalyst surface, the following exothermic reactions take place:

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{P} \quad r_1 = k_1 C_\mathrm{A} \tag{1}$$
$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{Q} \quad r_2 = k_2 C_\mathrm{A} \tag{2}$$
$$\mathrm{P} + \mathrm{B} \rightarrow \mathrm{R} \quad r_3 = k_3 C_\mathrm{P} \tag{3}$$

P is a desired product, and Q and R are by-products. The temperature-dependent rate constant k_i in each reaction is presented by Arrhenius form:

$$k_i = A_i \exp\left(-E_i / RT_S\right) \quad , \quad i = 1,2,3 \tag{4}$$

Table 1 shows the reaction parameters. A and P are treated as key components, and the TMR can be described by the following mass and energy balance equations:

$$\frac{\partial C_i}{\partial t} = -v\frac{\partial C_i}{\partial z} + D\frac{\partial^2 C_i}{\partial z^2} + D\frac{\partial^2 C_i}{\partial r^2} + D\left(\frac{1}{r}\right)\frac{\partial C_i}{\partial r} \tag{5}$$

$$\frac{\partial Tf}{\partial t} = -v\frac{\partial Tf}{\partial z} + \frac{k_f}{\rho_f C_{pf}}\frac{\partial^2 Tf}{\partial z^2} + \frac{k_f}{\rho_f C_{pf}}\frac{\partial^2 Tf}{\partial r^2} + \frac{k_f}{\rho_f C_{pf}}\left(\frac{1}{r}\right)\frac{\partial Tf}{\partial r} \tag{6}$$

$$\frac{\partial C_{\mathrm{AS}}}{\partial t} = -v\frac{\partial C_{\mathrm{AS}}}{\partial z} + D\frac{\partial^2 C_{\mathrm{AS}}}{\partial z^2} - \frac{k_C}{\delta}\left(C_{\mathrm{AS}} - C_\mathrm{A}\big|_{r=\frac{d}{2}}\right) - k_1 C_{\mathrm{AS}} - k_2 C_{\mathrm{AS}} \tag{7}$$

$$\frac{\partial C_{\mathrm{PS}}}{\partial t} = -v\frac{\partial C_{\mathrm{PS}}}{\partial z} + D\frac{\partial^2 C_{\mathrm{PS}}}{\partial z^2} - \frac{k_C}{\delta}\left(C_{\mathrm{PS}} - C_\mathrm{P}\big|_{r=\frac{d}{2}}\right) + k_1 C_{\mathrm{AS}} - k_3 C_{\mathrm{PS}} \tag{8}$$

$$\frac{\partial T_S}{\partial t} = -v\frac{\partial T_S}{\partial z} + \frac{k_f}{\rho_f C_{pf}}\frac{\partial^2 T_S}{\partial z^2} - \frac{U}{\delta}\left(T_S - T_f\big|_{r=\frac{d}{2}}\right) - \frac{U}{\delta}\left(T_S - T_w\big|_{r=\frac{d}{2}}\right) \tag{9}$$
$$-\left(\Delta H_1 k_1 C_{\mathrm{AS}} + \Delta H_2 k_2 C_{\mathrm{AS}} + \Delta H_3 k_3 C_{\mathrm{PS}}\right)/(\rho_f C_{pf})$$

$$\frac{\partial T_W}{\partial t} = \frac{k_w}{\rho_w C_{pw}}\frac{\partial^2 T_W}{\partial z^2} + \frac{k_w}{\rho_w C_{pw}}\frac{\partial^2 T_W}{\partial r^2} + \frac{k_w}{\rho_w C_{pw}}\left(\frac{1}{r}\right)\frac{\partial T_W}{\partial r} \tag{10}$$

where z and r are the axial and radial space coordinates [m], and other variables are summarized in Table 2. Subscripts s, f, and w are catalyst surface, fluid, and wall, respectively. Catalyst thickness, δ, is set to 0.2 mm.

3 PROCESS MODEL FORMULATION

TMR's physical model described in the previous section is regarded as a real process. Fault diagnosis will be based on a state space model, which is derived from the distributed parameter model (5)-(10).

3.1 Process Model

Using the orthogonal collocation method, each state variable is approximated by the following:

Table 1. Reaction parameters.

Reaction	Ai [1/s]	Ei [J/mol]	ΔHi [kJ/mol]
(1)	86760	71711.7	- 2980
(2)	37260	71711.7	- 4622
(3)	149.4	36026.3	- 1664

Table 2. Model parameters.

Parameter	Value	
Reactant velocity v	1	m/s
Mass diffusion coefficient D	1 x 10^{-5}	m^2/s
Heat diffusion coefficient k_f	0.041	J/m K s
Heat conductivity of wall k_w	16.3	J/m K s
Density of reactant ρ_f	1.01	kg/m^3
Density of wall ρ_w	8000	kg/m^3
Viscosity of fluid μ	2.92 x 10^{-5}	Pa s
Heat capacity of reactant C_{pf}	1090	J/kg K
Heat capacity of wall C_{pw}	500	J/kg K
Reactor length L	1	m
Channel diameter d	1	mm
Wall thickness d_w	1	mm
Inlet conc. of species A C_A	4	mol/m^3
Inlet conc. of species P C_P	0	mol/m^3
Inlet temp. of reactant $T_{f,in}$	733	K
Coolant temp. T_c	733	K

$$X(t,z,r) \approx \sum_{i=1}^{n}\sum_{j=1}^{m} L_i(z) L_j(r) X_{i,j}(t), \quad X = C_A, C_P, T_f, T_w \tag{11}$$

$$X(t,z,r) \approx \sum_{i=1}^{n} L_i(z) X_i(t), \quad X = C_{AS}, C_{PS}, T_S \tag{12}$$

where $X_{i,j}(t)$ and $X_i(t)$ are the value of $X(t, z, r)$ at the axial collocation points $z_1 \sim z_\mathrm{n}$ ($0 = z_1 < z_2 < \ldots < z_\mathrm{n} = L$) and the radial collocation points $r_1 \sim r_\mathrm{m}$ ($0 = r_1 < r_2 < \ldots < r_\mathrm{m} = d/2$), respectively. $L_i(z)$ and $L_j(r)$ are Lagrange polynomials. In this study, n and m are set to 30 and 5, respectively, and the collocation points are chosen as roots of a Chebyshev polynomial. The above approximation is also applied to the states at boundaries. Equations (5)-(10) are transformed into the following:

$$\dot{\boldsymbol{x}}(t) = f\left(\boldsymbol{x}(t), \boldsymbol{u}_{in}\right) \tag{13}$$

where $\boldsymbol{u}_\mathrm{in}$ denotes the input vector and $\boldsymbol{x}$ the state vector:

$$\boldsymbol{x} = [\ \boldsymbol{C}_\mathrm{A}{}^T \quad \boldsymbol{C}_\mathrm{P}{}^T \quad \boldsymbol{C}_{\mathrm{AS}}{}^T \quad \boldsymbol{C}_{\mathrm{PS}}{}^T \quad \boldsymbol{T}_f{}^T \quad \boldsymbol{T}_S{}^T \quad \boldsymbol{T}_w{}^T\ ]^T . \tag{14}$$

3.2 Observer Design

Nonlinear estimation problems in this research are solved with the extended Kalman filter (EKF), which is based upon the principle of linearization of the state transition matrix and the observation matrix with Taylor series expansions. Wall temperature measurements are used as observed variables. To obtain the best estimates, the locations of the available

sensors must be selected carefully. The sensor location candidates in the axial direction of TMR are prepared according to the collocation points, and the optimal sensor locations are selected according to the following performance index,

$$J = 1 \Big/ \left\{ \frac{1}{N} \sum_{k=1}^{N} \left\| \boldsymbol{\theta}_{real} - \boldsymbol{\theta}_{est}(k) \right\|^2 \right\} \tag{15}$$

where N is the number of iteration steps, $\boldsymbol{\theta}_{\text{real}}$ and $\boldsymbol{\theta}_{\text{est}}$ the real and estimated values of parameters. J means the rate of convergence towards real values. As J becomes larger, the estimation performance becomes higher.

4. SIMULATION RESULTS

Three different examples of fault diagnosis of TMR are presented in this section. In addition, the conventional observability measures are examined for the selection of optimal sensor locations in TMR.

4.1 CASE 1: Catalyst Deterioration

In CASE 1, an optimal sensor location problem for estimation of catalyst effectiveness α is investigated. It is assumed that reaction rate constant k_1 includes α:

$$k_1(t,z) = \alpha A_1 \exp\left(-E_1 / RT_S(t,z)\right) \tag{16}$$

An initial value of $\alpha = 1$ is considered, followed by an abrupt change from 1 to 0.8 at time $t = t_s$. After catalyst deterioration, α is estimated from one temperature measurement by using EKF. The normal steady state ($\alpha = 1$) is used as the initial state of parameter estimation. J is calculated at every candidate for sensor locations. As shown in Fig. 2 (left), the largest value of J can be found near the inlet of TMR. This result is well illustrated by Fig. 2 (right). The solid and dotted curves in Fig. 2 (right) correspond to the wall temperature profiles along TMR having $\alpha = 1$ and 0.8, respectively. The great differences between the solid and dotted curves mean the high responses of temperature to a parameter change. This physical interpretation confirms that the optimal sensor location is near the inlet of TMR.

4.2 CASE 2: Channel Blockage

In CASE 2, a blockage diagnosis problem in TMR is investigated. Specifically, the inlet flow rate is constant, and one temperature sensor is used to estimate fluid velocity v. An initial value of $v = 1$ m/s is considered, followed by an abrupt change from 1 m/s to 1.2 m/s at time $t = t_s$. After channel blockage, v is estimated by using EKF. The normal steady state ($v = 1$ m/s) is used as the initial state of parameter estimation. J is plotted as a function of sensor position. The relative large values of J can be found in the latter part of TMR, as seen in Fig. 3 (left). This result is well illustrated by Fig. 3 (right). As well as CASE 1, the large differences between both profiles mean the high responses of temperature to a fluid velocity change. All things considered,

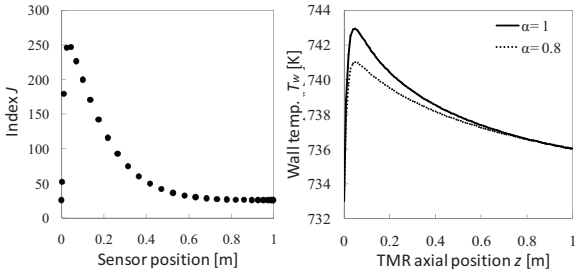


Fig. 2. Estimation results of CASE 1.

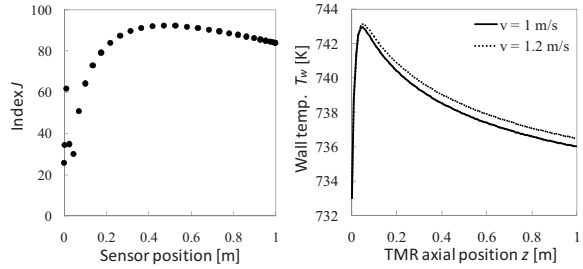


Fig. 3. Estimation results of CASE 2.

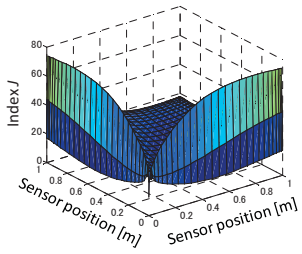


Fig. 4. Results of CASE 3.

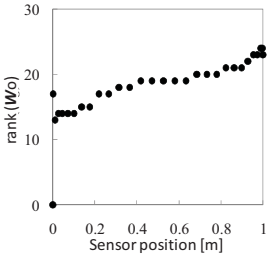


Fig. 5. Rank of $\boldsymbol{W}_O$.

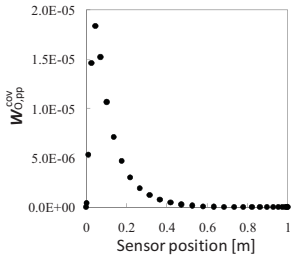


Fig. 6. $\boldsymbol{W}_{\text{O, pp}}^{\text{cov}}$ *of CASE 1.*

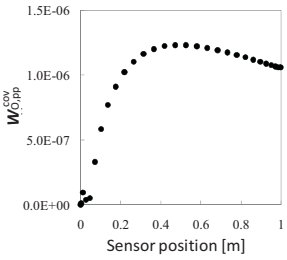


Fig. 7. $\boldsymbol{W}_{\text{O, pp}}^{\text{cov}}$ *of CASE 2.*

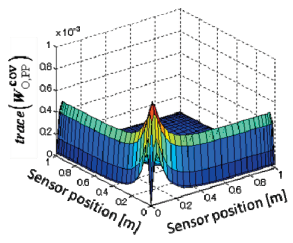


Fig. 8. $\boldsymbol{W}_{\text{O, pp}}^{\text{cov}}$ *of CASE 3.*

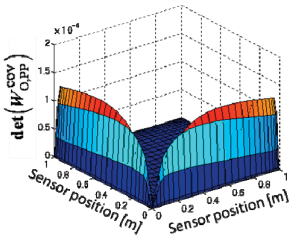


Fig. 9. $\boldsymbol{W}_{\text{O, pp}}^{\text{cov}}$ *of CASE 3.*

the optimal sensor location is in the latter part of TMR, which differs from CASE 1.

4.3 CASE 3: Simultaneous Diagnosis of Multi-Faults

This study is similar to the previous cases, but the two unknown parameters, α and v, are simultaneously estimated by using two temperature sensors. The simulation conditions of the observer are the same as used in the previous cases. Figure 4 suggests that one sensor should be located near the inlet of TMR and the other in the latter part of TMR to realize successful fault diagnosis.

4.4 Observability Measures

Over the years, several studies on sensor locations for estimation in conventional processes have been reported. The representative approaches are to define optimal criteria based on the observability Gramian $\boldsymbol{W}_{\mathrm{O}}$. An overview of several criteria is summarized below. Muller et al. (1972) examine the smallest singular value, the determinant, and the trace of $\boldsymbol{W}_{\mathrm{O}}$ as a measure for sensor location. Dochain et al. (1997) present the condition number of $\boldsymbol{W}_{\mathrm{O}}$, and van den Berg et al. (2000) use the trace of $\boldsymbol{W}_{\mathrm{O}}$ as a criterion for sensor location. While the above presented $\boldsymbol{W}_{\mathrm{O}}$ is suitable only for linear systems, one alternative is to use the observability covariance matrix $\boldsymbol{W}_{\mathrm{O}}^{\mathrm{cov}}$ if systems are nonlinear. Singh et al. (2005) present the trace of $\boldsymbol{W}_{\mathrm{O}}^{\mathrm{cov}}$ for sensor location. The aim of this section is to assess the effectiveness of the existent sensor location criteria for parameter estimation problems in TMR.

In the above presented TMR, observability analysis is performed by determining the rank of $\boldsymbol{W}_{\mathrm{O}}$. Figure 5 shows the rank of $\boldsymbol{W}_{\mathrm{O}}$ at every possible sensor location. Since $\boldsymbol{W}_{\mathrm{O}}$ at every location is rank deficient, then the system is not observable. In such a situation, the smallest singular value of $\boldsymbol{W}_{\mathrm{O}}$ is zero. Accordingly, it is not suitable to use criteria such as the smallest singular value, the smallest eigenvalue, the determinant, and the condition number of $\boldsymbol{W}_{\mathrm{O}}$. Therefore, the trace of $\boldsymbol{W}_{\mathrm{O}}^{\mathrm{cov}}$ is appropriate to the determination of the optimal sensor location for parameter estimation. However, for this work, $\boldsymbol{W}_{\mathrm{O}}^{\mathrm{cov}}$ is used instead of $\boldsymbol{W}_{\mathrm{O}}$ because the process is a time-variant system in case of parameter estimation problem. $\boldsymbol{W}_{\mathrm{O}}^{\mathrm{cov}}$ can be decomposed into (Singh et al., 2005):

$$\boldsymbol{W}_{\mathrm{O}}^{\mathrm{cov}} = \begin{bmatrix} W_{\mathrm{O,nn}}^{\mathrm{cov}} & W_{\mathrm{O,pn}}^{\mathrm{cov}} \\ W_{\mathrm{O,np}}^{\mathrm{cov}} & W_{\mathrm{O,pp}}^{\mathrm{cov}} \end{bmatrix} \tag{17}$$

where $\boldsymbol{W}_{\mathrm{O,nn}}^{\mathrm{cov}}$ the observability covariance matrix of the system, $\boldsymbol{W}_{\mathrm{O,pp}}^{\mathrm{cov}}$ the covariance of the outputs caused by changes in the parameters, and $\boldsymbol{W}_{\mathrm{O,np}}^{\mathrm{cov}}$ and $\boldsymbol{W}_{\mathrm{O,pn}}^{\mathrm{cov}}$ the covariance of the outputs resulting from changes in the states and parameters. The optimal sensor location for parameter estimation is computed by maximizing the traces of $\boldsymbol{W}_{\mathrm{O,pp}}^{\mathrm{cov}}$. The traces of $\boldsymbol{W}_{\mathrm{O,pp}}^{\mathrm{cov}}$ in CASEs 1 and 2 are plotted for possible sensor locations in Figs. 6 and 7. Figs. 6 and 7 are similar to Figs. 2 (left) and 3 (left), respectively. That is, the trace of $\boldsymbol{W}_{\mathrm{O,pp}}^{\mathrm{cov}}$ is useful as a criterion for judging where the sensors should be located. On the other hand, the trace and determinant of $\boldsymbol{W}_{\mathrm{O,pp}}^{\mathrm{cov}}$ in CASE 3 are plotted in Figs. 8 and 9, respectively. As compared with Fig. 4, it is clarified that the determinant of $\boldsymbol{W}_{\mathrm{O,pp}}^{\mathrm{cov}}$ is effective as a criterion of optimal sensor locations for estimating multi-parameters.

5. OPERATION POLICY FOR MCPs

The production capacity of MCPs is usually increased by numbering-up, which means the repetition of a microdevice. One of the critical operational issues of MCPs with numbering-up structure is to keep a uniform flow distribution among parallelized microdevices even when blockage occurs in one or more microdevice. Since it is not practical to install flow controllers in all the microdevices, a simple and effective operation policy against blockage occurrence needs to be developed. In this work, two types of operation policies, total flow control and pressure drop control, are compared from the viewpoint of flow uniformity when blockage occurs.

5.1 Total Flow Control and Pressure Drop Control

To maintain the desired product quality, it is important to keep a uniform flow rate in each microdevice of the micro chemical plant when blockage occurs, because flow maldistribution worsens the performance of the micro chemical plant. In this research, pressure drop control is proposed to achieve the uniform flow distribution.

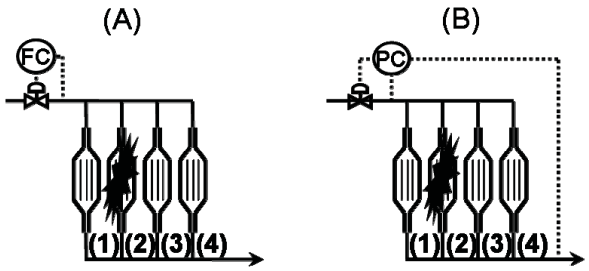


Fig. 10. Parallelized microdevices under two operation policies: (A) total flow control and (B) pressure drop control.

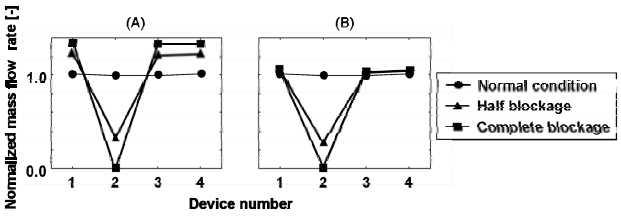


Fig. 11. Mass flow rate of each microdevice under blockage occurrence: (A) total flow control and (B) pressure drop control.

A micro chemical plant consisting of four parallelized microdevices in Fig. 10 is used to demonstrate the difference of two operation policies, total flow control and pressure drop control. When blockage occurs in microdevice 2, the flow distribution under the pressure drop control is derived by simulation and compared with that under the total flow control. In Fig. 10, reactant is fed to the parallelized microdevices at 0.1 m/s in the normal condition. The physical properties of the reactant are assumed to be the same as water (293 K). The results are shown in Fig. 11. In the case of blockage occurrence, mass flow rates of unblocked microdevices are significantly increased under the total flow control (Fig. 11 left), while they are kept constant at the value in a normal condition under pressure drop control (Fig. 11 right). These results show that the proposed pressure drop control is effective in making flow distribution uniform even when blockage occurs.

5.2 Comparison of Control Structures in Pressure Drop Control

In the previous section, it was confirmed that pressure drop control is superior to total flow control in realizing uniform flow distribution among unblocked microdevices when blockage occurs. In this section, two different control structures based on pressure drop control, pumping pressure control and pressure drop control over the parallelized section, are investigated.

5.2.1 Experimental Apparatus

Micro chemical plants having four or eight parallelized microdevices are used to grasp the distinction between two control structures. A schematic drawing of a micro chemical plant having four parallelized microdevices is shown in Fig. 12. Reactant is fed with a double plunger pump. The product line is open to the atmosphere. The flow rate of each microdevice is measured by using an in-line mass flow meter, and blockage in each microdevice is artificially realized by closing the valve located between the pump and each microdevice. Each microdevice consists of an SUS tube having 0.3 mm inner diameter and 2 m in length. In addition, an SUS tube with 0.5 mm in inner diameter and various lengths is installed after the parallelized microdevices to represent units that are not necessary to be parallelized. Hereafter, this section is referred to as a residence time section. The pressure drops over the parallelized section and the residence time section are denoted by ΔPa and ΔPb, respectively. The ratio of ΔPa to ΔPb is changed by adjusting the length of the residence time section.

5.2.2 Pumping Pressure Control

Under pumping pressure control, the double plunger pump is operated at constant pumping pressure. In experiments, pumping pressure is kept at a gauge pressure of 500 kPa – 1 MPa. Pressure drop over the whole micro chemical plant is kept constant under pumping pressure control, because the product line is open to the atmosphere.

The influence of blockage on flow distribution under pumping pressure control is investigated through both simulations and experiments with changing the ratio $\Delta Pa/\Delta Pb$ in the range of one-fifth to seven. The first step in the experimental procedure is to adjust the pumping pressure to realize a total flow rate of 12 mL/min. This operating condition is regarded as the normal condition. Then, microdevice 1 is artificially blocked by closing the valve. In 300 s, the micro chemical plant is returned to the normal condition by opening the valve. These procedures are repeated for the other valves to imitate blockage in the other microdevices.

Figure 13 shows the normalized average mass flow rate, which is defined as the ratio of average mass flow rate of unblocked microdevices under blockage occurrence to that under the normal condition at each $\Delta Pa/\Delta Pb$. There is little difference between the results of experiments and those of CFD simulations. The normalized average mass flow rate becomes closer to the flow rate under the normal condition as $\Delta Pa/\Delta Pb$ becomes larger. In other words, ΔPa should be significantly larger than ΔPb to keep the flowrate of unblocked microdevices unchanged when blockage occurs. It is concluded that pumping pressure control is effective to realize uniform flow distribution when the pressure drop over the parallelized section is dominant.

5.2.3 Pressure Drop Control Over the Parallelized Section

The flow uniformity achieved by pumping pressure control depends on $\Delta Pa/\Delta Pb$, which is the ratio of the pressure drop over the parallelized section to that over the residence time section. The flow uniformity in the parallelized microdevices deteriorates when $\Delta Pa/\Delta Pb$ is small. In this subsection,

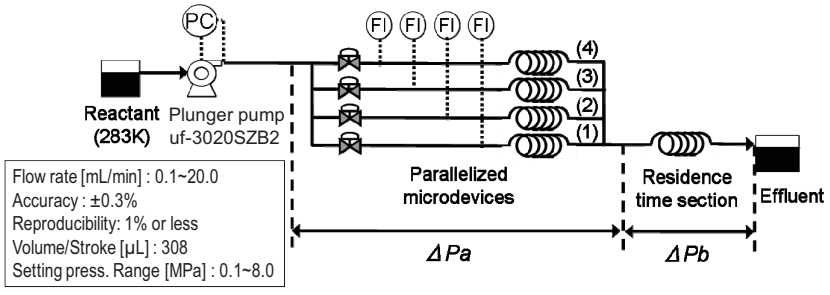


Fig. 12. Micro chemical plant under pumping pressure control.

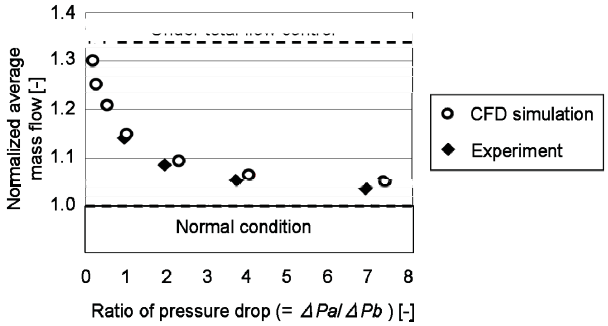


Fig. 13. Influence of blockage on the mass flow rate.

another pressure drop control structure where ΔPa is directly controlled by manipulating the flow rate of the bypass line is proposed.

The performance of the proposed control structure is evaluated experimentally by using the micro chemical plants with four (Type A) and eight (Type B) parallelized microdevices. The plunger pump is operated so that the total flow rate is kept constant. The other conditions are the same as those in the previous subsection.

The experimental results for a Type A plant are shown in Fig. 14. The top and bottom figures correspond to the case where $\Delta Pa/\Delta Pb = 50$ and $\Delta Pa/\Delta Pb = 1$, respectively. In the range of 0–300 s, the micro chemical plant is operated under the normal condition. The difference in flow rate between microdevices 1 and 2 is due to the degree of precision in the fabrication. When blockage occurs in microdevice 1 at 300 s, the flow rate of microdevice 1 goes to zero instantaneously and the flow rate of microdevice 2 deviates from its set-point. However, the flow rate of microdevice 2 returns to the normal level in a few tens of seconds. Flow rate deviation becomes small when blockage occurs gradually. The transient responses of microdevices 3 and 4 are similar to that of microdevice 2; they are not shown in Fig. 14 to identify the transient response of each microdevice easier. The top and bottom figures in Fig. 14 show almost the same profiles. This result shows that the efficiency of the proposed control structure does not depend on $\Delta Pa/\Delta Pb$.

The experimental result of a Type B plant is almost the same as that of the Type A plant. These results show that the proposed control structure has the function of keeping the flow rate of the unblocked devices constant regardless of the changes in $\Delta Pa/\Delta Pb$ and the number of parallelized microdevices.

6. CONCLUSIONS

In this study, sensor locations for effective fault diagnosis of TMR are investigated. It is clarified that two different faults are accurately diagnosed by using only two wall temperature sensors, which are optimally located in the axial direction of TMR. In addition, the optimality criteria for sensor locations in TMRs are investigated. The results obtained from case studies demonstrate that the criteria based on observability covariance matrix are effective and their maximization allows

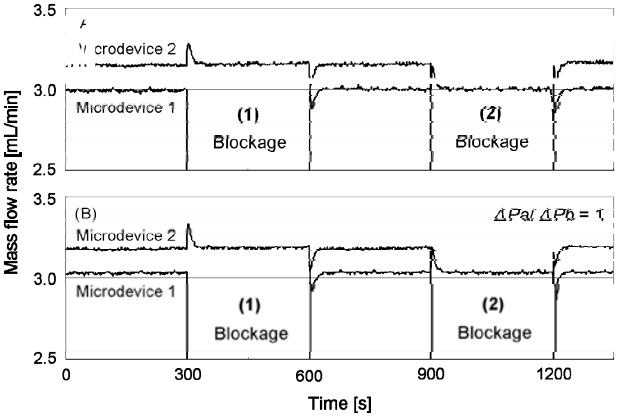


Fig. 14. The time series of mass flow rate in each microdevice in Type A: (A) $\Delta Pa/\Delta Pb = 50$, (B) $\Delta Pa/\Delta Pb = 1$.

one to determine where the sensors should be located. Finally, we have discussed operation policies and control structures for micro chemical plants with an external numbering-up structure. Two operation policies, total flow control and pressure drop control, were compared. The simulation result shows that the pressure drop control is effective to keep a uniform flow distribution among the parallelized microdevices even when blockage occurs. In addition, two control structures based on pressure drop control, pumping pressure control and pressure drop control over the parallelized section, were investigated experimentally. The former control structure is simple. However, this structure functions only when the ratio of the pressure drop over the parallelized section to that over the residence time section, $\Delta Pa/\Delta Pb$, is large. On the other hand, the latter control structure can make the flow distribution uniform for any $\Delta Pa/\Delta Pb$.

REFERENCES

Dochain, D., Tali-Maamar, N., and Babary, J. P. (1997). On Modeling, Monitoring and Control of Fixed Bed Bioreactors, *Comp. Chem. Eng.*, 21 (11), 1255-1266.

Hessel, V., Hardt, S., and Lowe, H. (2005). Chemical Micro Process Engineering: Fun-damentals, Modelling and Reactions. *Wiley, VCH, Weinheim*.

Kano, M., Fujioka, T., Tonomura, O., Hasebe, S., and Noda, M. (2007). Data-based and Model-based Blockage Diagnosis for Stacked Microchemical Processes. *Chem. Eng. Sci.*, 62 (4), 1073-1080.

Muller, P. C. and Weber, H. I. (1972). Analysis and Optimization of Certain Qualities of Controllability and Observability for Linear Dynamical Systems, *Automatica*, 8, 237-246.

Singh, A. K. and Hahn, J. (2005). Determining Optimal Sensor Locations for State and Parameter Estimation for Stable Nonlinear Systems, *Ind. Eng. Chem. Res.*, 44, 5645-5659.

van den Berg, F. W. J., Hoefsloot, H. C. J., Boelens, H. F. M., and Smilde, A.K. (2000). Selection of Optimal Sensor Position in a Tubular Reactor Using Robust Degree of Observability Criteria, *Chem. Eng. Sci.*, 55, 827-837.

Data-driven Control Loop Diagnosis: Dealing with Temporal Dependency in Bayesian Methods ⋆

Fei Qi *, **Biao Huang** *,

* *Department of Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, T6G 2G6, Canada(e-mail: fei.qi@ualberta.ca, biao.huang@ualberta.ca)*

Abstract: Conventional Bayesian methods commonly assume that the evidences are temporally independent. This condition does not hold for most practical engineering problems. With evidence transition information being considered, the temporal domain information can be synthesized within the Bayesian framework to improve the diagnosis performance. A data-driven algorithm is developed to estimate the evidence transition probabilities. The application in a pilot scale process is presented to demonstrate the data dependency handling ability of the proposed approach.

Keywords: Performance monitoring, Performance assessment, Bayesian diagnosis, Evidence dependency

1. INTRODUCTION

Control loop performance assessment and diagnosis has been an active area of research in the process control community. A number of control performance methods are available, including the ones based on minimum variance control (MVC), linear quadratic Gaussian control (LQG), historical data trajectories, and user-specified control, etc (Huang and Shah, 1999; Harris et al., 1999; Qin, 1998; Jelali, 2006; Schafer and Cinar, 2004; Patwardhan and Shah, 2002). Several surveys on the control performance assessment research are available (Harris et al., 1999; Qin, 1998; Hoo et al., 2003; Hugo, 2006; Jelali, 2006). Besides performance evaluation of control loops, significant progress has also been made in the development of monitoring algorithms for process and instrument components within the control loops, such as sensor monitor, valve stiction monitor, process model validation monitor (Qin and Li, 2001; Ahmed et al., 2009; Choudhury et al., 2008; Mehranbod et al., 2005). A number of successful industry applications of the process monitors have been reported. However, many practical problems remain. One of the outstanding problems is that the monitoring algorithms are often designed for one specific problem. An implicit assumption that other unattended components are in good shape is made. Clearly this assumption does not always hold, and thus it may lead to misleading results. It is desirable to develop approaches that not only monitor the performances of single components, but also are capable of synthesizing the information from different monitor outputs to isolate underlying source of problematic control performance.

According to Huang (2008), several challenging issues exist for the process monitor synthesizing problem. The first one is the similar symptoms among different problem sources. For instance, oscillations can either be invoked by a sticky valve or an improperly tuned controller. Another problem is that no process monitor has 100% detection rate and 0% false alarm rate, and thus a probabilistic framework should be built to represent the uncertainties. Third, a large number of the developed monitoring algorithms are purely data based without any *a priori* process information. Incorporating *a prior* process knowledge into the diagnosis framework is challenging, but better diagnosis performance can be expected by doing so.

The Bayesian method sheds lights on the problem solutions by providing a probabilistic information synthesizing framework. Applications of the Bayesian methods have been reported in medical science, image processing, target recognition, pattern matching, information retrieval, reliability analysis, and engineering diagnosis (Dey and Stori, 2005; Mehranbod et al., 2005; Steinder and Sethi, 2004; Chien et al., 2002). It is one of the most widely applied techniques in probabilistic inferencing. Built upon previous work in Bayesian fault diagnosis by Pernestal (2007) and a framework laid out by Huang (2008), Qi and Huang (2008) developed a data-driven Bayesian algorithm for control loop diagnosis with consideration of missing data. The algorithm is tested through simulation, where the information synthesizing ability of the proposed approach is demonstrated. However, the existing Bayesian methods have not considered temporal dependency problem. In this paper, a new algorithm is developed with consideration of temporal dependency, so as to achieve more reliable and better diagnosis performance.

⋆ This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Alberta Ingenuity Fund.

The remainder of this paper is organized as follows. In Section 2, the control loop diagnosis problem and related preliminaries are described, and the data-driven Bayesian approach developed in (Qi and Huang, 2008) is briefly revisited. The rationale to consider evidence temporal dependency is detailed in Section 3. The estimation algorithm for the evidence transition probability is developed in Section 4. Section 5 presents application of the proposed approach to a pilot scale process. Finally the Section 6 concludes this paper.

2. DATA-DRIVEN BAYESIAN DIAGNOSIS METHOD REVISIT

2.1 Control Loop Diagnosis Problem

Generally a control loop consists of the following components: controller, actuator, process, and sensor. These components may all suffer from malfunctions. In this work, monitors are assumed to be available for some or all of the components of interest. These monitors, however, are all subject to disturbances and thus can produce false alarms, and each monitor can be sensitive to abnormalities of other problem sources. Our target is to pinpoint the source of problematic control performance based on the collected monitor output data.

To adopt the Bayesian method for control loop diagnosis, several notations need to be introduced (Qi and Huang, 2008).

Mode M Assume that a control loop under diagnosis consists of P components of interest: C_1, C_2, $\cdots$, C_P, among which the problem source may lie in. Each component is said to have a set of discrete operating status. For instance, the sensor might be "biased" or "unbiased". An assignment of operating status to all the components of interest in the control loop is called a mode, and denoted as M; M can take different values and a specific value is denoted by m. For example, m=(C_1=*well tuned controller*, C_2=*valve with stiction*, $\cdots$). Suppose that component C_i has q_i different status. Then the total number of possible modes is

$$Q = \prod_{i=1}^{P} q_i,$$

and the set of all possible modes can be denoted as

$$\mathcal{M} = \{m_1, m_2, \cdots, m_Q\}.$$

Evidence E The monitor readings, called evidence, are the input to the diagnostic system, and are denoted as $E = (\pi_1, \pi_2, \cdots, \pi_L)$, where π_i is the output of the i-th monitor, and L is the total number of the monitors. Often the continuous monitor readings are discretized according to predefined thresholds. In this work, monitor readings all take discrete values. For example, the control performance monitor may indicate "optimal", "normal", or "poor". A specific value of evidence E is denoted as e; for example, e=(π_1=*optimal control performance*, π_2=*no sensor bias*, $\cdots$). Suppose that the single monitor output π_i has k_i different discrete values. Then there are totally

$$K = \prod_{i=1}^{L} k_i$$

different evidences, and the set of all possible evidence values can be denoted as

$$\mathcal{E} = \{e_1, e_2, \cdots, e_K\}.$$

Historical evidence data set $\mathcal{D}$ In this paper, process data refer to the readings from physical instruments such as temperature, pressure, etc. The evidence data refer to the readings of monitors which are calculated typically from a section (window) of process data. Historical evidence data are retrieved from the past record where the mode of the control loop, namely, status of the components of interest in the control loop, is available, and the monitor readings are also recorded. Each sample d^t at time t in the historical evidence data set $\mathcal{D}$ consists of the evidence E^t and the underlying mode M^t. This can be denoted as $d^t = (E^t, M^t)$, and the set of historical evidence data is denoted as

$$\mathcal{D} = \{d^1, d^2, \cdots, d^{\tilde{N}}\},$$

where $\tilde{N}$ is the number of historical evidence data samples. In (Qi and Huang, 2008), all the historical evidence data samples are assumed to be independent as commonly assumed in the data-driven Bayesian approaches.

2.2 Data-driven Bayesian Diagnosis Approach

This section will give a brief review of the data-driven Bayesian approach proposed by Qi and Huang (2008). Given current evidence E, historical evidence data set $\mathcal{D}$, the posterior probability of each possible operating mode can be calculated according to Bayes' rule:

$$p(M|E, \mathcal{D}) \propto p(E|M, \mathcal{D})p(M), \tag{1}$$

where $p(E|M, \mathcal{D})$ is the likelihood probability; $p(M)$ is the prior probability of mode M. Among all the possible modes, generally the one with the largest posterior probability is taken as the underlying mode based on the maximum *a posterior* (MAP) principle, and the abnormality associated with this mode is normally diagnosed as the problem source.

Since prior probabilities are determined by *a priori* information, the main task of building a Bayesian diagnostic system is the estimation of the likelihood probabilities with historical evidence data $\mathcal{D}$. In (Qi and Huang, 2008), a data-driven Bayesian algorithm for estimation of the likelihood probability is proposed based on the work by Pernestal (2007) and Huang (2008).

Suppose that the likelihood of evidence $E = e_i$ under mode $M = m_j$ is to be calculated, where

$$e_i \in \mathcal{E} = \{e_1, \cdots, e_L\},$$

and

$$m_j \in \mathcal{M} = \{m_1, \cdots, m_Q\}.$$

The following result can be obtained for calculating the likelihood (Pernestal, 2007):

$$p(E = e_i|M = m_j, \mathcal{D}) = \frac{n_{i|m_j} + a_{i|m_j}}{N_{m_j} + A_{m_j}}, \tag{2}$$

where $n_{i|m_j}$ is the number of historical evidence samples with the evidence $E = e_i$, and mode $M = m_j$; $a_{i|m_j}$ is the number of prior samples that is assigned to

evidence e_i under mode m_j; $N_{m_j} = \sum_i n_{i|m_j}$, and $A_{m_j} = \sum_i a_{i|m_j}$.

3. DEPENDENCY IN HISTORICAL EVIDENCE DATA

Note that in the approach described in Section 2, an assumption is that the current evidence only depends on current mode, and is independent on the previous samples. This assumption is true for appropriate designed monitors, as explained below.

The independency among evidences relies on how the evidence data are sampled, and how the disturbance affects the monitor outputs. If the evidence samples are collected with sufficiently large intervals, or if the disturbance has no or weak correlation among the evidence samples, the evidences may be considered as independent. Generally the first requirement regarding the sampling interval can be easily satisfied by leaving sufficient gap between consecutive monitor readings. However, there is no guarantee that the disturbance is uncorrelated in practical applications. If disturbance has long-term autocorrelation and the gap between consecutive monitor readings is not large enough, then the temporal independency assumption of monitor readings can not apply. A simple practical example of long-term autocorrelation of the disturbance is the ambient temperature change. Consider that each monitor reading is calculated based on 1-hour data and there is no overlap in the use of data. Assume that some of the monitor outputs are affected by the ambient temperature. Due to the cyclic change of temperature within 24 hours, the evidence samples should follow a predictable pattern. Apparently it is more justifiable to consider the dependency between those evidence samples than ignoring it in this example.

Besides the practical issues, another limitation with the conventional Bayesian approach ignoring evidence dependency is its inability to capture all time domain information. An illustrative problem is presented in the following. Suppose that the system under diagnosis has two modes m_1 and m_2. One monitor π, with two discrete outcomes, 0 and 1, is available. A set of 100 samples of the monitor outputs is shown in Figure 1. The title in each plot indicates the underlying operating mode under which the data are collected.

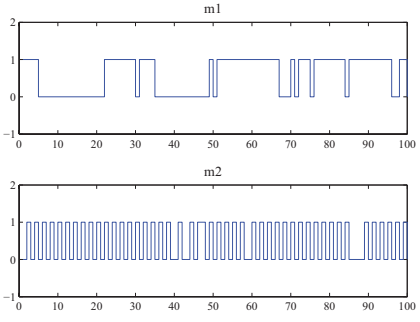


Fig. 1. Monitor outputs of the illustrative problem

The likelihood probability of evidence being 0 or 1 is calculated according to Equation 2, as summarized in Table 1. Clearly the likelihood of the evidence being 0 or 1 under the two modes is almost identical. This may invoke confusion in the diagnosis, which will lead to higher false diagnosis rate. By looking at the data plot in Figure 1, one can argue that distinguishing the two modes should not be such a difficult task. Although the evidences under m_1 and m_2 share similar likelihood, the frequencies of the evidence change apparently differ far from each other. The limitation with the conventional Bayesian method without considering evidence dependency is that the temporal information has not been completely used, leading to less efficient diagnosis performance. In summary it is desirable to take the evidence dependency into consideration when building the diagnostic model.

Table 1. Likelihood estimation of the illustrative problem

	$e = 0$	$e = 1$
m_1	0.46	0.54
m_2	0.48	0.52

With the consideration of evidence dependency, the mode posterior probability is calculated as

$$p(M^t|E^{t-1}, E^t, \mathcal{D}) \propto p(E^t|M, E^{t-1}, \mathcal{D})p(M). \quad (3)$$

Comparing the difference between Equation 1 and Equation 3, the main task of building a Bayesian diagnostic system boils down to the estimation of the evidence transition likelihood probability with historical evidence data $\mathcal{D}$, $p(E^t|M, E^{t-1}, \mathcal{D})$.

4. EVIDENCE TRANSITION PROBABILITY ESTIMATION

The intention of the estimation of evidence transition probability is to make the estimated likelihood probabilities be consistent with historical evidence data set $\mathcal{D}$ in which the evidence dependency exists. Our goal is to calculate the likelihood probability of an evidence E^t given current underlying mode M^t and previous evidence E^{t-1} to reflect the dependency with the Markov property, so every composite evidence sample for evidence transition probability estimation purpose should include three elements,

$$d_E^{t-1} = \{M^t, E^{t-1}, E^t\}. \quad (4)$$

Accordingly, the new composite evidence data set $\mathcal{D}_E$, which is assembled from historical evidence data set $\mathcal{D}$ to estimate transition probability, is defined as

$$\begin{aligned}\mathcal{D}_E &= \{d_E^1, \cdots, d_E^{t-1}\} \\ &= \{(M^2, E^1, E^2), \cdots, (M^t, E^{t-1}, E^t)\}, \end{aligned} \quad (5)$$

Figure 2 depicts how the original collected historical evidence data are divided to form composite evidence samples. In Figure 2, the part highlighted with shadows or gray and enclosed by the dash-lined or solid-lined frame is a single composite evidence sample described by Equation 4.

Suppose that the evidence transition probability from $E^{t-1} = e_s$ to $E^t = e_t$ under mode $M^t = m_k$ is to be estimated from the composite evidence data set,

$$p(E^t|E^{t-1}, M^t, \mathcal{D}_E) = p(e_t|e_s, m_k, \mathcal{D}_E) \quad (6)$$

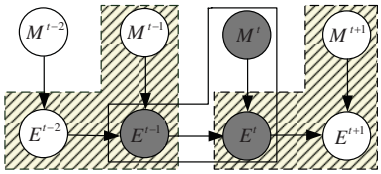


Fig. 2. Bayesian model considering dependent evidence

where

$$e_s, e_t \in \mathcal{E} = \{e_1, \cdots, e_L\}, \tag{7}$$

and

$$m_k \in \mathcal{M} = \{m_1, \cdots, m_Q\}. \tag{8}$$

The transition probability $p(e_t|e_s, m_k, \mathcal{D}_E)$ can only be estimated from the composite evidence data subset $\mathcal{D}_{E|m_k}$ where the mode $M^t = m_k$,

$$\begin{aligned} p(e_t|e_s, m_k, \mathcal{D}_E) &= p(e_t|e_s, m_k, \mathcal{D}_{E|m_k}, \mathcal{D}_{E|\neg m_k}) \\ &= p(e_t|e_s, m_k, \mathcal{D}_{E|m_k}), \end{aligned} \tag{9}$$

where $D_{E|\neg m_k}$ is the composite evidence data set whose underlying mode M^t is not m_k. To simplify notations, the subscript m_k will be omitted when it is clear from the context.

Define $\Phi_s = \{\phi_{s,1}, \phi_{s,2}, \cdots, \phi_{s,K}\}$ as the likelihood parameters for all possible evidence transition from evidence e_s under mode m_k, where $\phi_{s,j} = p(e_j|e_s, m_k)$ is the transition probability from evidence e_s to e_j, and K is the total number of possible evidences. The likelihood probability can be computed by marginalization over all possible evidence transition likelihood parameters,

$$\begin{aligned} &p(e_t|e_s, m_k, \mathcal{D}_E) \\ &= \int_{\Psi_1, \cdots, \Psi_K} p(e_t|\Phi_1, \cdots, \Phi_K, e_s, m_k, \mathcal{D}_E) \\ &\qquad \cdot f(\Phi_1, \cdots, \Phi_K|e_s, m_k, \mathcal{D}_E) d\Phi_1 \cdots \Phi_K \\ &= \int_{\Psi_1, \cdots, \Psi_K} \phi_{s,t} \cdot f(\Phi_1, \cdots, \Phi_K|e_s, m_k, \mathcal{D}_E) d\Phi_1 \cdots \Phi_K, \end{aligned} \tag{10}$$

where Ψ_i is the space of all the likelihood parameters Φ_i.

$f(\Phi_1, \cdots, \Phi_K|e_s, m_k, \mathcal{D}_E)$ can be calculated according to Bayes' rule,

$$\begin{aligned} &f(\Phi_1, \cdots, \Phi_K|e_s, m_k, \mathcal{D}_E) \\ &\propto p(\mathcal{D}_E|e_s, m_k, \Phi_1, \cdots, \Phi_K) f(\Phi_1, \cdots, \Phi_K|e_s, m_k). \end{aligned} \tag{11}$$

In Equation 11, the first term, $p(\mathcal{D}_E|e_i, m_k, \Phi_1, \cdots, \Phi_K)$ is the composite evidence data likelihood given parameter sets $\{\Phi_1, \cdots, \Phi_K\}$. It should be noted that likelihood of composite evidence data $\mathcal{D}_E$ is solely determined by the mode and parameter sets $\{\Phi_1, \cdots, \Phi_K\}$, and thereby is independent of e_s given the mode and the likelihood parameters, i.e.,

$$\begin{aligned} p(\mathcal{D}_E|e_s, m_k, \Phi_1, \cdots, \Phi_K) &= p(\mathcal{D}_E|m_k, \Phi_1, \cdots, \Phi_K) \\ &= \prod_{i=1}^{K} \prod_{j=1}^{K} \phi_{i,j}^{\tilde{n}_{i,j}}, \end{aligned} \tag{12}$$

where $\tilde{n}_{i,j}$ is the number of evidence transition from e_i to e_j in the composite evidence data set.

Assume that the priors for different parameter sets Φ_i and Φ_j, for $i \neq j$, are independent (Pernestal, 2007),

$$f(\Phi_1, \cdots, \Phi_K|e_s, m_k) = f(\Phi_1|e_s, m_k) \cdots f(\Phi_K|e_s, m_k). \tag{13}$$

Dirichlet distribution is commonly used to model priors of the likelihood parameters with parameters $b_{i1}, \cdots, b_{iK}$,

$$f(\Phi_i|e_s, m_k) = \frac{\Gamma(\sum_{j=1}^{K} b_{ij})}{\prod_{j=1}^{K} \Gamma(b_{ij})} \prod_{j=1}^{K} \phi_{ij}^{b_{ij}-1}, \tag{14}$$

where b_{ij} can be interpreted as the number of prior samples for evidence transition from e_i to e_j. $\Gamma(\cdot)$ is the gamma function,

$$\Gamma(x) = (x-1)!, \tag{15}$$

where x is positive integer.

Substituting Equation 14 and Equation 12 in Equation 11, and then combining it with Equation 10, the following result is obtained,

$$p(e_t|e_s, m_k, \mathcal{D}_E) = \frac{\tilde{n}_{s,t} + b_{s,t}}{\tilde{N}_s + B_s}, \tag{16}$$

where $\tilde{N}_i = \sum_j n_{i,j}$ is the total number of historical data samples with evidence transition from e_i under mode m_k, and $B_i = \sum_j b_{i,j}$ is the corresponding total number of prior samples.

By comparing Equation 2 and Equation 16, we can see that the evidence transition probabilities are also determined by both prior samples and historical samples, similar to the evidence likelihood calculation when the evidences are independent. The difference lies in how the numbers of prior and historical evidence samples are counted. In Equation 2 the prior and historical evidence samples refer to a simple count of the evidence samples corresponding to a certain mode, while in Equation 16 the prior and historical evidence samples refer to the count of composite evidence samples corresponding to a evidence transition under the target mode. Readers are referred to (Qi and Huang, 2008) for detailed explanation of the likelihood calculation.

5. PILOT SCALE EXPERIMENT

5.1 Process Description

The experiment setup is a water tank with one inlet flow and two outlet flows. The schematic diagram of the process is shown in Figure 3. The inlet flow is driven by a

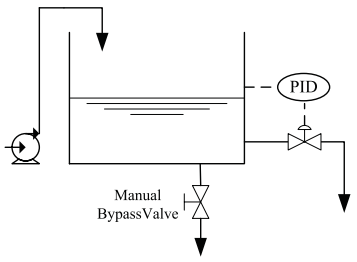


Fig. 3. Pilot scale tank process

pump. Of the two outlet flow valves, one is adjusted by a

PID controller to provide level control for the tank, and the other one is a manual bypass valve. It is assumed that the bypass valve is closed when the system in its normal operation condition.

Three operating modes are defined, including the *normal functioning* (NF) mode, and two problematic modes *leakage* and *bias*. The problems associated with the two faulty modes are: the tank leakage problem defined as *leakage* mode, implemented by opening the bypass valve manually, and the sensor bias problem defined as *bias* mode, implemented by adding a constant bias to the sensor output. The two problems share similar symptoms in terms of shifting the steady state operation point of the process. For instance, when there is a leakage in the tank, the valve adjusted by the PID controller will decrease to maintain the water level; when there is a negative sensor bias, the valve will also decrease. Thus it is not obvious how to distinguish the two faulty modes without any advanced information synthesizing approach. To make things worse, the external disturbance introduced by changing the pump input will also shift the operation point. Thus the operation point may also change during normal operation.

Random binary sequence is introduced into the inlet pump input to simulate temporal dependent disturbances. By defining the high value as 1, and the low value as 0, the disturbance transitions are designed to follow the transition probability matrices presented in Equation 17.

$$P_{NF}^{dis} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{pmatrix} \end{array}, P_{leakage}^{dis} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 0.1 & 0.9 \\ 0.8 & 0.2 \end{pmatrix} \end{array},$$

$$P_{bias}^{dis} = \begin{array}{c} \\ 0 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \end{array}. \quad (17)$$

Two process monitors, process model validation monitor and sensor bias monitor, are designed. Since we mainly focus on the study of the information retrieving and synthesizing abilities of Bayesian approaches with different diagnosis strategies, the selected monitor algorithms are not necessary to have good performances.

The output of process model validation monitor π_1 is given by the squared sum of the nominal model output residuals, scalded by the magnitude of the process output. Let the simulated output of the nominal model be $\hat{y}_t$ at each sampling instance t, and the real output be y_t. The output of the model validation monitor π_1 is calculated as

$$\pi_1 = \frac{\sum_{t=1}^{N}(y_t - \hat{y}_t)^2}{\bar{y}}, \quad (18)$$

where $\bar{y} = \frac{1}{N}\sum_{t=1}^{N} y_i$ is the mean value of the process output over one monitor reading period, and N is the length of data segment over the one monitor reading.

The sensor bias monitor output π_2 is obtained by examining the operation point shift. For illustration, consider the scenario when a negative sensor bias occurs. The steady state in terms of the sensor output will not change, since it is controlled by the PID. The steady state output of the controller, i.e., the valve position, however, will decrease. The valve position will reverse in the presence of the positive sensor bias. Thus we can detect the sensor bias by monitoring the deviation of the controller output mean value from the nominal operation point. The output of the sensor bias monitor π_2 is calculated as

$$\pi_2 = \left| u_0 - \frac{1}{N}\sum_{t=1}^{N} u_t \right|, \quad (19)$$

where u_0 is the nominal operation point of the controller output, u_t is the controller output at each sampling instance t, and N is the length of process data segment for a single monitor reading. Note that this monitor will fail for the transition data, thus only steady state data are collected and used in this example.

5.2 Diagnosis Settings and Results

Process data are collected for the three predefined modes. The sampling interval is set to be one second. Every 100 seconds of process data are used for calculation of one monitor reading. Totally 600 monitor readings are calculated from 16.5 hours of process data samples. The collected evidence data of the three modes are divided into two parts for estimation of the likelihood, and for cross-validation respectively. Table 2 summarizes the Bayesian diagnosis parameters.

Table 2. Summary of Bayesian diagnosis parameters

Discretizaion	$k_i = 2$, $K = 2^2 = 4$
Historical data	120 monitor readings for each mode
Prior samples	Uniformly distributed with prior sample, for single evidence space, and evidence transition space
Prior probabilities	$p(NF) = p(m_{other}) = 1/3$
Evaluation data	80 independently generated cross-validation monitor readings for each mode

With the data-driven Bayesian approaches of two different strategies, namely, considering and ignoring the evidence dependency, the diagnosis results in Figure 4 are obtained based on the cross-validation data. In the plot, the gray bars are the numbers of the underlying modes occurred in the validation data set; the light gray and dark bars are the numbers of the diagnosed mode by two diagnostic approaches respectively.

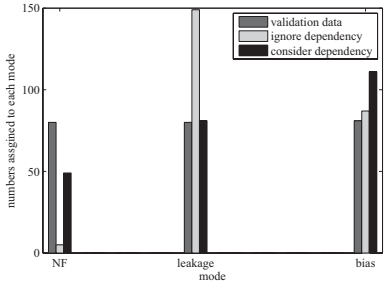


Fig. 4. Numbers assigned to each mode

Owing to the dependent external disturbance, the Bayesian approach ignoring evidence dependency significantly overestimates the number of leakage mode occurrence, and underestimates the number of NF mode. Therefore, its overall correct diagnosis rate is only 51.45%, and is much lower in comparison to the diagnosis rate of the proposed method, which is 73.86%. Not only can better overall performance be obtained by the proposed approach, the diagnosis performance of each single mode, as will be also investigated, is more favorable.

Figure 5 summarizes the diagnosis results in the form of average posterior probabilities. The title of each plot denotes the true underlying mode, and the posterior probability corresponding to the true underlying mode is highlighted with light gray bars. The left panel summarizes the diagnosis results calculated by the approach ignoring evidence dependency; the right panel summarizes the diagnosis results obtained by the approach with consideration of evidence dependency. It is observed that for the three modes, the posterior probabilities assigned to the true underlying modes by the proposed approach are all higher than these assigned by the method ignoring dependency. Thus we can conclude that the proposed approach has better performance for diagnosis of all modes. This conclusion is confirmed by computing the correct diagnosis rate for each mode, as presented in Table 3.

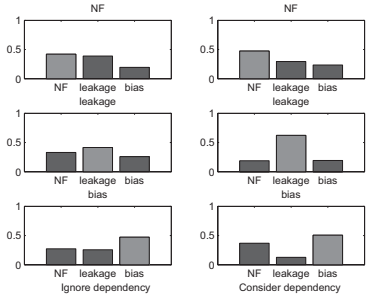


Fig. 5. Average posterior probability for each mode

Table 3. Correct diagnosis rate for each single mode

	NF	*leakage*	*bias*
Ignore evidence dependency	6.25%	73.75%	70%
Consider evidence dependency	55%	78.75%	92.5%

6. CONCLUSION

In this work, a data-driven approach considering evidence dependency is presented. Temporal dependency of monitor outputs is taken into consideration to obtain more accurate diagnosis results. The evidence transition probabilities are estimated from historical data with the developed data-driven algorithm. The method is applied to a pilot scale process, where the performance of the proposed approach is shown superior to that of the method ignoring evidence dependency. In summary, the more information from the time domain is synthesized, the better diagnosis performance is expected.

REFERENCES

Ahmed, S., Huang, B., and Shah, S. (2009). Validation of continuous-time models with delay. *Chemical Engineering Science*, 64(3), 443–454.

Chien, C.F., Chen, S.L., and Lin, Y.S. (2002). Using bayesian network for fault location on distribution feeder. *IEEE Transactions on Power Delivery*, 17(13), 785–793.

Choudhury, M.A.A.S., Jain, M., and Shah, S. (2008). Stiction - definition, modelling, detection and quantification. *Journal of Process Control*, 18(3-4), 232–243.

Dey, S. and Stori, J. (2005). A bayesian network approach to root cause diagnosis of process variations. *International Journal of Machine Tools*, 45(1), 75–91.

Harris, T., Seppala, C., and Desborough, L. (1999). A review of performance monitoring and assessment techniques for univariate and multivariate control systems. *Journal of Process Control*, 9(1), 1–17.

Hoo, K., Piovoso, M., Sheneller, P., and Rowan, D. (2003). Process and controller performance monitoring: overview with industrial applications. *International Journal of Adaptive Control and Signal Processing*, 17, 635–662.

Huang, B. (2008). Bayesian methods for control loop monitoring and diagnosis. *Journal of Process Control*, 18(9), 826–838.

Huang, B. and Shah, S. (1999). *Performance assessment of control loops*. Springer-Verlag.

Hugo, A. (2006). Performance assessment of single-loop industrial controllers. *Journal of Process Control*, 16, 784–794.

Jelali, M. (2006). An overview of control performance assessment technology and industrial applications. *Control Engineering Practice*, 14(5), 441–466.

Mehranbod, N., Soroush, M., and Panjapornpon, C. (2005). A method of sensor fault detection and identification. *Journal of Process Control*, 15(3), 321–339.

Patwardhan, R. and Shah, S. (2002). Issues in performance diagnostics of model-based controllers. *Journal of Process Control*, 12(3), 413–427.

Pernestal, A. (2007). *A Bayesian approach to fault isolation with application to diesel engine diagnose*. Ph.D. thesis, KTH School of Electrical Engineering.

Qi, F. and Huang, B. (2008). Data-driven bayesian approach for control loop diagnosis. In *Proceeding of Amecican Control Conference*. Seattle, U.S.A.

Qin, S. (1998). Control performance monitoring - a review and assessment. *Computers Chem. Engng*, 23(2), 173–186.

Qin, S. and Li, W. (2001). Detection and identification of faulty sensors in dynamic process. *AIChE Journal*, 47(7), 1581–1593.

Schafer, J. and Cinar, A. (2004). Multivariable mpc system performance assessment, monitoring, and diagnosis. *Journal of process control*, 14(2), 113–129.

Steinder, M. and Sethi, A. (2004). Probabilistic fault localization in communication systems using belief networks. *IEEE/ACM Transactions on Networking*, 12(5), 809–822.

Data-based Fault Detection and Isolation Using Output Feedback Control ⋆

Benjamin J. Ohran [*] **Jinfeng Liu** [*]
David Muñoz de la Peña [***] **Panagiotis D. Christofides** [*,**,1]
James F. Davis [*]

[*] *Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095-1592 USA.*
[**] *Department of Electrical Engineering, University of California, Los Angeles, CA 90095-1592, USA.*
[***] *Departamento de Ingeniería de Sistemas y Automática, Universidad de Sevilla, Camino de los Descubrimientos S/N, 41092, Sevilla, Spain.*

Abstract: This work focuses on data-based fault detection and isolation (FDI) of nonlinear process systems. Working within the framework of controller-enhanced fault detection and isolation that we recently introduced, we address and solve an unresolved, practical problem. We consider the case where only output measurements are available and design appropriate state estimator-based output feedback controllers to achieve controller-enhanced fault detection and isolation in the closed-loop system. The necessary conditions for achieving fault detection and isolation using output feedback control are provided. We use a nonlinear chemical process example to demonstrate the applicability and effectiveness of the proposed method.

Keywords: Process control, process monitoring, state estimation, fault detection and isolation

1. INTRODUCTION

Advanced automation technology has changed how the chemical process industry operates in many ways. Over the last few decades, advancements in plant operations have led to higher efficiency and improved economics through better control and monitoring of process systems. These technological advances have resulted in process systems becoming increasingly automated, no longer requiring operators to open and close valves in order to manually perform process control. In general, there is a trend towards such"smart" plants that are capable of highly automated control with decision making at the plant level taking into account environmental, health, safety and economic considerations (Christofides et al. (2007)). With increased amounts of sensors and actuators, it becomes possible to design systems capable of detecting and handling process or control system abnormalities through fault-tolerant control (FTC) (see for example, Mhaskar et al. (2006, 2007)). This is an important area of research as abnormal situations cost U.S. industries over $20 billion each year (Nimmo (1995)). A key element of a successful FTC system is a fast, accurate method for detecting faulty process behavior and isolating its cause. The fault detection and isolation (FDI) problem is the focus of the present work.

In a previous work (Ohran et al. (2008)), we developed an FDI method that takes advantage of both model-based and data-based approaches. This method brought together elements of model-based controller design and statistical process monitoring. In this method, the controller is designed with the FDI scheme in mind in addition to stability and performance criteria. By enforcing an isolable structure in the closed-loop system, it becomes possible to perform FDI based on statistical evaluation of process measurements. The purpose of the present work is to further develop the approach proposed in Ohran et al. (2008) by relaxing the requirement of full state feedback control. Specifically, we consider the case where only output measurements are available and design appropriate state estimator-based output feedback controllers to achieve controller-enhanced fault detection and isolation in the closed-loop system. This is demonstrated using a nonlinear chemical process example to show the applicability and effectiveness of the proposed method.

2. PRELIMINARIES

2.1 Process system structure

We consider nonlinear process systems with the following general state-space description:

$$\dot{x} = f(x, u, d) \tag{1}$$

where $x \in \mathbb{R}^n$ is the vector of process state variables, $u \in \mathbb{R}^m$ is the vector of manipulated input variables and $d \in \mathbb{R}^p$ is the vector of p possible actuator faults or disturbances. Vector d is equal to zero when the system is under normal operating conditions. When fault k, with $k = 1, ..., p$ occurs, d_k can take any time-varying value. The approach of controller enhanced FDI was introduced in Ohran et al. (2008) as a method of dividing the state vector into a number of partially decoupled subvectors. These subvectors can be monitored using measured process data.

⋆ Financial support from NSF, CTS-0529295, is gratefully acknowledged.
[1] Corresponding author: P.D. Christofides, pdc@seas.ucla.edu

Based on their responses and the system structure enforced by the decoupling controller, it is possible to discriminate between individual faults or groups of faults. In order to understand the necessary structure to perform isolation, we review the definitions of the incidence graph, the reduced incidence graph and the isolability graph (Ohran et al. (2008)).

Definition 1. The incidence graph of the system of Eq.1 is a directed graph defined by n nodes, one for each state, $x_i,\ i = 1 \dots n$, of the system. A directed arc with origin in node x_i and destination in node x_j exists if and only if $\frac{\partial f_j}{\partial x_i} \neq 0$.

The arcs in the incidence graph illustrate dependencies within the states of the system. A path through more than one arc that starts and ends at the same node is denoted as a loop.

Definition 2. The reduced incidence graph of the system of Eq.1 is the directed graph of N nodes, one for each q_i, $i = 1 \dots N$, where N is the maximum number of nodes that satisfy the following conditions:

- Each node q_i corresponds to a set of states $X_i = \{x_j\}$. These sets of states are a partition of the state vector of the system, i.e.,
$$\bigcup X_i = \{x_1, \dots x_n\}, \quad X_i \bigcap X_j = \emptyset,\ \forall i \neq j.$$
- A directed arc with origin q_i and destination q_j exists if and only if $\frac{\partial f_l}{\partial x_k} \neq 0$ for some $x_l \in X_i$, $x_k \in X_j$.
- There are no loops in the graph.

The reduced incidence graph reveals the partially decoupled subsystems within the structure of the states in x.

Definition 3. The isolability graph of the system of Eq.1 is a directed graph made of the N nodes of the reduced incidence graph and p additional nodes, one for each possible fault d_k. In addition, a directed arc with origin in fault node d_k and destination to a state node q_j exists if and only if $\frac{\partial f_l}{\partial d_k} \neq 0$ for some $x_l \in X_j$.

These definitions present the basic dependencies within a state vector. In most nonlinear process systems, the states are fully coupled and the isolability graph contains a single node representing all of the states in the system. However, in systems with partially decoupled dynamics these figures demonstrate graphically the subsets of the state vector.

With the isolability graph of a system, we can perform fault isolation based upon monitoring the subsystems. For this purpose, it is necessary to review the definition of a fault signature given below (Ohran et al. (2008)):

Definition 4. The signature of a fault d_k of the system of Eq.1 is a binary vector W^k of dimension N, where N is the number of nodes of the reduced incidence graph of the system. The i^{th} component of W^k, denoted W^k_i, is equal to 1 if there exists a path in the isolability graph from the node corresponding to fault d_k to the node q_i corresponding to the set of states X_i, or 0 otherwise.

2.2 Process monitoring

For the purpose of monitoring whether or not a state has deviated from its normal behavior, we use statistical process monitoring methods. Specifically, we use Hotelling's T^2 statistic developed in Hotelling (1947), a well established method in statistical process control that monitors multivariate normal (Gaussian) data using a single statistic. Because of its suitability for continuous, serially correlated chemical processes, the method of using single observations is employed (Tracy et al. (1992)). Given a multivariate state vector x of dimension n, the T^2 statistic can be computed using the mean $\bar{x}$ and the estimated covariance matrix S of process data obtained under normal operating conditions (see, for example, Kourti and MacGregor (1996)), as follows:

$$T^2 = (x - \bar{x})^T S^{-1} (x - \bar{x}). \tag{2}$$

The upper control limit (UCL) for the T^2 statistic can be calculated from its distribution, under the assumption that the data are multivariate normal, according to the following formula:

$$T^2_{UCL} = \frac{(h^2 - 1)n}{h(h - n)} F_\alpha(n, h - n) \tag{3}$$

where h is the number of historical measurements used in estimating S, $F_\alpha(n, h-n)$ is the value on the F distribution with $(n, h - n)$ degrees of freedom for which there is probability α of a greater or equal value occurring.

In order to perform FDI, the T^2 statistic based on the full state vector x with upper control limit T^2_{UCL} is first used to detect the presence of a fault. Subsequently, the T^2_i statistic is used to monitor the status of each subset of the state vector with an upper control limit T^2_{UCLi} where $i = 1, \dots, N$ that is based on each of the subvectors and their states $x_j \in X_i$. The fault detection and isolation procedure then follows the steps given below (Ohran et al. (2008)):

1. A fault is detected if $T^2(t) > T^2_{UCL}\ \forall t\ t_f \leq t \leq t_f + T_P$ where t_f is last time when T^2 crossed the UCL and T_P is the fault detection window chosen. Choosing T_P depends on the process time constants and on historical information of past process behavior.
2. Fault isolation can be performed by comparing fault signatures with the process signature $W(t_f, T_P)$ which can be built as follows:
$T^2_i(t) > T^2_{UCLi}\ \forall t\ t_f \leq t \leq t_f + T_P \rightarrow W_i(t_f, T_P) = 1.$
$T^2_i(t) \not> T^2_{UCLi}\ \forall t\ t_f \leq t \leq t_f + T_P \rightarrow W_i(t_f, T_P) = 0.$
A fault d_k is isolated at time $t_f + T_P$ if $W(t_f, T_P) = W^k$. If two or more faults are defined by the same signature, further isolation between them is not possible on the basis of the fault signature.

2.3 Controller design for enhanced FDI

Decoupling controller design The approach to fault detection and isolation discussed in the previous section can be applied if the signatures of the faults in the closed-loop system are distinct. The uniqueness of a fault depends on the structure of the closed-loop system and the faults considered. In general, complex nonlinear systems are fully coupled (i.e., cannot be broken down into partially decoupled subvectors). However, an isolable structure in the closed-loop system may still be achieved through the application of appropriately designed nonlinear control laws. As an example, consider a controller that can be

applied to nonlinear systems with the following state space description:

$$\begin{aligned} \dot{x}_1 &= f_{11}(x_1) + f_{12}(x_1, x_2) + g_1(x_1, x_2)u + d_1 \\ \dot{x}_2 &= f_2(x_1, x_2) + d_2 \end{aligned} \tag{4}$$

where $x_1 \in R$, $x_2 \in \mathbb{R}^n$, $u \in \mathbb{R}$ and $g_1(x_1, x_2) \neq 0$ for all $x_1 \in R$, $x_2 \in \mathbb{R}^n$. With a nonlinear state feedback controller of the form:

$$u(x_1, x_2) = -\frac{f_{12}(x_1, x_2) - v(x_1)}{g_1(x_1, x_2)} \tag{5}$$

the closed-loop system takes the form

$$\begin{aligned} \dot{x}_1 &= f_{11}(x_1) + v(x_1) + d_1 \\ \dot{x}_2 &= f_2(x_1, x_2) + d_2 \end{aligned} \tag{6}$$

where $v(x_1)$ has to be designed in order to achieve asymptotic stability of the origin of the x_1 subsystem when $d_1 = 0$. In this case, the controller of Eq.5 guarantees asymptotic stability of the closed-loop system, as well as different signatures for faults d_1 and d_2. For more detailed results, see Ohran et al. (2008).

Input/output linearizable nonlinear systems Input/output linearizable nonlinear systems constitute a special class of nonlinear systems for which it is possible to systematically design nonlinear controllers to achieve controller-enhanced fault detection and isolation. Using a feedback-linearizing control law that takes the following general form,

$$u(x) = \frac{1}{L_g L_f^{r-1} h(x)} [v(x) - L_f^r h(x)] \tag{7}$$

where $L_f^r h(x)$ is the r-th order Lie derivative, $L_g L_f^{r-1} h(x)$ is a mixed Lie derivative and $v(x)$ is an external controller for the purpose of stabilizing the system, the system under closed loop operation will have linear input-output dynamics.

If the state-feedback law given in Eq.7 is applied to an input/output linearizable system, faults affecting the system can be isolated into two different groups: those that affect the output and those that do not affect the output. The induced structure of the closed-loop system provides different signatures for the faults depending on the relative degree of the output with respect to the fault and the relative degree of the output with respect to the input. Faults with relative degree higher than the relative degree of the input will not affect the output. Thus, when a fault occurs, taking into account whether the trajectory of the output has deviated from the normal case or not, it is possible to isolate to which group the fault belongs. For the definitions of relative degree and an in depth discussion of feed-back linearization in this context, see Ohran et al. (2008).

3. CONTROLLER ENHANCED FDI USING OUTPUT FEEDBACK CONTROL

3.1 State estimation

In order to perform controller enhanced FDI using output feedback control, any unknown process state variable must be quickly and accurately estimated from the available output measurements so that the decoupling state feedback controller designs of subsections 2.3.1 and 2.3.2 can be implemented. The state estimation is performed for the state vector x (or a subset thereof) with the outputs, or measured states, defined as $y = Cx$. In this work, we consider only outputs of the form $y_i = x_i$, $i = 1, \ldots, q < n$. In other words, C is a matrix with one and only one non-zero entry in each row and that entry is equal to unity. This set-up is appropriate in chemical process control applications where measurements of a few states like temperature and concentrations of a few species, like key products, are available, but concentrations of some species are not measured. This set-up also allows obtaining a clear picture of the use of output feedback instead of full state feedback in controller enhanced FDI. The theory for the state estimator design is based upon a linear system, but can also be applied to nonlinear systems, using a local stability analysis around the operating point (origin). Specifically, the linearized model of the nonlinear system of Eq.1 takes the following form:

$$\begin{aligned} \dot{x} &= Ax + Bu + Wd \\ y &= Cx \end{aligned} \tag{8}$$

where A is the Jacobian matrix of the nonlinear system at the operating point, u is the manipulated input vector and d is the fault vector. The matrices B and W can be computed from the linearization of Eq.1 around the origin. Under the assumption that (A, C) forms an observable pair, each state variable x can be estimated by the following dynamic equation:

$$\dot{\hat{x}} = A\hat{x} + Bu + L(y - C\hat{x}) \tag{9}$$

where $\hat{x}$ is the state estimate and L is the estimator gain that can be chosen so that all the eigenvalues of the matrix $(A - LC)$ are placed at appropriate locations in the left-half of the complex plane to guarantee a desirable rate of convergence of the estimation error to zero. The computation of L can be done using standard pole placement techniques or via a Kalman filtering framework by adding process and measurement noise in the linearized model of Eq.8. In either case, the linearized state estimation error equation with $d(t) = 0$ takes the form:

$$\dot{e} = (A - LC)e. \tag{10}$$

where $e = x - \hat{x}$ is the estimation error. While it is possible to perform state estimation using the full state vector in the state estimator of Eq.9 when $d(t) \equiv 0$, it becomes necessary to use a reduced-order process model when designing a state estimator-based output feedback controller to enhance FDI. This need for a reduced-order model arises due to faults that affect the state estimator and introduce error into the estimate (i.e., the full state estimation scheme of Eq.9 works when $d(t) = 0$, but not when $d(t) \neq 0$). Specifically, if the error vector d on the right-hand side of Eq.8 is nonzero, the new equation for the estimator error becomes $\dot{e} = (A - LC)e + Wd$. Thus, in the presence of a fault, the state estimates no longer converge to their actual values, and the isolable structure attained in the closed-loop system under state feedback control cannot be maintained. However, it is possible in some process systems to perform the state estimation task using a subset of the states that are not directly affected by the expected faults, i.e., effectively eliminating d in the estimation error system. The general structure of the model in Eqs.8-10 remains the same for the reduced-order system, but it is based on a subset of the full state vector, $x_r \subset x$. To mathematically realize this notion, consider a system with the following structure, where time derivatives of the states

x_r are not functions of d and include all unknown states to be estimated along with some measured states, and x_d includes the remaining measured states, whose dynamic equations may be functions of d. Specifically, we consider the following decomposition of the vectors and matrices of the linearized system of Eq.8

$$x = \begin{bmatrix} x_r \\ x_d \end{bmatrix},\ A = \begin{bmatrix} A_r & A_{rd} \\ A_{dr} & A_d \end{bmatrix},\ W = \begin{bmatrix} 0 \\ W_d \end{bmatrix}$$
$$B = \begin{bmatrix} B_r \\ B_d \end{bmatrix},\ C = \begin{bmatrix} C_r & 0 \\ 0 & C_d \end{bmatrix},\quad y = \begin{bmatrix} y_r \\ y_d \end{bmatrix}. \tag{11}$$

Provided that the pair (A_r, C_r) is observable, the state estimator based on the reduced-order system then takes the form:

$$\dot{\hat{x}}_r = A_r\hat{x}_r + A_{rd}x_d + B_r u + L_r(y_r - C_r\hat{x}_r) \tag{12}$$

Eq.12 uses the actual measured values for all of the states in x_d. We can break x_r down further into measured states and unmeasured states, $x_r = [x_{rm}^T\ x_{ru}^T]^T$. Note that x_{rm} must include enough measured states independent of d for the system to be observable. Given the restrictions on C, this implies that $y_r = C_r x_r = x_{rm}$ and $C_d = I$ (i.e., $y_d = x_d$). Finally, we define a vector with full state information by combining the measured and estimated data, $\hat{x} = [x_{rm}^T\ \hat{x}_{ru}^T\ x_d^T]^T$. Note that $\hat{x}_{rm}$ is only used as the driving force for convergence of the state estimator. With these definitions, the reduced-order state estimator of Eq.12 is not a direct function of d and the dynamics of the estimation error, $e_r = x_r - \hat{x}_r$, take the form $\dot{e}_r = (A_r - L_r C_r)e_r$ which implies that $e_r(t)$ will converge to zero even in the presence of a change in d.

Once the estimator gain obtained from the linearized model of the system is calculated, it can then be used to estimate the states of the process using the nonlinear model dynamics. Once again, for the nonlinear system, the state vector, x, decomposes into the one of the reduced-order system (independent of d) and the remaining states, i.e., $x = [x_r^T\ x_d^T]^T$ and $f([x_r^T\ x_d^T]^T, u, d) = [f_r(x_r, x_d, u)^T\ f_d(x_r, x_d, u, d)^T]^T$. The nonlinear dynamic equations for the reduced-order system are then combined with the estimator gain and the output error to create a nonlinear state estimator as follows:

$$\dot{\hat{x}}_r = f_r(\hat{x}_r, x_d, u) + L_r(y_r - h_r(\hat{x}_r)) \tag{13}$$

where the measured values are used for the states in x_d, i.e., by assumption $y_d = x_d$. Note that following the previous assumption, $h_r(x_r) = C_r x_r$. Combining the nonlinear state estimator of Eq.13 with a nonlinear state feedback controller, $u = p_{DC}(x)$, that enforces an isolable structure in the closed-loop system and can be designed following the approaches presented in subsections 2.3.1 and 2.3.2, we obtain the following dynamic nonlinear output feedback controller:

$$\begin{aligned} \dot{\hat{x}}_r &= f_r(\hat{x}_r, x_d, p_{DC}(\hat{x})) + L_r(y_r - C_r\hat{x}_r) \\ u &= p_{DC}(\hat{x}) \end{aligned} \tag{14}$$

Due to the effect of estimation error, it is not possible to achieve complete decoupling. However, it is possible to achieve a near isolable structure that is sufficient for practical purposes. In this sense, we consider a near isolable structure to be one where the closed-loop system under output feedback control can be seen as an $O(e_r)$ regular perturbation of the closed-loop system under state feedback control which is locally exponentially stable and has an isolable structure. Thus, the estimation error can be viewed a small perturbation error that will be accounted for by the FDI thresholds designed to filter out normal process variation. Theorem 1 below summarizes the main analysis and controller design result of this section as well as the closed-loop FDI properties.

Theorem 1. Consider the closed-loop system of Eq.1 under the nonlinear output feedback controller of Eq.14 and assume that the pair (A_r, C_r) is observable and L_r is designed such that the matrix $(A_r - L_r C_r)$ has all of its eigenvalues in the left-half of the complex plane. Then, there exist δ, ϵ and T_y such that if f is continuously differentiable on $D = \{x \in \mathbb{R}^n |\ \|x\|_2 < \delta\}$, the Jacobian of f is bounded and Lipschitz on D and $\max\{\|x(t_0)\|_2, \|\hat{x}_r(t_0)\|_2\} < \delta$ then $\|x_r(t) - \hat{x}_r(t)\|_2 < \epsilon,\ \forall t > t_0 + T_y$, and a near isolable structure is enforced in the closed-loop system.

Proof. Under the control law of Eq.14, the closed-loop system of Eq.1 takes the form,

$$\begin{aligned} \dot{x} &= f(x, p_{DC}(\hat{x}), d),\ \ y = h(x) \\ \dot{\hat{x}}_r &= f_r(\hat{x}_r, x_d, p_{DC}(\hat{x})) + L_r(y_r - h_r(\hat{x}_r)). \end{aligned} \tag{15}$$

Linearizing the closed-loop system of Eq15 around the equilibrium point (origin) yields,

$$\dot{x} = Ax + Bp_{DC}(\hat{x}),\ \ y = Cx \tag{16}$$
$$\dot{\hat{x}}_r = A_r\hat{x}_r + A_{rd}x_d + B_r p_{DC}(\hat{x}) + L_r(y_r - C_r\hat{x}_r). \tag{17}$$

The error between the actual and estimated states of the reduced-order, linearized system is then $e_r = x_r - \hat{x}_r$ with the dynamics $\dot{e}_r = (A_r - L_r C_r)e_r$. Assuming that the pair (A_r, C_r) is observable and that L_r is chosen such that the matrix $A_r - L_r C_r$ has eigenvalues in the left-half of the complex plane, the estimation error, e_r, in the linearized system has exponentially stable dynamics. If the vector field of the nonlinear system, $f(x, p_{DC}(\hat{x}), d)$, is continuously differentiable and the Jacobian matrix is bounded and Lipschitz on $D = \{x \in \mathbb{R}^n |\ \|x\|_2 < \delta\}$, then the nonlinear system dynamics are also locally, exponentially stable within some region around the equilibrium point Khalil (1992). For some initial condition $max\{\|x_0\|_2, \|x_{r0}\|_2\} < \delta$, the state estimation error, e_r, will be bounded such that $\|x_r - \hat{x}_r\| < \epsilon\ \forall t > t_0 + T_y$, where T_y is a time interval of $O(\epsilon)$. Thus, the output feedback control approaches state feedback control with error of order ϵ, i.e., $x_r = \hat{x}_r + O(\epsilon)\ \forall t > t_0 + T_y$. For sufficiently small ϵ, this leads to a near isolable structure in the closed-loop system for almost all times since the state feedback controller $p_{DC}(x)$ enforces an isolable structure in the closed-loop system.

3.2 Application to a CSTR example

The example considered is a well-mixed CSTR in which a feed component A is converted to an intermediate species B and finally to the desired product C, according to the reaction scheme

$$A \overset{1}{\rightleftharpoons} B \overset{2}{\rightleftharpoons} C.$$

Both steps are elementary, reversible reactions and are governed by the following Arrhenius relationships:

$$r_1 = k_{10} e^{\frac{-E_1}{RT}} C_A, \;\; r_{-1} = k_{-10} e^{\frac{-E_{-1}}{RT}} C_B \tag{18}$$

$$r_2 = k_{20} e^{\frac{-E_2}{RT}} C_B, \;\; r_{-2} = k_{-20} e^{\frac{-E_{-2}}{RT}} C_C \tag{19}$$

where k_{i0} is the pre-exponential factor and E_i is the activation energy of the i^{th} reaction where the subscripts $1, -1,\ 2, -2$ refer to the forward and reverse reactions of steps 1 and 2. R is the gas constant, while C_A, C_B and C_C are the molar concentrations of species A, B and C, respectively. The feed to the reactor consists of pure A at flow rate F, concentration C_{A0} and temperature T_0. The state variables of the system include the concentrations of the three main components C_A, C_B, and C_C as well as the temperature of the reactor, T. Using first principles and standard modeling assumptions, the following mathematical model of the process is obtained

$$\begin{aligned}
\dot{C_A} &= \frac{F}{V}(C_{A0} - C_A) - r_1 + r_{-1} + d_1 \\
\dot{C_B} &= -\frac{F}{V}C_B + r_1 - r_{-1} - r_2 + r_{-2} \\
\dot{C_C} &= -\frac{F}{V}C_C + r_2 - r_{-2} \\
\dot{T} &= \frac{F}{V}(T_0 - T) + \frac{(-\Delta H_1)}{\rho c_p}(r_1 - r_{-1}) \\
&\quad + \frac{(-\Delta H_2)}{\rho c_p}(r_2 - r_{-2}) + u + d_2
\end{aligned} \tag{20}$$

where V is the reactor volume, ΔH_1 and ΔH_2 are the enthalpies of the first and second reactions, respectively, ρ is the fluid density, c_p is the fluid heat capacity, $u = Q/\rho c_p$ is the manipulated input, where Q is the heat input to the system, d_1 denotes a disturbance in the inlet concentration and d_2 denotes a fault in the control actuator. The system of Eq.20 is modeled with sensor measurement noise and autoregressive process noise. For details on noise generation and for complete system parameter values, please refer to Ohran et al. (2008).

In order to obtain the estimated trajectory for C_B, a state estimator as in Eq.13 was implemented using the reduced-order system $\hat{x}_r = [\hat{C}_B\ \hat{C}_C]^T$. The process measurements for C_A and T were used in computing the dynamics of $\hat{x}_r$. Note that although C_C is measured, it is used in the reduced-order state estimator so that the reduced-order system is observable. The control input was updated at each sampling interval with the measured values for C_A, T and C_C and the estimated value of $\hat{C}_B$. As discussed in subsection 3.1, C_A and T should not be modeled as dynamic states in the estimator since they are directly affected by the faults d_1 and d_2. Thus, the measured data for C_A and T must be used in modeling the estimator, and the final form of the state estimator based on the reduced subsystem $\hat{x}_r = [\hat{C}_B\ \hat{C}_C]^T$ is as given below:

$$\begin{aligned}
\dot{\hat{C}}_B &= -\frac{F}{V}\hat{C}_B + r_1 - r_{-1} - r_2 + r_{-2} + L_1(C_C - \hat{C}_C) \\
\dot{\hat{C}}_C &= -\frac{F}{V}\hat{C}_C + r_2 - r_{-2} + L_2(C_C - \hat{C}_C)
\end{aligned} \tag{21}$$

with

$$r_1 = k_{10} e^{\frac{-E_1}{RT}} C_A, \;\; r_{-1} = k_{-10} e^{\frac{-E_{-1}}{RT}} \hat{C}_B$$
$$r_2 = k_{20} e^{\frac{-E_2}{RT}} \hat{C}_B, \;\; r_{-2} = k_{-20} e^{\frac{-E_{-2}}{RT}} \hat{C}_C$$

where L is the filter gain obtained using Kalman-filtering theory based on the reduced-order system.The resulting value for L_r is $[L_{r1}\ L_{r2}]^T = [0.0081\ 0.0559]^T$.

The controlled output of the system, for the purpose of feedback linearization, is defined as the concentration of the desired product $y = h(x) = C_C$ (although, the measured output vector is $y_m = [C_A\ T\ C_C]^T$.) We consider only faults d_1 and d_2, which represent undesired changes in C_{A0} (disturbance) and Q (actuator fault), respectively. In this process, the manipulated input u appears in the temperature dynamics and the output, $y = C_C$, has relative degree 2 with respect to u. The fault d_1 appears only in the dynamics of C_A and the output, $y = C_C$, has relative degree 3 with respect to d_1. Finally, the output has relative degree 2 with respect to d_2. Based on the relative degrees of the output with respect to the input and with respect to the faults, under feedback linearizing control the system structure will be such that the state vector can be separated into two subsets: $X_1 = \{C_A, \hat{C}_B, T\}$ and $X_2 = \{C_C\}$. Thus, the fault signature for $d_1 = [1\ 0]^T$ and for $d_2 = [1\ 1]^T$. During the simulation, the T^2 for the full state vector is monitored in order to perform fault detection (substituting the estimate $\hat{C}_B$ for the unknown state C_B.) Each of the subsystems is monitored to compute the system signature upon detection of a fault. Based on observation of the system dynamic behavior, a fault detection window, T_P, of 1 min is used.

The control objective is to regulate the system at the equilibrium point

$$\begin{aligned}
&C_{As} = 2.06\frac{kmol}{m^3},\ C_{Bs} = 1.00\frac{kmol}{m^3},\ C_{Cs} = 0.937\frac{kmol}{m^3}, \\
&T_s = 312.6K,\ u_s = 0K/s
\end{aligned} \tag{22}$$

where the subscript s refers to the steady state values of the variables. It should be noted that the CSTR system of Eq.20 belongs to the class of systems of Eq.1 with $x = [C_A - C_{As},\ T - T_s,\ C_B - C_{Bs},\ C_C - C_{Cs}]^T$ where C_B is replaced with $\hat{C}_B$ in the definition of $\hat{x}$. This implies that we can apply the output feedback scheme presented using the controlled output $y = C_C$. Using Eq.7, the feedback-linearizing controller takes the following form:

$$u = \frac{v - L_f^2 h(\hat{x})}{L_g L_f h(\hat{x})} \tag{23}$$

with

$$v = [-2\zeta_1 - 2\zeta_2].$$

where

$$\begin{aligned}
\zeta_1 &= C_C, \;\; \zeta_2 = -\frac{F}{V}C_C + r_2 - r_{-2} \\
r_2 &= k_{20} e^{\frac{-E_2}{RT}} \hat{C}_B, \;\; r_{-2} = k_{-20} e^{\frac{-E_{-2}}{RT}} C_C.
\end{aligned}$$

The state variables are in the transformed space and are shifted so that the origin represents the desired set-point.

The closed-loop system was simulated for each of the two faults considered. Each simulation was run for a process time of 1 hour with the fault occurring at $t = 40$ min. The values for the faults were each zero prior to the fault

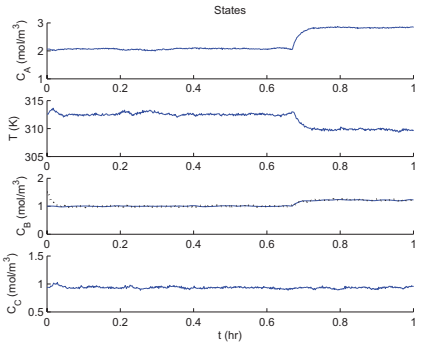


Fig. 1. Plot of measured state values for the CSTR under output feedback decoupling control with fault d_1. C_B shows both actual (solid) and estimated (dotted) values.

occurring and took constant values of $d_1 = 1\ kmol/m^3 min$ and $d_2 = 10\ K/min$ at $t = 40\ min$. The state estimator was initialized far from the operating point at $\hat{C}_B(0) = 1.5\ kmol/m^3$ and $\hat{C}_C(0) = C_C(0) = C_{Cs}$ in order to demonstrate convergence.

Figure 1 shows the trajectories for each of the states in the simulation with a failure in d_1. The fault is apparent at approximately $t = 40\ min$ $(0.667hr)$. We can readily see from the state trajectories, that the decoupling scheme was effective as evidenced by the fact that the output, C_C, is unaffected by the fault. Also, we see that the state estimator converged at around $t = 3\ min$.

For the system with a failure in d_1, Figure 2 shows the Hotelling's T^2 statistic for the two subvectors X_1 and X_2 as well as for the full state vector. From the graph, we can see that a fault is clearly detected at the expected time $t = 40\ min$ as shown in the plot of the T^2 statistic for the full state vector (T_3^2). Although there were a few single incidents of data breaching the upper control limit, none of them represented sustained departures for the length of the fault detection window, T_P. Also note that values above the upper control limit before $t = 0.1hr$ were due to the state estimator not having converged. Upon detection of the fault, the system signature can be computed as $W = [1\ 0]^T$ due to the fact that the T^2 statistic for the subvector X_1 exceeded the upper control limit for a sustained period and the T^2 for the subvector X_2 remained within the bounds of normal operation. Because the system signature matches that of the fault signature for d_1, a fault in d_1 is declared at time $t \approx 41\ min$. In Figure 3, we see the simulation results for the same system with a failure in d_2. Again, the failure is evident around $t = 40\ min$. However, in this case we see that both subsystems are affected. The process signature obtained from the T^2 statistics in Figure 3 shows that both subvectors were affected and this process signature matches the fault signature of d_2.

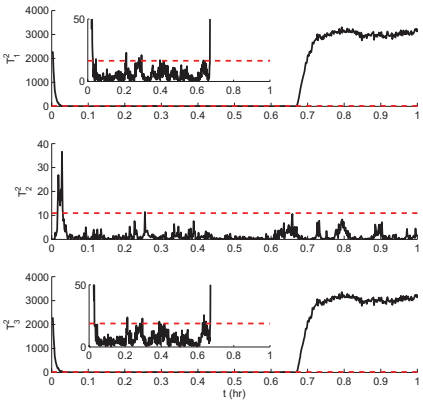


Fig. 2. T^2 statistics for the CSTR under output feedback decoupling control with fault d_1 for the subsystem X_1 (T_1^2), the subsystem X_2 (T_2^2) and the full system x (T_3^2).

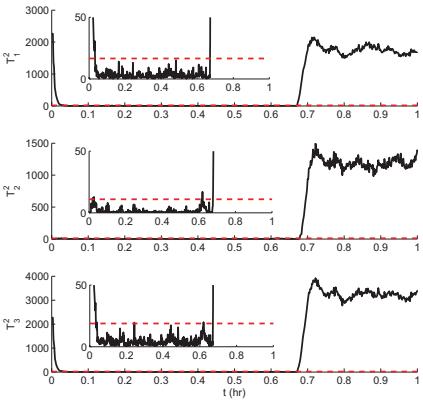


Fig. 3. T^2 statistics for the CSTR under output feedback decoupling control with fault d_2 for the subsystem X_1 (T_1^2), the subsystem X_2 (T_2^2) and the full system x (T_3^2).

REFERENCES

Christofides, P.D., Davis, J.F., El-Farra, N.H., Clark, D., Harris, K.R.D., and Gipson, J.N. (2007). Smart plant operations: Vision, progress and challenges. *AIChE Journal*, 53, 2734–2741.

Hotelling, H. (1947). Multivariate quality control. In O. Eisenhart (ed.), *Techniques of Statistical Analysis*, 113–184. McGraw-Hill.

Khalil, H. (1992). *Nonlinear Systems*. Macmillan Publishing Comapny.

Kourti, T. and MacGregor, J. (1996). Multivariate SPC methods for process and product monitoring. *Journal of Quality Technology*, 28, 409–428.

Mhaskar, P., Gani, A., and Christofides, P.D. (2006). Fault-tolerant process control: Performance-based reconfiguration and robustness. *International Journal of Robust & Nonlinear Control*, 16, 91–111.

Mhaskar, P., Gani, A., McFall, C., Christofides, P.D., and Davis, J.F. (2007). Fault-tolerant control of nonlinear systems subject to sensor data losses. *AIChE Journal*, 53, 654–668.

Nimmo, I. (1995). Adequately address abnormal operations. *Chem. Eng. Prog.*, 91, 36–45.

Ohran, B.J., Muñoz de la Peña, D., Christofides, P.D., and Davis, J.F. (2008). Enhancing data-based fault isolation through nonlinear control. *AIChE Journal*, 54, 223–241.

Tracy, N.D., Young, J.C., and Mason, R.L. (1992). Multivariate control charts for individual observations. *Journal of Quality Technology*, 24, 88–95.

Monitoring and Hybrid Control of Industrial Processing Systems (Invited Session)

Oral Session

Data reconciliation and optimal management of hydrogen networks of a petrol refinery

D. Sarabia*, S. Cristea*, E. Gómez*, G. Gutierrez*, C.A. Méndez, J.M. Sola***, C. de Prada***

**University of Valladolid, c/ Real de Burgos s/n, 47011 Valladolid, Spain (Tel: 34-983184647; e-mail: dsarabia@autom.uva.es).*
*** INTEC (UNL-CONICET), Güemes 3450, 3000 Santa Fe,Argentina (e-mail: cmendez@intec.unl.edu.ar)*
**** PETRONOR, Edificio Muñatones, San Martin, 5, 48550 Muskiz, Spain (e-mail: jmsola@repsol.com).*

Abstract: This paper describes the main problems associated to the management of hydrogen networks in petrol refineries and presents an approach to deal with them with the aim of operating the installation in the most profitable way. In particular, the problems of data reconciliation, economic optimization and interaction with the underlying basic control structure are reviewed. The paper provides also a proposal for the implementation of the system and illustrates the approach with results obtained using real data from an industrial site.

Keywords: Hydrogen networks, Process optimization, Data reconciliation, Control of networks.

1. INTRODUCTION

Hydrogen has become one of the main products in petrol refineries due to several factors, among them the new legislation about the reduction in the polluting compounds content (sulphur, nitrogen, aromatics, etc.), the need to convert heavy into light products to improve the economic balance of the refineries, and the installation of platformer plants, with the purpose of increasing the octane degree of the gasoline, as an alternative path to the use of lead compounds, operations that involve the use of large amounts of hydrogen.

As a result, hydrogen management plays a key role in the production of the different commercial oil fractions. Three different types of units are involved in a typical plant: dedicated hydrogen production units, hydrogen consuming units, and production units where hydrogen is the by-product of another process. All these kinds of units are interconnected through a hydrogen pipeline network.

Hydrogen is quite often produced on site from hydrocarbons in reformer ovens Control of these units, its temperature in particular, is not easy and MPC is frequently used to direct its operation. Demands of H_2 from the consumers change from time to time and constitute at the same time, a disturbance to reject with respect to the H_2 purity and a target to follow with regards to the mass flow. Adaptation to these demands plays an important role in order to operate with minimum losses while satisfying the orders from other units. The hydrogen production is fed at a given pressure and purity to the hydrogen pipeline network for distribution.

Most of the consumer plants have as a goal the desulphurization of different oil fractions and are named with the acronym HDS. They receive a mixture of hydrocarbons which react with H_2 at the appropriate temperatures and with the adequate catalysers in the HDS reactors. In order to secure the life of the catalysers, a given excess ratio hydrogen / hydrocarbons must be kept on the reactors. The surplus hydrogen from the reactors is separated and partially recycled, the excess being sent to the fuel-gas network. In the recovery of H_2, flash units can be employed as well as special membranes, which are used to separate high-purity H_2 from other gases.

The operation of the HDS is quite complex and its is affected by different disturbances, in particular the supply of hydrocarbons that may change in quantity as well as in composition according to the type of crude being processed and the global production aims. Important operating constraints are linked to the hydrogen / hydrocarbon ratio and to the operation of the compressors that maintain the hydrogen flows and inject it from the H_2 distribution network. This one should be able to provide the required amounts requested by the changing operation of the HDS along time.

There are a final set of plants, mainly platformers, which increase the octane index of the gasoline (catalytic reforming process) and generate hydrogen as result of these reactions. These plants generate a positive net flow of low purity hydrogen (between 75 % and 85 %) as a by-product, which is incorporated to the hydrogen pipeline network for use in other plants. Being a secondary product, there is no direct control of the H_2 production, so that it can be considered as a disturbance in flow and purity from the point of view of the network conditions.

All these types of plants are interconnected by several kilometres diverse pipes forming a distribution network with different purities, capacities and operating at several pressures. Fig. 1 displays the structure of one of such networks with three main hydrogen collectors, high purity

manifold (C-H4), medium purity manifold (C-H3) and low purity manifold (C-BP). The boxes represent the different types of production (H3 and H4), consumer and net production units (P1 and P2). Production and net production units dump hydrogen to the collectors (C-H4, C-H3, C-P1N1 and C-P2N2), while the HDS are fed from the different sources according to the choice of the operators.

In the picture we can see also the outputs from the plants to fuel-gas network, where the excess hydrogen is sent to be consumed in furnaces. Part of this flow also comes from the pressure controllers of the collectors (i.e. from manifold C-BP on the left). In order to guarantee that enough hydrogen is available to the consumer units when need it, a surplus of it must be maintained in the collectors, the excess being released by the pressure controllers to the fuel-gas network.

Hydrogen networks have received attention in the literature from the point of view of its (re)design, but very few from the one of real-time operation and, as far as we know, no commercial software is available in the sector for this purpose. The most used method of analysis is the so-called hydrogen pinch to evaluate the scope for hydrogen savings, Alves (1999). On the other hand, Hallale and Liu (2001) developed an improved methodology for hydrogen network retrofit that considers pressure constraints as well as the existing compressors.

To improve the day-to-day operation of the whole hydrogen supply network, this paper presents an integrated optimization based framework to optimize the distribution of the available hydrogen from producers to consumer facilities, so as to take advantage of low purity hydrogen supplies by combining streams of different purity levels and flows and, at the same time, ensuring operational restrictions. This work is carried out within an industrial project in close collaboration with Petronor, an oil refining company belonging to the Repsol-YPF Group, Spain. The major aim of the project is to provide an effective and integrated decision support system for on-line, open-loop optimization and data reconciliation. The proposed optimization tool has been validated with real-world data provided by the Petronor refinery.

The paper is organized as follows: after the introduction, the main problems of the hydrogen network operation and a proposal for a decision support system (DSS) are described in section two. The formulation of the hydrogen network model is given in section three, then, data reconciliation and hydrogen optimal management problems are described in sections four and five respectively while results obtained using plant data are presented in section 6. The paper ends with some conclusions and a short bibliography.

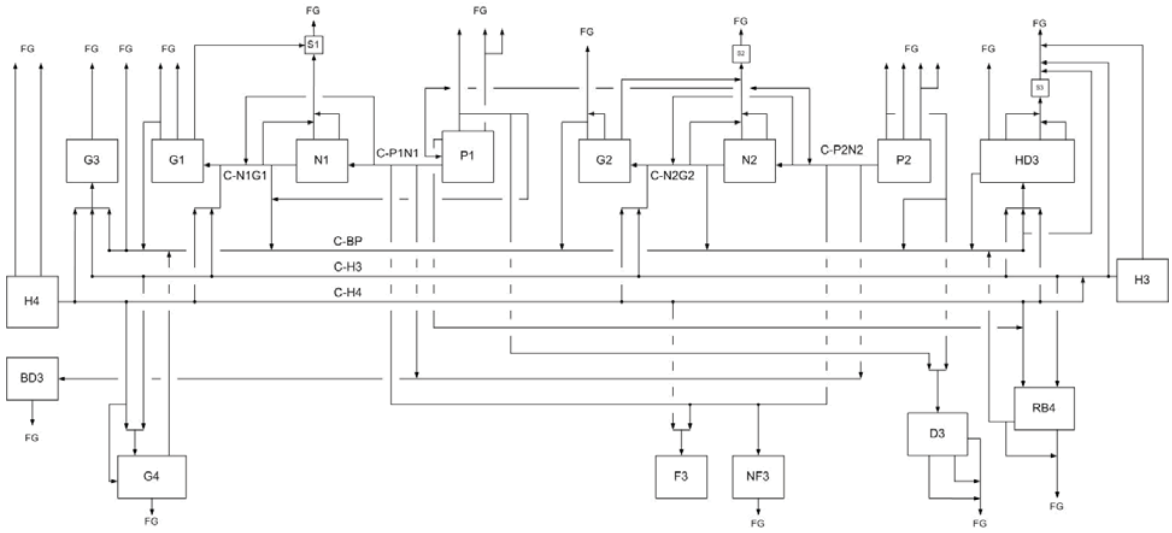


Fig. 1. A typical hydrogen network of a petrol refinery.

2. THE HYDROGEN NETWORK

2.1 Operational problems

Being a product heavily used and expensive to produce, optimizing the use of hydrogen is a clear target in any refinery. The problem can be formulated as of balancing the hydrogen that is being produced and consumed in the refinery and distribute it through the existing pipeline network in such a way that an economic target is optimized, while satisfying a set of operational constraints. Many aims appear as possible targets for the problem. For instance, minimizing the production of H_2, maximizing the use of lower cost hydrogen, minimizing the flow of H_2 to the fuel gas network, maximizing the use of low purity H_2, etc. the choice of one of them or a suitable combination being dependent of the particular situation of the refinery.

Several problems are related to the hydrogen management and optimization that are worth to mention, among them, the lack of reliable information about many streams and compositions and the large scale of the system that creates additional difficulties.

Regarding the first one, it is clear that reliable information of the network is required if we wish to perform optimal

decisions. Part of the uncertainty comes from the measurement system, but mainly from unmeasured variables and from partial measurements. In particular, gas flows, which are the main variables of the process, are measured usually in terms of volumetric flows that require compensation in order to be converted to mass or normalized flows required for the models, based on mass balances. This compensation involves pressure, temperature and molecular weight of the streams. Nevertheless, the last one are quite often non available, partly because of the price and reliability of the instruments measuring hydrogen purity and partly because the purity of the flows do not reflect directly its composition. With other gases this would not be a problem, but hydrogen has a molecular weight of only two, so that a small change in the composition of the (unmeasured) impurities, for instance from methane to propane, can have a significant impact on the molecular weight of the stream and hence on its mass flow.

Consequently, improving the information about the hydrogen network implies then the need of a data reconciliation system able to correct the readings of the process transmitters and estimate the unknown variables Cronkwright (2007).

Regarding the large scale of the system, it imposes computational barriers for a global solution of the problem. Firstly, because the size of the problem, but also for the wide range of time scales involved. The problem is dynamic in nature, being one of its aspects the need to adapt the rhythm of production of hydrogen to its consumption in order to minimize losses to the fuel-gas network. It operates with the changes in global production at the time scale of hours-days, changes in the operation of the producer and consumer units at the time scale of minutes-hours and the fast dynamics of the pressures and gas flows in the order of seconds. Trying to find solutions involving all these elements at the same time would be unrealistic, but the division in time scales allows separating the decision problems in different layers, facilitating in this way the solution as a set of linked sub-problems. The separation can be considered also from a functional point of view: producer and consumer units can perform local optimizations of its functioning provided that they have predictions of its future loads, while the optimal distribution of these loads must be performed in the network, which operates with a much faster dynamics and can be considered static in relation to the slower producer and consumer units.

Finally, notice that the implementation of optimally computed targets for the units and the distribution network will require a control layer that takes into account its dynamics and associated constraints. Alternatively, a decision support system (DSS) could give recommendations to the operators of the control room about these targets, being them the ones in charge of the implementation using the existing plant controllers.

2.2 Proposed architecture

In view of the above mentioned problems, the following supervisory architecture, depicted in Fig. 2, is proposed. It consists of four stages: the first one, corresponding to data reconciliation, allows fitting periodically the network and units models to the state of the plant. The second stage uses simplified models of the consumer units to compute the future profile of the hydrogen consumption at the unit hydrogen entrance, required to treat the future loads. This profile can be locally optimized or taken as the one corresponding to the current operating policy. The third stage considers the whole distribution network and, using a model of it, computes the optimal production profile of each production unit as well as the optimal distribution that satisfies the consumer units needs. Notice that, formulated in this way this problem can be considered as a series of constraint programming problems. Finally, the last stage is performed either by local MPC controllers (model predictive controllers) that implement the required distribution of flows, or as a DSS that gives the recommendations to the operators.

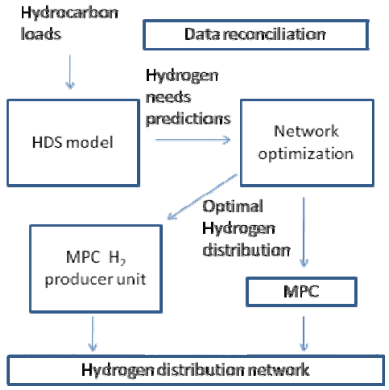


Fig. 2. A schematic of the proposed control and optimization system, with only one producer and one consumer unit.

In this paper, the reconciliation and optimization problems corresponding to the third stage are described assuming constant demands from the consumer units.

3. THE NLP MATHEMATICAL MODEL

All nodes in the complete hydrogen network are modelled by mass balances in terms of purity, flow and molecular weight for every stream, considering also a mixture of ideal gases. For example, a node consisting in one input stream F_1 and two output streams F_2 and F_3 with hydrogen purities X_1, X_2 and X_3 and molecular weights MW_1, MW_2 and MW_3 respectively is described by,

$$\begin{aligned} F_1 &= F_2 + F_3 \\ F_1 X_1 &= F_2 X_2 + F_3 X_3 \\ F_1 MW_1 &= F_2 MW_2 + F_3 MW_3 \end{aligned} \qquad (1)$$

On the other hand, the molecular weight of every stream is calculated from the hydrogen purity X, hydrogen molecular weight and molecular weights of all impurities MW^{imp} according to:

$$MW_i = (2X_i + MW_i^{imp}(100 - X_i))/100 \qquad i = 1,2,3 \qquad (2)$$

Volumetric flows F_i (Nm^3/h) in (1) are measured at standard temperature (0ºC) and pressure (1 Atm) conditions and purities are measured in percentage of volume.

In the particular industrial case considered, the nonlinear model consists of 142 equations like (1) and (2) and 263 variables, 137 of them are flows, 42 are purities, 42 are molecular weights of the streams and 42 are molecular weights of impurities of every stream. From a mathematical point of view it is necessary to define 121 boundary variables and the remaining 142 are considered explicit variables. On the other hand, there are 138 measured data from the process, so, 121 are assigned to boundary variables and the remaining 17 are redundant (explicit variables but with a measured data available).

4. DATA RECONCILIATION PROBLEM

The data reconciliation problem can be formulated as to compute the decision variables $F_{i,dec}$, $X_{i,dec}$ and $MW^{imp}_{i,dec}$ (flows, purities and molecular weights of impurities that minimize the function J (3) given by the sum of the squares of the deviations between the (compensated) measured data ($F_{i,med}, X_{i,med}, F^{red}_{i,med}, X^{red}_{i,med}$) and the calculated variables ($F_{i,dec}, X_{i,dec}, F^{red}_{i,exp}, X^{red}_{i,exp}$), while satisfying the nonlinear model (4), and the ranges on the explicit and decision variables (5), (6).

$$\begin{aligned}\min_{\{F_{i,dec},X_{i,dec},MW^{imp}_{i,dec}\}} J = &\sum_{i=1}^{89}\frac{w_i}{\sigma_i^2}\left(F_{i,dec}/Fc_i - F_{i,med}\right)^2 \\ &+\sum_{i=1}^{16}\frac{w_i}{\sigma_i^2}\left(X_{i,dec} - X_{i,med}\right)^2 \\ &+\sum_{i=1}^{15}\frac{w_i}{\sigma_i^2}\left(F^{red}_{i,exp}/Fc^{red}_i - F^{red}_{i,med}\right)^2 \\ &+\sum_{i=1}^{2}\frac{w_i}{\sigma_i^2}\left(X^{red}_{i,\exp} - X^{red}_{i,med}\right)^2\end{aligned} \tag{3}$$

Subject to:

$$\begin{aligned} F_{i,exp} &= g(F_{j,dec}, X_{k,dec}, MW^{imp}_{l,dec}) && i = 1,...,33 \\ F^{red}_{i,exp} &= g(F_{j,dec}, X_{k,dec}, MW^{imp}_{l,dec}) && i = 1,...,15 \\ X_{i,exp} &= g(F_{j,dec}, X_{k,dec}, MW^{imp}_{l,dec}) && i = 1,...,24 \\ X^{red}_{i,exp} &= g(F_{j,dec}, X_{k,dec}, MW^{imp}_{l,dec}) && i = 1,...,2 \\ MW_{i,exp} &= g(F_{j,dec}, X_{k,dec}, MW^{imp}_{l,dec}) && i = 1,...,42 \\ MW^{imp}_{i,\exp} &= g(F_{j,dec}, X_{k,dec}, MW^{imp}_{l,dec}) && i = 1,...,26 \end{aligned} \tag{4}$$

$$\begin{aligned} 0 \le F_{i,exp} &\le F_{max} && i = 1,...,33 \\ F^{red}_{i,\min} \le F^{red}_{i,exp} &\le F^{red}_{i,\max} && i = 1,...,15 \\ 0 \le X_{i,exp} &\le 100 && i = 1,...,24 \\ X^{red}_{i,\min} \le X^{red}_{i,exp} &\le X^{red}_{i,\max} && i = 1,...,2 \end{aligned} \tag{5}$$

$$\begin{aligned} F_{i,\min} \le F_{i,dec} &\le F_{i,\max} && i = 1,...,89 \\ X_{i,med} \le X_{i,dec} &\le X_{i,med} && i = 1,...,16 \\ MW^{imp}_{i,\min} \le MW^{imp}_{i,dec} &\le MW^{imp}_{i,\max} && i = 1,...,16 \end{aligned} \tag{6}$$

The stationary model of the hydrogen network is represented by (4), where explicit variables $F_{i,exp}$, $X_{i,exp}$, $MW_{i,\exp}$, $MW^{imp}_{i,exp}$, $F^{red}_{i,exp}$ and $X^{red}_{i,exp}$ are calculated solving the model $g(\bullet)$ with values of the boundary variables $F_{i,dec}$, $X_{i,dec}$ and $MW^{imp}_{i,dec}$ respectively.

The lower and upper limits of the decision variables associated to flows and purities come from the range of their corresponding instruments. But the limits for molecular weights of impurities are obtained through historical data of laboratory analysis in the associated streams, because there are not measured online. Finally, all terms in the cost function (3) have been normalized by means of the variance (σ_i^2) of data measured and can also be weighted by w_i (from 0 to 1) which indicates the level of importance of the corresponding instrument. The problem is a NLP (nonlinear programming) one that consists of 121 decision variables, 142 nonlinear equations (network model), 148 nonlinear constraints (74 lower limits and 74 upper limits of explicit variables) and 121 linear inequalities (lower and upper limits of decision variables).

There is another issue associated to flow measurements which must be taken into account: Most flowmeters in the refinery are orifice plates and they provide volumetric flows at standard conditions considering a specific pressure, temperature and molecular weight of design. However, these values change during the operation, being necessary compensate the corresponding measured flow. The compensated flow is given by,

$$F_{i,compensate} = F_{i,med} Fc_i \Rightarrow F_{i,med} = F_{i,compensate} / Fc_i \tag{7}$$

where $F_{i,med}$ is the measured flow and Fc_i is the factor of compensation defined for each orifice plate,

$$Fc_i = \sqrt{\frac{T_{i,dis}+273}{(P_{i,dis}+1)MW_{i,dis}}}\sqrt{\frac{(P_{i,op}+1)MW_{i,op}}{T_{i,op}+273}} \tag{8}$$

$P_{i,dis}$, $T_{i,dis}$ and $MW_{i,dis}$ are the design values for pressure, temperature and molecular weight of the stream (in kg/cm^2, ºC and g/mol respectively) and "op" are the operating values. Then, pressures and temperatures are also measured data from the process, and the molecular weight of every stream is a variable of the model, which is calculated through the model (4) or equation (2) in the small example.

So, the compensation factor Fc_i is a function of the hydrogen purity and molecular weight of each stream and indirectly a function of molecular weight of impurities for every stream and has been included in cost function (3). In this way, the reconciliation of mass and volume is made simultaneously in a rigorous manner.

5. OPTIMAL MANAGEMENT PROBLEM

According to the policy depicted in Fig.2, the main goal in this step is to distribute the hydrogen in the network and recirculate most of the excess of hydrogen from consumer units into the low purity manifold (C-BP), minimizing the hydrogen production from units H3 and H4 and all flows to the fuel gas manifold, while satisfying predefined hydrocarbon production targets, actual topological restrictions (10) as well as the exact demand in flow and purity of the hydrogen makeup flowing from the different sources to each consumer unit. The cost function *Jc* is shown bellow and the 11 flows to be minimized are shown in Table 2.

$$\min_{\{F_{i,dec},Pu_{i,dec}\}} Jc = \sum_{i=1}^{11} w_i F_i \tag{9}$$

Subject to:

$$\begin{aligned} F_{i,exp} &= g(F_{j,dec}, X_{k,dec}) \qquad i = 1,...,33 \\ F^{red}_{i,exp} &= g(F_{j,dec}, X_{k,dec}) \qquad i = 1,...,15 \\ X_{i,exp} &= g(F_{j,dec}, X_{k,dec}) \qquad i = 1,...,24 \\ X^{red}_{i,exp} &= g(F_{j,dec}, X_{k,dec}) \qquad i = 1,...,2 \end{aligned} \tag{10}$$

$$\begin{aligned} F_{i,\min} \le F_{i,exp} \le F_{i,max} \qquad i = 1,...,33 \\ F^{red}_{i,\min} \le F^{red}_{i,exp} \le F^{red}_{i,\max} \qquad i = 1,...,15 \\ X_{i,\min} \le X_{i,exp} \le X_{i.\max} \qquad i = 1,...,24 \\ X^{red}_{i,\min} \le X^{red}_{i,exp} \le X^{red}_{i,\max} \qquad i = 1,...,2 \end{aligned} \tag{11}$$

$$\begin{aligned} F_{i,\min} \le F_{i,dec} \le F_{i,\max} \quad i = 1,...,89 \\ X_{i,\min} \le X_{i,dec} \le X_{i,\max} \quad i = 1,...,16 \end{aligned} \tag{12}$$

This problem assumes that the dynamics of the network is faster than the one of the production and consumer ones, distributing in a better way the hydrogen available in the refinery. This is possible because in several units the excess hydrogen in the reactions can be sent to fuel gas manifold or recirculated to the low purity manifold (C-BP). Moreover, medium purity manifold (C-H3), low purity manifold (C-BP), manifold from unit N1 to G1 (C-N1G1) and manifold from unit N2 to G2 (C-N2G2) can send hydrogen to fuel gas if there is an overpressure, that is, if the hydrogen in these manifolds is not consumed/used in other units. Table 1. lists the decision variables of the problem: hydrogen production flow in units H4 and H3 (H4.F and H3.F) and all flows to fuel gas manifold which we want to minimize.

Notice that the model of the hydrogen network represented in (10), and used to solve the optimal management problem only includes flows and purities. The molecular weight of every stream, and the molecular weight of impurities are considered constant because all flows measured come from the solution of reconciliation problem previously solved. Equations (11) and (12) are the lower and upper limits of all flows and purities. In many cases these upper and lower limits are equal, forcing to maintain the exact flow and purity of hydrogen makeup in each consumer unit and forcing to maintain the exact excess of hydrogen and its purity from units as the current ones, letting unmodified in this way the internal operation of the HDS. For example, in unit G2 the decision variables are the inflow from manifold C-H4 (C-H4_G2.F), the inflow from manifold C-H3 (C-H3_G2.F) and the inflow from manifold C-N2G2 (C-N2G2_G2.F), imposing the constraints on the total inflow and purity to the unit. Others decision variables are the outflow from G2 to fuel gas (G2_FG.F) and from G2 to low purity manifold (G2_C-BP.F) being their sum fixed by the operation of the unit.

6. RESULTS AND DISCUSSION

The approach presented before has been tested with sets of real operation data of the refinery. Here we present some of them in a certain normalized scale. They corresponds to the average of two hours of operation and the corresponding standard deviation of the measured variables in this period. First, the data reconciliation methodology has been applied and then, the optimal management problem has been solved with all data reconciled. In both cases, the CPU time necessary to solve the optimization problem is lower than 3 minutes in a Intel Corel Duo with 2.13 GHz. Notice that all flows presented here have been scaled between 0 and 100 Nm^3/h and the purities of H_2 are in percentage (%).

6.1 Data reconciliation

The optimization problem (3) has been solved with a set of weights w_i equal to 1 for all terms in cost function (3), that is, we suppose that all measured data has the same accuracy. Fig. 3 shows the standard deviation times between the reconciled data (the solution) and measured data. Notice that, measured flows are not compensated but the solution flows are compensated in pressure, temperature and molecular weight, so, to compare both quantities the flows have been de-compensated. The solution of NLP problem provide a coherent close balance of hydrogen in all hydrogen network, 108 reconciled measures have a difference lower than 4 sigmas. These differences can be due to a bad flowmeter calibration or other causes. In order to eliminate its effect, the data reconciliation is repeated, this time with a weight w_i equal to zero in the potentially faulty variables, and the new reconciled data are used in the following step, while an order is given to recalibrate the defective instruments.

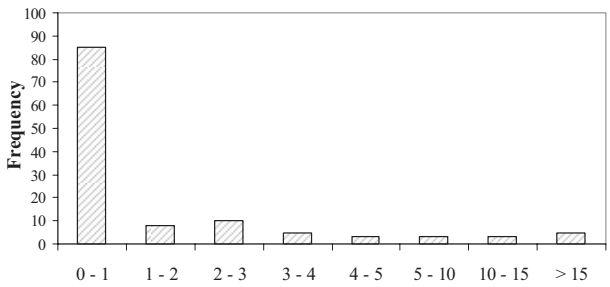


Fig. 3. Difference between reconciliated and measured data in number of standard deviations.

6.2 Optimal management of hydrogen network

The optimization problem (9) has been solved with the set of reconciled data obtained before. Results of the optimization are given in Table 1 besides their initial values(reconciled data): the cost function *Jc* and individual values of each term in this cost function. The cost function has been reduced from 100 Nm^3/h to 81.76 Nm^3/h. The optimal solution obtained reduces the total flow produced in units H4 and H3, from 73.22 Nm^3/h to 64.10 Nm^3/h and decrease the flow sent to fuel gas, from 24.98 Nm^3/h to 16.76 Nm^3/h, that is, the results show the possibility to reuse the hydrogen available in the refinery in a better way: i) without modifying the operation of each consumer unit ii) without increasing the production of hydrogen and iii) without increasing the purity of hydrogen produced. The minimum and maximum production of hydrogen allowed in unit H4 is 27.84 Nm^3/h and 62.77 Nm^3/h respectively and 15.57 Nm^3/h and 34.53 Nm^3/h for unit H3. Notice that the flow production of unit H3 (H3.F) has been reduced to the minimum production allowed, 15.57 Nm^3/h.

Table 1. Solution of the hydrogen optimal management problem

Flows (F_i)	Units	Data reconciled	Optimal solution
H4.F	Nm^3/h	50.76	48.53
H3.F	Nm^3/h	22.46	15.57
G1_FG.F	Nm^3/h	0.00	0.00
G2_FG.F	Nm^3/h	2.42	10.11
P1_FG.F	Nm^3/h	2.64	0.00
P2_FG.F	Nm^3/h	5.75	0.00
C-BP_FG.F	Nm^3/h	0.26	0.29
C-BP_FG2.F	Nm^3/h	12.72	6.36
C-N1G1_S1.F	Nm^3/h	0.08	0.00
C-N2G2_S2.F	Nm^3/h	1.72	0.89
C-H3_FG.F	Nm^3/h	1.20	0.00
Sum of all flows (*Jc*)	Nm^3/h	100.00	81.76
Economical cost (*Je*)	€/h	100.00	66.50

It is interesting to evaluate the economical cost of this solution *Je*. To do this, we are going to use the economical criteria used in the refinery. That is, the economical cost *Je* is the total hydrogen sent to fuel gas manifold times the total cost of hydrogen production minus the price of hydrogen as fuel,

$$Je = F_{H2-FG}\left(Cost_{ProductionH2} - Price_{combustibleH2}\right) \quad (13)$$

On the other hand, the total cost of hydrogen production is calculated by means,

$$Cost_{ProductionH2} = \frac{Cost_{H2inH4} H4.F + Cost_{H2inH3} H3.F}{H4.F + H3.F} \quad (14)$$

where H4.F and H3.F are the flow production in H4 and H3 respectively, the cost of hydrogen production in unit H4 is $Cost_{H2inH4}$ = 77.0€/KNm^3, in unit H3 is $Cost_{H2inH3}$ = 88.1€/KNm^3 and the price of hydrogen used as combustible in the fuel gas manifold is $Price_{combustibleH2}$=6.55 €/KNm^3. These are scaled values and they are related to pure hydrogen (100 % of purity). Table 1. shows the economical cost before hydrogen optimal management *Je* = 100.00 €/h and for the optimal solution *Je* = 66.5 €/h, so, it is possible a economical reduction of 33.5 %. Of course, this solution is not directly applicable to the refinery mainly due to the pressure constraints in hydrogen network. At present, further research is conducted to include the dynamical constraints imposed by the lower network control layer on the network optimization.

6. CONCLUSIONS

An approach has been presented to optimally manage complex hydrogen networks of refinery operations. The data reconciliation and optimal hydrogen distribution steps have been described with more detail using a NLP based optimization. The proposed method is able to systematically reduce utility cost by increasing hydrogen recovery in consumer units and reducing production cost in the alternative hydrogen suppliers. This paper is mainly focused on the treatment of hydrogen mass balances. Future work is aiming at extending the model to actual compression costs and other operational constraints as well as the use of alternative separation units (membranes) to recycle higher-purity off-gas to consumer units. In particular, including membranes in the model, convert it in a hybrid process, because membranes are formed by discrete package which can be turn on or turn off. Other improvements are related to the gross errors detection must be added to the DSS to enhance the quality and coherence of the reconciled data as well as better detect instrumentation malfunctions in the refinery.

ACKNOWLEDGMENTS

The valuable collaboration of Repsol-YPF and the Petronor refinery is gratefully acknowledged. The authors are also thankful for financial support received from project DPI2006-13593 of the Spanish CICYT and project GR. 085/2008 of JCyL.

REFERENCES

Alves, J. (1999). *Analysis and design of refinery hydrogen systems*, Ph.D. Thesis, UMIST.

Cronkwright, M., Frey, G., Brown, S. and Gulati, H. (2007) Refinery-wide data reconciliation case-study: operational decision support based on reconciled data. *NPRA 2007 O&A and Technology Forum*, 8-9.

Fonseca, A., Sá, V., Bento, H., Tavares, M.L.C., Pinto, G. and Gomes, A.C.N. (2008). Hydrogen distribution network optimization: a refinery case study. *Journal of Cleaner Production*, 16, 1755-1763.

Hallale, N. and Liub, F. (2001). Refinery hydrogen management for clean fuels production. *Advances in Environmental Research*, 6, 81-98.

Towler, G.P., Mann, R., Serriere, A.J-L, and Gabaude, C.M.D. (1996). Refinery hydrogen management: Cost analysis of chemically-integrated facilities. *Ind. Eng. Chem. Res.*, 35 (7), 2378-2388.

Performance Monitoring in Supermarket Refrigeration Systems - Synchronization of Refrigerated Display Cases

Liang Chen * **Torben Green** ** **Lars F. S. Larsen** **
Rafael Wisniewski * **Roozbeh Izadi-Zamanabadi** **

* *Department of Electronic Systems, Automation and Control, Aalborg University, Fredrik Bajers Vej 7C, DK-9220 Aalborg, Denmark (email: lchen.dhu@gmail.com, raf@es.aau.dk).*
** *Danfoss A/S, Refrigeration and Air conditioning, DK-6430 Nordborg, Denmark (email: Torben.Green@danfoss.com, Lars.Larsen@danfoss.com, Roozbeh@danfoss.com).*

Abstract: The overall task of a supermarket refrigeration system is to maintain the quality of the foodstuff. This is done by making use of a refrigeration cycle in which a refrigerant transport heat from the refrigerated display-cases to the outdoor surroundings. Typically the system is equipped with a decentralized control system neglecting interactions between subsystems. Though these interactions are minor they from time to time lead to a synchronization of the operation of the display-cases which causes an inferior control performance and increased energy consumption. In this paper we will analyze the synchronization using bifurcation theory and show that the system has a chaos-like behavior when it is not synchronized. Therefore, it is a good choice to de-synchronize the system by making the system chaotic. The positive maximal Lyapunov exponent is usually taken as an indication that the system is chaotic, it is used in this paper as a measure of performance for the tendency of the system to synchronize.

Keywords: Refrigeration system, Chaotic behavior, Hybrid system, Performance monitoring, Complex system.

1. INTRODUCTION

A supermarket refrigeration system consists of a central compressor rack that maintains the required flow of refrigerant to the refrigerated display cases located in the supermarket sales area. Each display case has an inlet valve for refrigerant that needs to be opened and closed such that the air temperature in the display case is kept within tight bounds to ensure a high quality of the goods. For many years, the control of supermarket refrigeration systems has been based on distributed control systems, which are flexible and simple. For example, each display case used to be equipped with an independent hysteresis controller that regulates the air temperature in the display case by manipulating the inlet valve. The major drawback, however, is that the control loops are vulnerable to self-inflicted disturbances caused by the interaction between the distributed control loops. In particular, practice and simulations show that the distributed hysteresis controllers have the tendency to synchronize [Larsen (2006)], meaning that the opening and closing actions of the valves coincide. Consequently, the compressor periodically has to work hard to keep up the required flow of refrigerant, which results in low efficiency, inferior control performance and a high wear on the compressor. The control problem is significantly complicated by the fact that many of the control inputs are restricted to discrete values, such as the opening/closing of the inlet valves and the stepwise control of the compressor. Furthermore, the system features switched dynamics turning the supermarket refrigeration system into a hybrid system.

The intense focus on limiting energy consumption and the global environmental awareness calls for energy efficient solutions. By monitoring the performance of the refrigeration system the "goodness" of the operation can be measured and early warnings about undesired behaviors can be given such that the control system can accommodate these and continuously optimize the system performance. In this paper we will focus on the monitoring of the synchronization phenomenon. By analyzing the system behavior using bifurcation and chaos theory [Crawford (1991), Devaney (2003)] it can be shown that the system has a chaos-like behavior. Bifurcation and chaos theory is most commonly applied to the mathematical study of dynamical systems to investigate dramatic changes in the qualitative or topological structure of a system. It can be dated back to 1975 when the first mathematical definition of 'chaos' was given [Li and Yorke (1975)]. The synchronization in the paper can be interpreted as a low order limit cycle [Wisniewski and Larsen (2008)]. It will be shown how the system jumps between low order and high order limit cycles varying the hysteresis bounds of the temperature controller. If the system converges towards a

low order limit cycle it can be seen as an indication of a risk that the system may synchronize. Therefore, it is a good choice to de-synchronize the system by making the system chaotic. The positive maximal Lyapunov exponent is usually taken as an indication that the system is chaotic. A huge number of references are available on calculation of the Lyapunov exponents [Müller (1995); Benettin et al. (1980); Wolf et al. (1985)]. We suggest using the maximal Lyapunov exponent as a measure of performance for the tendency of the system to synchronize.

2. SYSTEM DESCRIPTION

The overall task for a supermarket refrigeration system is to maintain the quality of the stored foodstuff. The goods are usually stored in open display cases in the sales area of the supermarket. The working principle of the supermarket system is a refrigeration cycle which utilizes a refrigerant to transport heat from the display cases to the outdoor surroundings.

A simplified supermarket refrigeration circuit is shown in Fig. 1. The compressors and the display cases are in the majority of supermarket refrigeration systems connected in parallel. The compressors supply the flow by compressing the low pressure refrigeration which is drained from the suction manifold. The refrigerant then passes through the condenser and into the liquid manifold. Each display case has an expansion valve which is connected to the liquid manifold where from the refrigerant then flows through the expansion valve and into the evaporator of the display case. In the evaporator the refrigerant absorbs heat from the stored goods and thereby changes phase. The vaporized refrigerant flows into the suction manifold, thus closing the refrigerant cycle. The typical layout of a

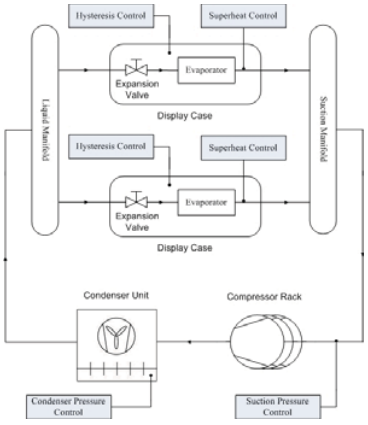


Fig. 1. Simplified supermarket refrigeration layout

display case can be seen in Fig. 2. The refrigerant is fed into the evaporator at the bottom of the display case and as the air is passed over the surface of the evaporator heat is absorbed which render a vaporization of the refrigerant. The resulting air flow creates an air curtain of cold air over the stored good. The air takes up the heat flow $\dot{Q}_{\text{goods}-\text{air}}$ from the stored goods and as a side effect the heat flow $\dot{Q}_{\text{load}}$ from the surroundings. The temperature of the air, T_{air}, is measured by a sensor mounted in the inlet air stream to the evaporator the goods to provide an indirect measure for the temperature of the goods.

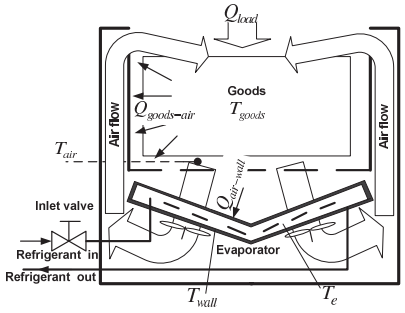


Fig. 2. Cross sectional view of a refrigerated display case.

2.1 Traditional Control

The typical controller structure for a supermarket refrigeration system is decentralized. Each of the display cases is fitted with an air temperature controller and a superheat controller, which ensures the desired filling of the evaporator. The compressor rack controls the suction pressure controller and the condenser fans control the condensing pressure. The only controllers considered in this paper is the suction pressure controller and the temperature controllers in the display cases.

The temperature in the display cases is controlled by a hysteresis controller that opens and closes the expansion valve, i.e. the valve opens when T_{air} reaches a predefined upper temperature limit and stay open until T_{air} decreases to the lower temperature limit and the valve closes again.

The suction pressure is controlled by switching compressors in the compressor rack on or off. A dead band around the reference is introduced to avoid excessive switching of the compressors. If the pressure exceeds the upper bound of the dead band one or more compressors are switched on. If the pressures drops below the lower bound of the dead band a compressor is switched off. This control strategy prevents moderate changes in the suction pressure from initiating compressor switching.

In a common supermarket many of the display cases will be alike and in addition be working under the same conditions. Thus, the switching frequency , for each of the expansion valves for the different display cases, will be close to each other. The display cases have a tendency to synchronize because there individual dynamics are coupled through the suction pressure. Synchronization of the display cases lead to periodic high and low amount of vaporized refrigerant flow into the suction manifold. Hence, large fluctuations in the suction pressure will be a consequence which then leads to higher switch frequency of the compressors and therefore excessive wear on the compressors. The result from synchronizing display cases can be seen in Fig. 3

3. MODEL OF THE REFRIGERATION SYSTEM

The model for the supermarket refrigeration system is composed of a number of sub-models which each represent a component in the refrigeration system. That is, individual models are made for the display cases, the suction manifold, the compressor rack, and the condensing unit. Because the emphasis of the paper is to examine

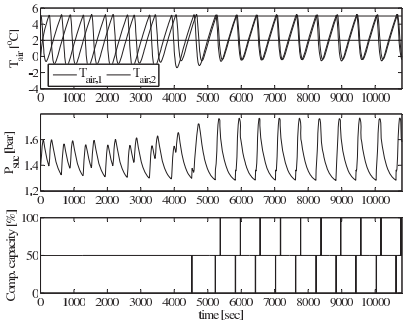


Fig. 3. The effect of synchronization

the synchronization phenomenon the modeling will be concentrated on the display cases and the suction manifold such that only the dynamics relevant for the control of the hysteresis control and the compressors are captured. The dynamic of the compressors are usually much faster than the dynamics of the rest of the refrigeration system. Thus, the modeling of the compressors dynamic is neglected.

The mathematical model presented in this section is a summary of the supermarket refrigeration model developed in [Larsen et al. (2007)]. The suction pressure P_{suc}, comprises the common state for the combined models. Each display case, $i = 1, \ldots, N$ where N is the number of display cases in the system, is described by four states. That is, the temperature of the goods $T_{\text{goods,i}}$, the air temperature $T_{\text{air,i}}$, the temperature of the evaporator wall $T_{\text{wall,i}}$, and the mass of the refrigerant in the evaporator $M_{r,i}$. The input is the volumetric flow generated by the compressors $\dot{V}_{\text{comp}}$, and the binary state of the ith inlet valve δ_i (closed or opened, $\delta_i \in \{0, 1\}$). The systems are affected by the heat load from the surroundings of the display case $\dot{Q}_{\text{load}}$.

$$\frac{\mathrm{d}T_{\text{goods,i}}}{\mathrm{d}t} = -\frac{\dot{Q}_{\text{goods-air,i}}(\cdot)}{M_{\text{goods,i}}\ C_{\text{p,goods,i}}} \tag{1}$$

$$\frac{\mathrm{d}T_{\text{wall,i}}}{\mathrm{d}t} = \frac{\dot{Q}_{\text{air-wall,i}}(\cdot) - \dot{Q}_{\text{e,i}}(\cdot)}{M_{\text{wall,i}} C_{\text{p,wall,i}}} \tag{2}$$

$$\frac{\mathrm{d}T_{\text{air,i}}}{\mathrm{d}t} = \frac{\dot{Q}_{\text{goods-air,i}}(\cdot) + \dot{Q}_{\text{load,i}}(\cdot) - \dot{Q}_{\text{air-wall,i}}(\cdot)}{M_{\text{air}} C_{\text{p,air,i}}} \tag{3}$$

$$\frac{\mathrm{d}M_{r,i}}{\mathrm{d}t} = \begin{cases} \dfrac{M_{\text{r,max,i}} - M_{\text{r,i}}}{\tau_{\text{fill,i}}} & \text{if } \delta_i = 1 \\ -\dfrac{\dot{Q}_{\text{e,i}}(\cdot)}{\Delta h_{\text{lg}}(P_{\text{suc}})} & \text{if } \delta_i = 0 \text{ and } M_{\text{r,i}} \geq 0 \\ 0 & \text{if } \delta_i = 0 \text{ and } M_{\text{r,i}} = 0 \end{cases} \tag{4}$$

$$\frac{\mathrm{d}P_{\text{suc}}}{\mathrm{d}t} = \frac{\dot{m}_{\text{in-suc}}(\cdot) + \dot{m}_{\text{r,const}} - \dot{V}_{\text{comp}}\ \rho_{\text{suc}}(P_{\text{suc}})}{V_{\text{suc}}\ \nabla\rho_{\text{suc}}(P_{\text{suc}})} \tag{5}$$

The enthalpy difference across the two-phase region of the evaporator is denoted by Δh_{lg}, the density of the refrigerant is denoted by ρ_{suc}, and $\nabla\rho_{\text{suc}}$ denotes the pressure derivative of the refrigerant density. T_e is the evaporation temperature for the refrigerant. The subscripts for the mass M and the heat capacity C_p denotes the media. The heat flow is denoted by $\dot{Q}$ where the subscript indicates the media between which the thermal energy is exchanged. In addition, the varies heat flows are defined by the following functions:

$$\dot{Q}_{\text{goods-air,i}}(T_{\text{goods,i}}, T_{\text{air,i}}) = UA_{\text{goods-air,i}} \cdot (T_{\text{goods,i}} - T_{\text{air,i}}) \tag{6}$$

$$\dot{Q}_{\text{air-wall,i}}(T_{\text{air,i}}, T_{\text{wall,i}}) = UA_{\text{air-wall,i}} \cdot (T_{\text{air,i}} - T_{\text{wall,i}}) \tag{7}$$

$$\dot{Q}_{e,i}(M_{r,i}, T_{\text{wall,i}}, P_{\text{suc}}) = UA_{\text{wall-ref,i}}(M_{r,i}) \cdot (T_{\text{wall,i}} - T_e(P_{\text{suc}})) \tag{8}$$

$$UA_{\text{wall-ref,i}}(M_{r,i}) = UA_{\text{wall-ref,max,i}} \cdot \frac{M_{\text{r,i}}}{M_{\text{r,max,i}}} \tag{9}$$

The overall heat transfer coefficient is denoted by UA and the subscript denotes the media from which the heat is transferred. In addition the mass flow rate in the suction manifold is given by:

$$\dot{m}_{\text{in-suc}}(M_{\text{r,i}}, T_{\text{wall,i}}, P_{\text{suc}}) = \sum_{i=1}^{N} \frac{\dot{Q}_{\text{e,i}}(\cdot)}{\Delta h_{\text{lg}}(P_{\text{suc}})} \tag{10}$$

The functions Δh_{lg}, ρ_{suc}, and T_e are refrigerant specific. Detailed description of these functions are given in [Larsen et al. (2007)]. In (4) it can be seen that the system have a hybrid nature due to the the discrete input which represents the opening and closing of the expansion valves.

3.1 Simplified model

In order to obtain a model that is suitable for analyzing the synchronization phenomenon the equation system (1) through (5) are further simplified to a second order affine switched system.

The simplification of the model is based on the following assumptions:

- The heat capacity of the goods is large, thus the temperature of the goods in a display case is constant and equal T_{g0}.
- The heat capacity of the air is small.
- The evaporator is instantly filled (emptied) when the inlet valve is opened (closed).
- The mass flow out of the display case when the valve is open is constant and equal $\dot{m}_0$.
- The evaporation temperature T_{e} and the density ρ_{suc} of the refrigerant in the suction manifold are affine functions of suction pressure P_{suc},

$$T_{\text{e}} = a_T P_{\text{suc}} + b_T \text{ and } \rho_{\text{suc}} = a_\rho P_{\text{suc}} + b_\rho$$

- The gradient $\nabla\rho_{\text{suc}}(P_{\text{suc}}) \equiv \nabla\rho_{\text{suc0}}(P_{\text{suc0}})$ is constant.
- The compressor delivers a constant volume flow $\dot{V}_{\text{comp}}$.
- The heat load $\dot{Q}_{\text{load}}$ on the display cases is constant.

Based on these assumptions the dynamic of the air temperature $T_{\text{air,i}}$ in the ith display case can be formulated as follows:

$$\frac{\mathrm{d}T_{\mathrm{air,i}}}{\mathrm{d}t} = \frac{\dot{Q}_{\mathrm{goods-air,i}} + \dot{Q}_{\mathrm{load,i}} - \delta_i \dot{Q}_{\mathrm{e,max,i}}}{\left(1 + \frac{UA_{\mathrm{goods-air,i}}}{UA_{\mathrm{air-wall,i}}}\right) M_{\mathrm{wall,i}} Cp_{\mathrm{wall,i}}} \quad \text{with} \tag{11}$$

$$T_{\mathrm{wall,i}} = T_{\mathrm{air,i}} - \frac{\dot{Q}_{\mathrm{goods-air,i}} + \dot{Q}_{\mathrm{load,i}}}{UA_{\mathrm{air-wall,i}}}, \tag{12}$$

$$\dot{Q}_{\mathrm{goods-air,i}} = UA_{\mathrm{goods-air,i}}(T_{\mathrm{g0,i}} - T_{\mathrm{air,i}}), \tag{13}$$

$$\dot{Q}_{\mathrm{e,max,i}} = UA_{\mathrm{wall-ref,max,i}}(T_{\mathrm{wall,i}} - a_T P_{\mathrm{suc}} - b_T), \tag{14}$$

The suction manifold dynamics is governed by the expression

$$\frac{\mathrm{d}P_{\mathrm{suc}}}{\mathrm{d}t} = \frac{\sum_{i=1}^{N} \delta_i \dot{m}_{0,\mathrm{i}} + \dot{m}_{\mathrm{r,const}} - \dot{V}_{\mathrm{comp}}(a_\rho P_{\mathrm{suc}} + b_\rho)}{V_{\mathrm{suc}} \cdot \nabla\rho_{\mathrm{suc0}}}. \tag{15}$$

Thus, the non-linear hybrid system has been reduced to a two order (for each display case) affine system with discrete inputs. For a refrigeration system with two display cases, the system states of the simplified model are $T_{\mathrm{air,i}}(i = 1, 2)$ and P_{suc}. The discrete inputs are $\delta_i \in \{0, 1\}$, which indicate if the valves are closed or open. The input δ is controlled by a hysteresis controller which changes the value of δ in the following way:

$$\delta_i(k+1) = \begin{cases} 1 & \text{if } T_{air,i} \geq \overline{T_{\mathrm{air,i}}} \\ 0 & \text{if } T_{air,i} \leq \underline{T_{\mathrm{air,i}}} \\ \delta_i(k) & \text{if } \underline{T_{\mathrm{air,i}}} < T_{air,i} < \overline{T_{\mathrm{air,i}}}, \end{cases} \tag{16}$$

where k denotes the time index, $\overline{T_{air,i}}$ is the upper bound the air temperature and $\underline{T_{air,i}}$ is the lower bound .

4. DYNAMICAL ANALYSIS

In this section, we will analyze dynamics of the simplified refrigeration model through bifurcation and chaos theory. The theory is most commonly applied to the mathematical study of dynamical systems. The aim of the theory is to investigate dramatic changes in the qualitative or topological structure of a system by changing smoothly a system parameter. For the refrigeration system we will analyze the system behavior by changing smoothly the lower bound in one of the display cases. The resulting behavior will be depicted in a so-called bifurcation diagram, from which the synchronization phenomenon will be studied, thereafter a measure will be developed to evaluate the tendency of synchronization. All simulation results in the section are based on the following parameter settings:

4.1 Phase plots w.r.t $\underline{T_{\mathrm{air,2}}}$

We shall study bifurcation, i.e. the influence of changes of system parameters on the system behavior at large. There are two system parameters, i.e. the upper bound and the lower bound of the temperature in the air temperature control of the display case. Here, we examine the lower bound of the second display case $\underline{T_{\mathrm{air,2}}}$. This will provide an example which will help understand how the system behaves with the varying parameter.

Fig. 4 shows some typical phase plots of the system states $T_{\mathrm{air,1}}$ and $T_{\mathrm{air,2}}$. When the parameter $\underline{T_{\mathrm{air,2}}} = 0$, the limiting behavior of the system switches between the two points $(0, 0)$ and $(5, 5)$ within an accepted tolerance ($1E-6$ in the paper). Here, we call it as a 2-periodic limit cycle. The period of a limit cycle is defined by the sum of a number of switching points on the boundary $\partial\square = \partial([\underline{T_{\mathrm{air,1}}}, \overline{T_{\mathrm{air,1}}}] \times [\underline{T_{\mathrm{air,2}}}, \overline{T_{\mathrm{air,2}}}])$. The phase plot of the 2-periodic limit cycle corresponds to the synchronization phenomenon mentioned in the above section, where the two states $T_{\mathrm{air,1}}$ and $T_{\mathrm{air,2}}$ agree all the time. When the parameter $\underline{T_{\mathrm{air,2}}}$ increases slightly to the value of 0.15, a 4-periodic limit cycle appears, which is similar to the synchronization but with a bigger difference of the two states; we will call it the quasi-synchronization. When $\underline{T_{\mathrm{air,2}}} = 0.2$, another topology of 4-periodic limit cycle appears in the phase plot, which is totally different from the state agreement in the synchronization. If we continue increasing $\underline{T_{\mathrm{air,2}}}$ to the value of 0.3, we will find that the system tends to a high-periodic limit cycle with many switching points in the boundary of $\partial\square$. It looks like chaos, the common phenomenon in the nonlinear system [Devaney (2003)]. It is far away from the synchronization.

Table 1. Parameters for a simplified supermarket refrigeration system

Display cases					
$UA_{\mathrm{wall-ref,max}}$	500	$\frac{J}{s \cdot K}$	T_{g0}	3.0	^0C
$UA_{\mathrm{goods-air}}$	300	$\frac{J}{s \cdot K}$	$\dot{m}_0$	1.0	kg/s
$UA_{\mathrm{air-wall}}$	500	$\frac{J}{s \cdot K}$	$\dot{Q}_{\mathrm{load}}$	3000	J/s
$\dot{m}_{\mathrm{r,const}}$	0.2	$\frac{kg}{s}$	M_{wall}	260	kg
$\nabla\rho_{\mathrm{suc0}}$	4.6	$\frac{kg}{m^3 bar}$	$C_{\mathrm{p,wall}}$	385	$\frac{J}{kg \cdot K}$
The same parameters has been used for all disp.					
Compressor					
$\dot{V}_{\mathrm{comp}}$	0.28	$\frac{m^3}{s}$			
Suction manifold					
V_{suc}	5.00	m^3			
Air temperature control					
$\underline{T_{\mathrm{air,i}}}$	0.00	^0C	$\overline{T_{\mathrm{air,i}}}$	5.00	^0C
i for the disp.					
Coefficients					
$a_T = 16.2072$		$b_T = 41.9095$	$a_\rho = 4.6$		$b_\rho = 0.4$

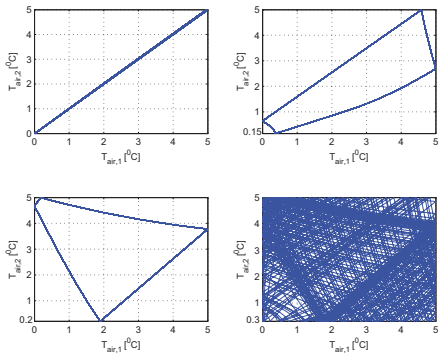


Fig. 4. Typical limiting behaviors with the various values of the parameter $\underline{T_{\mathrm{air,2}}}$. (a) 2-periodic limit cycle (synchronization, $\underline{T_{\mathrm{air,2}}} = 0$), (b) 4-periodic limit cycle ($\underline{T_{\mathrm{air,2}}} = 0.15$), (c) another 4-periodic limit cycle ($\underline{T_{\mathrm{air,2}}} = 0.2$), (d) high-periodic limit cycle ($\underline{T_{\mathrm{air,2}}} = 0.3$).

Synchronization of the display cases leads to large fluctuations in the suction pressure which then result in higher switch frequency of the compressors. It reduces lifetime of the compressors and enlarges energy consumption. Fig. 5 shows the comparison of the suction pressures between the synchronization and the chaos-like situation. We can see that in the chaos-like situation, the fluctuation range of the suction pressure decreases occasionally; even for the part with the same fluctuation amplitude as the synchronization, the pressure jumps so fast that the traditional PI controller in the compressor can fix it. Therefore, we conclude that good control performance can be achieved if the system behaves like chaos.

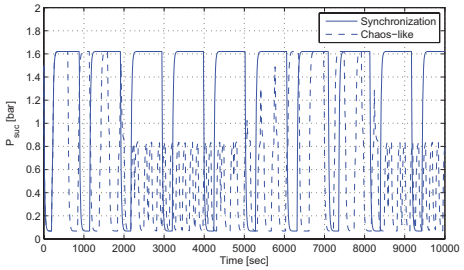


Fig. 5. The suction pressure P_{suc} in the situations of (a) synchronization and (b) chaos-like ($\underline{T_{\text{air},2}} = 0.3$).

4.2 Bifurcation diagram w.r.t $\underline{T_{\text{air},2}}$

To show how the system behaves with the smooth change of the parameter, we usually use the bifurcation diagram. A bifurcation diagram exhibits the possible long-term values (equilibria/fixed points or periodic orbits) of a system as a function of parameters in the system. A bifurcation occurs when a small smooth change made to the values of the bifurcation parameter causes a sudden 'qualitative' or topological change in its behavior. In this paper, the lower bound of the second display case $\underline{T_{\text{air},2}}$ is considered as the bifurcation parameter. Fig. 6 shows the limiting behavior of the state $T_{\text{air},1}$ with respect to the parameter $\underline{T_{\text{air},2}}$. The system exhibits very complicated behaviors. When $\underline{T_{\text{air},2}} = 0$, the system stabilizes at a 2-periodic limit cycle with the phase plot shown in Fig. 4a. As the parameter $\underline{T_{\text{air},2}}$ increases, the 2-periodic limit cycle becomes unstable, and a stable 4-periodic limit cycle appears (the phase plot is like Fig. 4b). The limit cycle retains the similar shape until another stable 4-periodic limit cycle occurs at $\underline{T_{\text{air},2}} = 0.186$ (the phase plot is like Fig. 4c). The shape keeps until $\underline{T_{\text{air},2}} = 0.234$, a higher order oscillation, like chaos, is generated (the phase plot is like Fig. 4d). If the parameter continuously rise, we can see that the system behavior becomes very complex and keeps switching between order and chaos-like oscillations.

The bifurcation diagram demonstrates that the simple refrigeration model is very sensitive to a change in the parameter $\underline{T_{\text{air},2}}$. If we change the parameter slightly, we obtain a totally different topology of the behavior. Therefore, we may ask: is it possible to suppress the synchronization phenomenon we found in the practice of the supermarket refrigeration system by adjusting a little

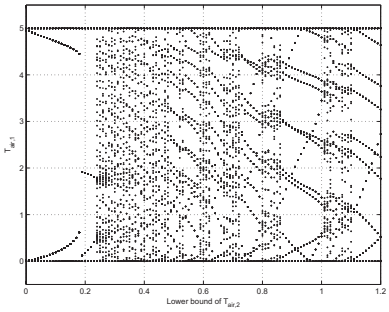


Fig. 6. Bifurcation diagrams of the system state $T_{\text{air},1}$ w.r.t the parameter $\underline{T_{\text{air},2}}$.

bit value of the bounds $\overline{T_{\text{air}}}$ or $\underline{T_{\text{air}}}$? It is obvious for a small 'toy' system with only two display cases that by changing the bound of the temperature, the system can be de-synchronized by making it chaotic. However, for a real plant with many different display cases, it is not obvious how to select the bounds. Hence, we are seeking a method to identify whether the bounds are selected such that the system will synchronize or not. For this purpose we will use the maximal Lyapunov exponent.

4.3 Description of Lyaponov exponent

The Lyapunov exponent characterizes the averaged rate of separation of two close trajectories in the phase space [Oseledec (1968)]. Quantitatively, two trajectories $x_1(t)$ and $x_2(t)$ in phase space with initial separation $\delta x(0)$ diverge

$$|\delta x(t)| \approx e^{\lambda t} |\delta x(0)|, \tag{17}$$

where $\delta x(t) = x_1(t) - x_2(t)$, λ is the Lyapunov exponent. The negative Lyapunov exponent measures the exponential convergence of trajectories, and the positive measures the exponential divergence of trajectories.

There are n Lyapunov exponents in the spectrum of an n-dimensional dynamical system. It is common to just refer to the largest one, i.e. the maximal Lyapunov exponent (MLE), which is defined as follows:

$$\lambda_{\max} = \lim_{t\to\infty} \frac{1}{t} \ln \frac{|\delta x(t)|}{|\delta x(0)|}. \tag{18}$$

The positive MLE is usually taken as an indication that the system is chaotic. It is used in this paper as a measure of performance for the tendency of the system to synchronize, that is, the higher value of the MLE the lower risk for synchronization.

4.4 Algorithm for computation the MLE

Algorithms for computing the Lyapunov exponents of the "smooth" dynamical system are well established [Benettin et al. (1980),Wolf et al. (1985)]. Let us consider the system

$$\dot{x} = f(x(t)), \tag{19}$$

where $f \in C^1$ is a continuously differentiable vector function with the initial condition $x(t_0) = x(0)$. The algorithm is based on the integration of the linearized equation (20) as follows:

$$\delta\dot{x} = J(t)\delta x, \tag{20}$$

where

$$J(t) = \left.\frac{\partial f(x)}{\partial x^T}\right|_{x=x(t)} \quad (21)$$

is the Jacobian matrix of f w.r.t the trajectory under consideration. The MLE $\lambda_{\max}$ is given as the average for some different initial conditions $\delta x(0)$ as Eq. (18).

For the "non-smooth" dynamical system with discontinuities like the refrigeration system in this paper, the above algorithm cannot be directly applied. One way to calculate the MLE is to supplement the transition conditions at the instants of discontinuities into the linearized equations [Müller (1995)]. This model-based method requires exact system information and the calculation complexity greatly increases with the number of the sub-systems. It is not suitable for the refrigeration system especially with many display cases in the supermarket. Besides, to apply this method one needs to know the switching sequence of the sub-systems in advance. It is also impossible in our case. We prefer to take the practical view of calculating the MLE from experimental data. Two well-known papers are [Wolf et al. (1985), S. Sato and Sawada (1987)]. Basic computing steps are given in the following:

(1) Based on the $N-$point time series $x_1, x_2, ..., x_N$, reconstruct the phase space:

$$Y_i = [x_i, x_{i+\tau}, ..., x_{i+(m-1)\tau}] \in R^n \ (i = 1, ..., M),$$

where $M = N - (m-1)\tau$, τ is the reconstruction delay, m is the embedding dimension.

(2) Find the nearest neighbor, $Y_{\hat{i}}$, by searching for the point that minimizes the distance to the particular reference point Y_i, that is, $d_i(0) = \min_{Y_{\hat{i}}} \|Y_i - Y_{\hat{i}}\|$. After j discrete-time steps, the distance $d_i(0)$ goes to $d_i(j)$.

(3) Estimate the averaged rate of distance separation as the MLE:

$$\lambda_{\max} = \frac{1}{j \cdot \Delta t} \cdot \frac{1}{(M-j)} \sum_{i=1}^{M-j} \frac{d_i(j)}{d_i(0)},$$

where Δt is the sampling period of the time series.

5. CONCLUSION

The main focus of this paper was on dynamic analysis of a simple system with two display cases by using bifurcation and chaos theory. Interpreting synchronization as a low periodic limit cycle and by varying the hysteresis bounds of the temperature controller it was shown that the system exhibited a complex chaos-like behavior when it was not synchronized, i.e. it switches between low and high periodic limit cycles. Synchronization of the system has proven to result in an inferior performance due to the resulting large pressure variations. It was, however, indicated in this paper that by de-synchronizing the system (by making it chaotic) it is possible to significantly reduce the pressure variations and hence improve the performance. The positive maximal Lyapunov exponent, usually used as the indication of chaos, was in this paper used as a measure of performance for the tendency of the system to synchronize. These findings for the small "toy" system may seem obvious, however they can easily be scaled to (realistic) large scale systems, where it is harder to distinguish and evaluate "good" and "bad" behavior.

REFERENCES

Benettin, G., Galgani, L., Giorgilli, A., and Strelcyn, J.M. (1980). Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part i: Theory, part ii: Numerical application. *Meccnnica*, 15:9–20, 21–30.

Crawford, J.D. (1991). Introduction to bifurcation theory. *Reviews of Modern Physics, Vol. 63, No. 4*, 991–1037.

Devaney, R.L. (2003). *An Introduction to Chaotic Dynamical Systems*. Westview Press.

Georgiou, I.T. and Schwartz, I.B. (1999). Dynamics of large scale coupled structural/ mechanical systems: A singular perturbation/proper orthogonal decomposition approach. *SIAM J. Appl. Math.*, 59:1178–1207.

Kennedy, M.P., Kolumban, G., and Kis, G. (2000). Chaotic modulation for robust digital communications over multipath channels. *Int. J. Bifur. Chaos*, 10:695–718.

Larsen, L.F.S. (2006). *Model Based Control of Refrigeration Systems*. Ph.D. thesis, Department of Control Engineering - Aalborg University, Aalborg.

Larsen, L., Izadi-Zamanabadi, R., Wisniewski, R., and Sonntag, C. (2007). Supermarket refrigeration systems - a benchmark for the optimal control of hybrid systems. Technical report, Hycon, http://astwww.bci.uni-dortmund.de/hycon4b/wprelated/sr.pdf.

Li, T.Y. and Yorke, J.A. (1975). Period three implies chaos. *Amer. Math. Monthly*, 82:481–485.

Müller, P.C. (1995). Calculation of lyapunov exponents for dynamic systems with discontinuities. *Chaos, Solitons and Fractals*, 9:1671–1681.

Oseledec, V.I. (1968). A multiplicative ergodic theorem: Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.*, 19:197–231.

Ottino, J.M., Muzzio, F.J., and Tjahjadi, M. (1992). Chaos, symmetry, and self-similarity: Exploiting order and disorder in mixing processes. *Science*, 257:754–760.

S. Sato, M.S. and Sawada, Y. (1987). Practical methods of measuring the generalized dimension and the largest lyapunov exponent in high dimensional chaotic systems. *Prog. Theor. Phys.*, 77:1.

Schiff, S.J., Jerger, K., and Duong, D.H. (1994). Controlling chaos in the brain. *Nature*, 370:615–620.

Wisniewski, R. and Larsen, L.F.S. (2008). Method for analysis of synchronization applied to supermarket refrigeration system. In *17th IFAC World Congress*. Seoul, Korea.

Wolf, A., Swift, J.B., Swinney, H.L., and Vastano, J.A. (1985). Determining lvapunov exponents from a time series. *Physica D*, 16:285–317.

A hierarchical approach to optimal control of a hybrid chromatographic batch process ⋆

Dmitry Gromov [*] **Suzhou Li** [**] **Jörg Raisch** [*, **]

[*] *Fachgebiet Regelungssysteme, Technische Universität Berlin*
e-mail: {gromov,raisch}@control.tu-berlin.de
[**] *Max-Planck-Institut für Dynamik komplexer technischer Systeme, Magdeburg, Germany*
e-mail: suzhou@mpi-magdeburg.mpg.de

Abstract: In this paper we consider a hierarchical approach to solve an optimal control problem for a hybrid chromatographic batch process. The plant consists of several chromatographic columns which can be connected in an arbitrary way. The plant configuration can therefore be considered as a discrete-valued control input. The dynamics of each chromatographic column is described by two coupled nonlinear partial differential equations. Hence, we have a hybrid optimisation problem with highly nonlinear dynamics.
To handle complexity, we propose a hierarchical two-level optimisation scheme: first, we solve a number of continuous optimisation problems that correspond to fixed configurations. In a subsequent step, on the basis of these solutions, we solve a discrete optimisation problem to generate the optimal configuration or configuration signal.
Because of the imposed structure, we can of course not expect the overall solution to be optimal. However, we demonstrate that, by using the plant configuration as an an additional control input, performance may be considerably improved when compared to the constant configuration scenario.

Keywords: Hierarchical optimisation, optimal control, hybrid systems, batch chromatography

1. INTRODUCTION

Column chromatography using solid and liquid phases is a key technique for the isolation and purification of valuable products, which has found a large number of successful applications in petrochemical, food and pharmaceutical industries during the last decades. Chromatographic separation processes can be operated in continuous or batch mode. Continuous separation is usually realised by the well-known simulated moving bed (SMB) process and its various modifications. This scheme has been proven to be very efficient for large-scale separation tasks. A drawback is, however, that the start-up procedure may take considerable time. Hence, this technology may not be suitable for the separation for relatively small amounts of mixtures. In this case, batch chromatography is an attractive alternative. Currently, a considerable number of chromatographic separations are operated in batch mode. Therefore the efficient operation and control of these processes is an important topic in order to exploit the economic potential and reduce the production cost.

There are a number of papers devoted to the optimisation and parameter identification of batch separation processes, see, e.g., (Dünnebier et al., 2001; Gao and Engell, 2005; Nagrath et al., 2003; Piątkowski, 2006) for details and references. These papers study the problem of optimisation of the process w.r.t. different performance criteria (e.g., productivity, or more specific criteria, like in (Felinger and Guiochon, 1996, 1998)). These criteria describe the overall (integral) performance of the system, but they are not very suitable if there are additional operational restrictions such as fixed batch volume, separation time and so on. Such restrictions naturally appear if the chromatographic system is a part of a complex chemical plant whose operation must follow a certain schedule. Moreover, in some applications the configuration of the plant is an additional degree of freedom, as the plant consists of a number of chromatographic columns that can be arranged in different ways. This degree of freedom has not been widely investigated up to now. An exception is (Ziomek et al., 2006), where the plant configuration is a design parameter, but constant over time. Also, additional operational restrictions (e.g., fixed batch size etc.) are not considered there.

In this contribution, we aim to develop a general framework to optimal (open loop) control of a chromatographic batch process. It includes several practically important problem statements and covers the scenario where the plant configuration may change during the operation of the plant. The latter introduces an additional, discrete-valued, degree of freedom, which makes the overall control problem an intrinsicly hybrid one. To deal with the inherent complexity of this hybrid problem, we suggest a hierarchical approach, where a lower control level determines the continuous inputs and a higher control level solves the remaining discrete optimisation problem.

⋆ Work has been done in the framework of the project "Methods from Discrete Mathematics for the Synthesis and Control of Chemical Processes", DFG-FG 468

This paper is organised as follows: In Section 2, we define the plant and a suitable PDE model. In Section 3, we motivate different optimisation problems for this plant. Section 4 suggests a hierarchical approach to solve these problems, and Section 5 presents a numerical example.

2. PLANT DESCRIPTION

2.1 Mathematical model

The system consists of N identical chromatographic columns, which can be arranged in $N_\ell \leq N$ parallel lines. Via valves connected to the columns, their configuration can be changed within a very short time. The column configuration can therefore be interpreted as a control input. The number of columns in the i-th line is denoted by $N^i_{col}(t)$, with the obvious restriction $\sum_{i=1}^{N_\ell(t)} N^i_{col}(t) = N$, $\forall t$.

Figure 1 shows a configuration with (at time t) $N = 5$, $N_\ell(t) = 2$, $N^1_{col}(t) = 3$, and $N^2_{col}(t) = 2$.

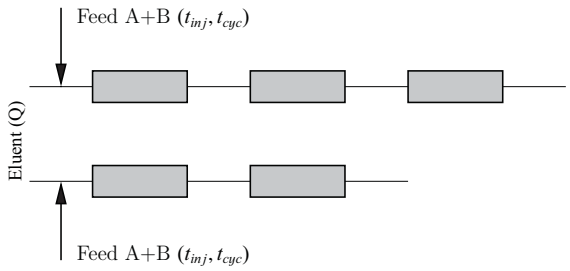


Fig. 1. Example configuration.

The dynamics of a single column is described by two nonlinearly coupled second order PDEs (so called equilibrium-dispersive model, (Guiochon et al., 2006)):

$$\frac{\partial c_k(t,x)}{\partial t} + F\frac{\partial q_k(t,x)}{\partial t} + u\frac{\partial c_k(t,x)}{\partial x} = D_{ap}\frac{\partial^2 c_k(t,x)}{\partial x^2}, \quad k \in \{A,B\},$$

where c_k and q_k are the liquid phase and solid phase concentrations, t, x are the temporal and spatial coordinates, u is the velocity of the liquid phase in the column, $F = (1-\epsilon_t)/\epsilon_t$ is the phase ratio, and ϵ_t is the column total void fraction. D_{ap} is the apparent axial dispersion coefficient defined as a function of u: $D_{ap} = uf(u)/2$, where $f(u) = (\alpha_D + \beta_D u)$ represents the linearised van Deemter equation, α_D and β_D are constant coefficients. In the following, we shall often use the volumetric flowrate $Q = u\pi D^2\epsilon_t/4$ instead of u, where D is the diameter of the column.

The concentrations of the components in the solid and liquid phases are related via the isotherm equation. To describe adsorption, the competitive Langmuir model is employed:

$$q_k(c_A, c_B) = \frac{H_k c_k}{1 + K_A c_A + K_B c_B}, \quad k \in \{A,B\},$$

where H_k and K_k are the Henry and the equilibrium constants. Component A is assumed to be less retained than component B, therefore, $H_A < H_B$.

The initial and boundary conditions are defined in a standard way: the initial concentration of the components in the columns is equal to zero. We employ the Danckwerts boundary conditions for the inlet of the first and the outlet of the last column in each line:

$$\left[uc_k(t,x) - D_{ap}\frac{\partial c_k(t,x)}{\partial x}\right]_{x=0} = uc_{k,in}(t)$$
$$\left.\frac{\partial c_k(t,x)}{\partial x}\right|_{x=L} = 0,$$

where $k \in \{A,B\}$, L is the column length. The "inter-column" boundary conditions reflect the continuity of the concentrations profiles.

The functioning of a single chromatographic column is shown schematically in Fig.2. The solution containing a binary mixture is injected at the inlet of the column during the time interval t_{inj}. Injections are repeated cyclically, with the interval between two subsequent injections t_{cyc} (Fig.2a). The mixture is transported through the column with velocity u. During transportation, separation occurs because one of the components (in our example component A) is less retained than the other one. Fig.2b shows a snapshot of the concentration profiles within the column at two time instants: at an "early" time instant, the separation effect is weak, and the concentration profiles for the two components are very close (dashed lines). At a later time instant, the profiles have moved further apart. Finally, component A is collected from the outlet during the fractionation interval t_{fr}. Fig.2c shows the concentration signals $c_{A,out}(t)$ and $c_{B,out}(t)$ at the outlet of the column. t_1 denotes the time instant when the concentration of the less retained component A exceeds a given threshold $c_{A,thr}$ and t_4 denotes the time when the concentration c_B becomes less than another threshold.

The following entities are used to characterise the separation process within each line and during each injection cycle. For component A, they are:

- The mass output
$$m_{A,out} = \int_{t_1}^{t_2} c_{A,out}(t)Q dt.$$
- The purity
$$Pur_A = \frac{\int_{t_1}^{t_2} c_{A,out}(t)dt}{\int_{t_1}^{t_2} c_{A,out}(t)dt + \int_{t_1}^{t_2} c_{B,out}(t)dt}.$$
- The yield
$$Y_A = \frac{m_{A,out}}{V_{inj} c_{A,in}},$$
where $V_{inj} = Q\, t_{inj}$ is the injection volume.
- The productivity
$$Pr_A = \frac{m_{A,out}}{t_{cyc}} = \frac{V_{inj} c_{A,in} Y_A}{t_{cyc}}. \quad (1)$$

For component B, they are defined in an analogous way.

There are also a number of technological constraints imposed on the system. Some of them are listed below (for details see Ziomek et al. (2006)). For example, for each line of columns, we have:

- Two restrictions on the volumetric flowrate. The first one is caused by the the maximal pressure drop ΔP_{max}, namely $Q \leq Q_{max}(\Delta P_{max})$. The maximal

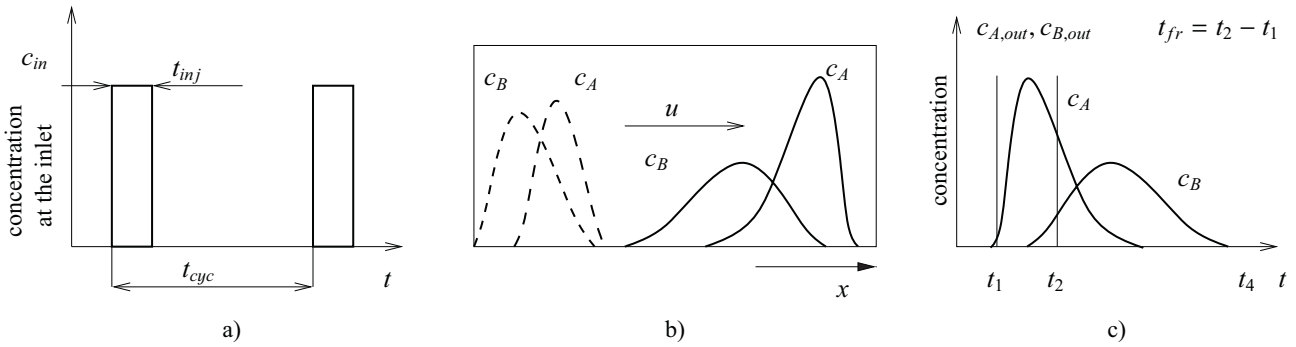


Fig. 2. a)input signal; b) concentration profiles within the column; c) concentration signals at the outlet.

pressure drop is calculated from the Darcy equation. Furthermore, there is a restriction on the maximal capacity of the pump Q^p_{max}. Hence, the resulting constraint is written as follows (Guiochon et al., 2006):

$$Q \leq \min(Q_{max}(\Delta P_{max}), Q^p_{max}).$$

- Minimal fractionation time caused by technical limitations: $t_{fr} \geq t_{fr,min}$.

2.2 Decision parameters

There are a number of parameters that can be chosen to optimise the process. We can divide them into two groups: discrete parameters, which can take values in a finite set, and continuous parameters, which can take values in a dense subset of $\mathbb{R}$. The continuous parameters are defined for each line $i \in \{1, N_\ell\}$:

(1) The velocity of the liquid phase, u^i,
(2) The injection time, t^i_{inj},
(3) The cycle period, t^i_{cyc},
(4) The fractionation time, t^i_{fr}.

The first parameter enters the PDEs directly whereas the second and the third one enter the boundary conditions of the first column within a line. The last parameter is usually determined to satisfy purity requirements.

The discrete parameters are the number of lines, N_ℓ, and the number of columns in the i-th line, N^i_{col}. These parameters describe the configuration of the plant. Furthermore, for a given configuration, the number of injections N^i_{inj} in line i is also a degree of freedom. In the following, we shall consider two cases:

(1) The configuration is constant during the entire operation of the process.
(2) The configuration changes over time.

In the second case, the sequence

$$\{(N_\ell(j), N^1_{col}(j), \ldots, N^{N_\ell(j)}_{col}(j), N^1_{inj}(j), \ldots, N^{N_\ell(j)}_{inj}(j), \tau(j))\}, \quad j = 0, N_s$$

can be interpreted as a control signal, with N_s the (a priori fixed) number of configuration changes, $\tau(j)$ the time interval during which the plant is operated in the j-th configuration, $N_\ell(j)$ the number of lines for this configuration, $N^i_{col}(j)$ the number of columns in the i-th line in the j-th configuration, and $N^i_{inj}(j)$ the number of injections for the i-th line in the j-th configuration.

In the remainder of this paper, we will use the following assumptions:

A1 The continuous control parameters for all parallel lines can be adjusted separately.
A2 The only component we are interested in is component A (less retained). In the following, we shall omit the subscript identifying the component if this is clear from the context.
A3 The continuous control parameters do not change within the interval of constant configuration $\tau(j)$.
A4 The cycle time is chosen to be equal to the duration of the chromatogram: $t^i_{cyc} = t^i_4 - t^i_1,\ i \in \{1, N_\ell\}$.
A5 The inlet concentrations $c_{k,in},\ k \in \{A, B\}$ are assumed to be known and fixed $\forall t$.

3. OPTIMISATION PROBLEMS

We shall investigate the following overall optimisation problems:

(1) Yield maximisation for fixed overall time span T_{max}:

$$\begin{aligned} &\text{maximise } Y_\Sigma \\ &s.t.\ T_\Sigma \leq T_{max}, \end{aligned} \tag{2}$$

where overall yield Y_Σ is defined as

$$Y_\Sigma = \frac{\sum_{j=0}^{N_s} V_\Sigma(j) Y_\Sigma(j)}{\sum_{j=0}^{N_s} V_\Sigma(j)}. \tag{3}$$

In (3), $V_\Sigma(j)$ and $Y_\Sigma(j)$ are the processed volume and the yield for the j-th configuration:

$$V_\Sigma(j) = \sum_{i=1}^{N_\ell(j)} N^i_{inj}(j) V^i_{inj}(j),$$

$$Y_\Sigma(j) = \frac{\sum_{i=1}^{N_\ell(j)} N^i_{inj}(j) m^i_{out}(j)}{c_{in} V_\Sigma(j)}.$$

Overall time T_Σ is given by $T_\Sigma = \sum_{j=0}^{N_s} \tau(j)$, where

$$\tau(j) \geq \max_{i=1,N_\ell(j)} (N^i_{inj}(j) t^i_{cyc}(j)).$$

(2) Produce required yield Y_{min} in minimal time:

$$\begin{aligned} &\text{minimise } T_\Sigma \\ &s.t.\ Y_\Sigma \geq Y_{min}. \end{aligned} \tag{4}$$

In both cases, we have additional constraints for overall batch size (volume): $V_\Sigma = V_{batch}$, and for overall purity:

$$Pur_\Sigma = \frac{\sum_{j=0}^{N_s} V_\Sigma(j) Y_\Sigma(j)}{\sum_{j=0}^{N_s} \frac{V_\Sigma(j) Y_\Sigma(j)}{Pur_\Sigma(j)}} \geq Pur_{min},$$

where

$$Pur_\Sigma(j) = \frac{\sum_{i=1}^{N_\ell(j)} N^i_{inj}(j) m^i_{out}(j)}{\sum_{i=1}^{N_\ell(j)} N^i_{inj}(j) \frac{m^i_{out}(j)}{Pur^i(j)}}.$$

Note that m^i_{out} and Pur^i are the mass output and the purity in the i-th line for one cycle and depend on the continuous decision parameters as well as on the number of columns in line i, N^i_{col}.

4. HIERARCHICAL APPROACH

The optimisation problems posed in Sec.3 are highly complex tasks. In the variable configuration case, there is a large number of discrete and continuous decision parameters. In each iteration step during the optimisation procedure, $2N_\ell$ partial differential equations have to be solved. Moreover, the values of the cost functions cannot be found analytically and have to be calculated from the results of numerical simulation. Since information about the derivatives of the cost functions is not available, derivative-free methods must be used, which substantially reduces the efficiency of numerical optimisation. Therefore, standard solvers normally fail to deliver a solution to these optimisation problems in reasonable time (if they provide a solution at all).

Therefore, we propose to use a hierarchical optimisation scheme to cope with complexity. In this scheme, optimisation of discrete and continuous decision variables is decoupled.

4.1 Low-level (local) optimisation

On the low level, we define a set of continuous optimisation problems for one line of columns over one cycle period. The cost function for the low-level problem is productivity (Eq.1). The set of low-level problems is parametrised by the number of columns n (n=1, N) in a line and a (finite) number of purity constraints $Pur \geq p_m$, $m = 1, M$:

$$\begin{aligned} &\underset{(Q, t_{inj}, t_{fr})}{\text{maximise}} \quad Pr \\ &s.t.\ N_{col} = n, \\ &\quad Pur \geq p_m. \end{aligned} \tag{5}$$

For each low-level problem, the optimal solution ($Q^*, V^*_{inj} = t^*_{inj} Q^*, t^*_{fr}$) and the corresponding values (t^*_{cyc}, m^*_{out}) need to be stored and will be used for the solution of a high-level problem. This information can be conveniently collected in the following table.

		Pur		
		p_1	...	p_M
N_{col}	1	.	.	.
	⋮	.	.	.
	n	.	.	Q^*, V^*_{inj}, t^*_{fr}, t^*_{cyc}, m^*_{out}
	⋮	.	.	.
	N	.	.	.

Table 1. Results of the local optimisation

4.2 High-level optimisation

a) Constant configuration case:

On the high level we maximise overall productivity

$$Pr_\Sigma = \frac{\sum_{i=1}^{N_\ell} N^i_{inj} m^i_{out}}{T_\Sigma} \tag{6}$$

$$\text{s.t. } Pur_\Sigma = \frac{\sum_{i=1}^{N_\ell} N^i_{inj} m^i_{out}}{\sum_{i=1}^{N_\ell} N^i_{inj} \frac{m^i_{out}}{Pur^i}} \geq Pur_{min} \tag{7}$$

$$V_\Sigma = \sum_{i=1}^{N_\ell} N^i_{inj} V^i_{inj} = V_{batch} \tag{8}$$

and such that

$$T_\Sigma = \max_{i=1,N_\ell} (N^i_{inj} t^i_{cyc}), \tag{9}$$

or such that

$$Y_\Sigma = \frac{\sum_{i=1}^{N_\ell} N^i_{inj} m^i_{out}}{c_{in} V_\Sigma}. \tag{10}$$

Note that for fixed time T_Σ and fixed batch volume V_Σ, the maximisation of overall productivity is equivalent to the maximisation of overall yield. Conversely, for fixed batch volume and fixed overall yield, the maximisation of productivity is equivalent to the minimisation of the overall processing time. Therefore, the optimisation problem (6-8) and (9) or (10) can be seen as a general formulation encompassing both our original optimisation problems described in Sec. 3.

The decision variables for the high-level optimal problem are N_ℓ, N^i_{inj}, $i = 1, N_\ell$, and pairs (n^i, m^i) representing the entry in the n^i-th row and m^i-th column of Table 1. Naturally, the restriction $\sum_{i=1}^{N_\ell} n^i = N$ has to hold. As N^i_{inj}, the number of injections in the i-th line, is always bounded by the problem setup, the high-level problem has

a finite search space. Note that to evaluate the overall cost function, only the values of the decision variables and the corresponding entries in Tab. 1 are needed. In particular, no numerical simulations are required. In effect, this means that we assemble an overall solution from the solutions of the low-level optimisation problem.

While we cannot expect the resulting solution to be globally optimal, it seems reasonable that we shall obtain a decent approximation.

b) Variable configuration case:

In the variable configuration case, the overall cost function (6) changes to

$$Pr_{\Sigma} = c_{in} \frac{\sum_{i=0}^{N_s} V_{\Sigma}(j) Y_{\Sigma}(j)}{\sum_{i=0}^{N_s} \tau(j)} \tag{11}$$

with the obvious corresponding changes in (7)-(10). The only difference in terms of the optimisation procedure is an increase of the cardinality of the search space.

From a practical point of view, the question whether a significant improvement is possible when allowing a variable configuration is of prior interest. Our example in the next section indicates that switching configuration may indeed improve performance.

5. NUMERICAL EXAMPLE

As an example of the proposed approach we consider a case with 5 chromatographic columns, $N = 5$. The numerical values of parameters and restrictions for a single column are listed in Appendix A. We investigate the maximisation of yield within a given time. The required minimal purity is 95%. Time T_{max} is equal to 20000s. Different overall batch volumes are considered.

5.1 Low-level optimisation

First, we solved a set of low-level optimisation problems of the form (5). Optimisation was performed for the following values of required purity: $p_m \in \{0.9, 0.91, \ldots, 0.99\}$ and for different numbers of columns in the line: $N_{col} \in \{1, 2, 3, 4, 5\}$. The optimal values of the continuous decision variables $(Q^*, V_{inj}^*, t_{fr}^*)$, as well as the corresponding values (t_{cyc}^*, m_{out}^*) were stored as a table as indicated in the previous section.

An improved derivative-free Nelder-Mead method (Nelder and Mead, 1965; Kelley, 1999) was used to solve these nonlinear constrained optimisation problems. In the proposed variant, the initial simplex was generated randomly and a multi-restart strategy was taken to increase the probability of locating the global optimum. For each subsequent restart, only the vertex with the best solution found previously was retained and the other vertices were replaced with new random points. Moreover, the algorithm was capable of handling the nonlinear constraints by using the penalty function technique. During the optimisation, a numerical simulation procedure was used to generate the outlet concentration signals for given values of decision variables. From this, we evaluate the objective function and check the constraints. This information was then used by the optimiser to find a new direction in the continuous search space to improve the cost.

The coupled PDE model of the considered process with N_{col} columns connected in series was discretised using the method of orthogonal collocations on finite elements (Ma and Guiochon, 1991; Kaczmarski et al., 1997). The resulting system of differential algebraic equations was solved by ode15s, a variable step-size and variable order integrator implemented in Matlab (Shampine and Reichelt, 1997). Information about the Jacobian contained in the discretised model equations was fully exploited by the solver, thereby significantly accelerating the integration.

5.2 High-level optimisation

a) Constant configuration case:

Next, we solved the discrete optimisation problem (2) for the constant configuration case, as described in Sec.4.2a. The results of the optimisation procedure are presented in Tab. 2.

Volume of the batch, V_{batch}, ml	Optimal configuration	Optimal purities, Pur^i, %	Resulting productivity Pr, g/s	Resulting yield Y
23000	{3 2}	0.94 0.97	5.942 e-4	0.517
25000	{3 2}	0.95 0.95	6.2404 e-4	0.4992
27000	{3 2}	0.96 0.94	6.4118 e-4	0.475
28000	{2 2 1}	0.95 0.96 0.97	6.0756 e-4	0.424
29000	{2 2 1}	0.95 0.95 0.98	6.2808 e-4	0.421
31000	{2 2 1}	0.95 0.96 0.94	6.6047 e-4	0.3901

Table 2. Results of high-level optimisation for constant configuration scenario

b) Variable configuration case:

It can be seen from Table 2 that the configuration {3 2} ensures high productivity for smaller batch volumes. For higher volumes, this configuration cannot process the required volume within the required time. Hence, from the certain batch volume V_{batch}, one has to use another configuration, for example {2 2 1}. This configuration ensures higher throughput at the cost of productivity. Therefore, it is intuitive to consider configuration as a time-variant degree of freedom, i.e., a control variable. In this way one expects to combine the advantages of both configurations.

We considered the case with one possible configuration change, i.e., $N_s = 1$. Indeed, our optimisation results show an increase of performance for the variable configuration case. The results are shown in Table 3. It is worth noting that the improvement could be much bigger if there were bigger differences in the productivity and in the throughput between the different configurations.

ACKNOWLEDGEMENTS

We acknowledge many helpful discussions with A. Seidel-Morgenstern, L.C. Keßler and G. Ziomek.

Volume of the batch, V_{batch}	Optimal configuration	Time, $\tau(j)$, $j = 0, 1$	Volume, $V_\Sigma(j)$	Optimal purities, $Pur^i(j)$	Resulting productivity, Pr_Σ	Resulting yield, Y_Σ	Improvement compared to constant configuration case
28000	{2 2 1}	10600	16100	0.95 0.95 0.96	6.4305 e-4	0.4603	+5.84%
	{3 2}	9400	11900	0.95 0.95			
29000	{2 2 1}	10800	16675	0.95 0.95 0.96	6.5115 e-4	0.4509	+3.7%
	{3 2}	9200	12325	0.96 0.94			

Table 3. Results of the high-level optimisation with the change of structure

REFERENCES

Dünnebier, G., Engell, S., Epping, A., Hanisch, F., Jupke, A., Klatt, K.U., and Schmidt-Traub, H. (2001). Model-based control of batch chromatography. *AIChE Journal*, 47(11), 2493–2502.

Felinger, A. and Guiochon, G. (1996). Optimizing preparative separations at high recovery yield. *Journal of Chromatography A*, 752, 31–40.

Felinger, A. and Guiochon, G. (1998). Comparing the optimum performance of the different modes of preparative liquid chromatography. *Journal of Chromatography A*, 796(1), 59–74.

Gao, W. and Engell, S. (2005). Iterative set-point optimization of batch chromatography. *Computers & Chemical Engineering*, 29(6), 1401–1409.

Guiochon, G., Felinger, A., and Shirazi, S. (2006). *Fundamentals of Preparative and Nonlinear Chromatography*. Elsevier.

Kaczmarski, K., Mazzotti, M., Storti, G., and Morbidelli, M. (1997). Modelling fixed-bed adsorption columns through orthogonal collocations on moving finite elements. *Comput. Chem. Eng.*, 21(6), 641–660.

Kelley, C. (1999). *Iterative Methods for Optimization*. SIAM.

Ma, Z. and Guiochon, G. (1991). Application of orthogonal collocation on finite elements in the simulation of nonlinear chromatography. *Comput. Chem. Eng.*, 15(6), 415–426.

Nagrath, D., Bequette, B.W., and Cramer, S.M. (2003). Evolutionary operation and control of chromatographic processes. *AIChE Journal*, 49(1), 82–95.

Nelder, J. and Mead, R. (1965). A simplex method for function minimization. *Comput. J.*, 7(4), 308–313.

Piątkowski, W. (2006). Analysis of the design and optimization of preparative chromatography on the basis of the separation of a real post-reaction mixture. *Acta Chromatographica*, 16, 92–118.

Shampine, L. and Reichelt, M. (1997). The MATLAB ODE suite. *SIAM J.Sci.Comput.*, 18.

Ziomek, G., Antos, D., Tobiska, L., and Seidel-Morgenstern, A. (2006). Comparision of possible arrangements of five identical columns in preparative chromatography. *Journal of Chromatography A*, 1116, 179–188.

Appendix A. LIST OF PARAMETERS

Name	Unit	Description	Value
Physical (geometric) parameters			
L	cm	Column length	25
D	cm	Column diameter	2
ϵ_t	-	Column total void fraction	0.78
α_D	-	Coefficient of the linearised van Deemter equation	0.012
β_D	-	Coefficient of the linearised van Deemter equation	0.156
H_1	-	Henry constant	5.7
H_2	-	Henry constant	7.4
K_1	ml/g	Equilibrium constant	170
K_2	ml/g	Equilibrium constant	370
Operating parameters			
$C_{i,in}$, i=A, B	g/ml	Feed concentrations	0.001
$C_{i,thr}$	g/ml	Threshold concentrations	$0.001C_{i,in}$
N	-	Number of columns	5
t_{fr}^{min}	s	Minimal collecting (fractionating) time	5
ΔP_{max}	bar	Maximal pressure drop	100
Q_{max}	ml/s	Volumetric flowrate corresponding to the maximal pump capacity	8.33

Sensitivity-based Predictive Control of a Large-scale Supermarket Refrigeration System ⋆

Christian Sonntag * **Malte Kölling** ** **Sebastian Engell** *

* *Process Dynamics and Operations Group (BCI-DYN), Dept. of Biochemical and Chemical Engineering, Technische Universität Dortmund, 44221 Dortmund, Germany E-mail: {christian.sonntag|sebastian.engell}@bci.tu-dortmund.de.*
** *Hydro Aluminium, 41515 Grevenbroich, Germany E-mail: malte.koelling@hydro.com*

Abstract: Today, many supermarket refrigeration systems are operated by decentralized control systems that often lead to excessive starting and stopping of the compressors which drive the cooling cycle and, consequently, to a large wear of the process equipment. In our previous work, a hierarchical model-predictive control scheme was proposed for supermarket refrigeration systems that overcomes this drawback. In this scheme, simple low-level temperature controllers are employed, and the high-level optimization task is the optimal adjustment of the parameters of these controllers. While this approach yields a good control performance, it is computationally too expensive for larger systems. In this paper, a more efficient approach is presented that is based on an approximation of the system dynamics using simple models that are computed from system sensitivities around simulated reference trajectories. The application of the new approach to a large hybrid model proves the real-time capabilities of the new technique.

Keywords: Predictive control, process control, supermarket refrigeration systems, discretely controlled continuous systems, hybrid systems.

1. INTRODUCTION

In supermarket refrigeration systems, a rack of compressors feeds liquid refrigerant to several open display cases that are used to cool edible goods. These systems exhibit both, discrete and continuous dynamics, and are thus hybrid systems: The control inputs (valves and compressors) can only be switched discretely, and the nonlinear continuous dynamics changes due to switching of the discrete inputs. Today, supermarket refrigeration systems are often controlled using decentralized schemes in which each display case is equipped with independent simple control loops (Larsen et al., 2005). Since this approach often causes a severe reduction of the efficiency of the process and of the lifespan of the equipment (see e.g. Wisniewski and Larsen (2008)), the suitability of advanced model-predictive schemes for the control of supermarket refrigeration systems has been investigated in previous work to overcome these problems.

In Larsen et al. (2005), the hybrid MPC approach from Bemporad and Morari (1999) is applied to a piecewise affine approximation of the nonlinear hybrid model of a supermarket refrigeration system that is also considered in this paper. Although this approach succeeds in keeping most process variables within pre-specified bounds, the frequency of the compressor switching is high due to the inaccuracy of the linear approximations of the nonlinear dynamics. In Sarabia et al. (2009), a nonlinear MPC scheme is proposed in which the cost function is evaluated by simulation of a nonlinear model. This approach is capable of keeping all process variables within the bounds, but the solution of complex NLP problems with many decision variables in each iteration leads to a large computational effort. In Sonntag et al. (2007, 2008), a hierarchical NMPC approach for supermarket refrigeration systems is presented. Here, the switching of the valves of the display cases is not optimized directly, but simple low-level controllers are employed that regulate the temperatures in the display cases with a high sampling frequency. The parameters of these controllers are adjusted by a high-level NMPC optimizer that operates on a longer time horizon, thus leaving more computation time for the NLP step in every NMPC iteration. Instead of considering a (complex) MINLP problem in each iteration, the discrete search is performed by solving a sequence of continuous optimization problems with an increasing number of switches, and the search is stopped as soon as a policy is found that meets the specification. This reflects the main control goal, the minimization of the number of switches of the compressors.

While the NLP-based approaches that are described above yield a good control performance, they cannot currently be applied to larger systems in real time due to the prohibitively large computational effort. To overcome this problem, this paper presents a new and computationally

⋆ The authors gratefully acknowledge the financial support by the EU-funded Network of Excellence HYCON, contract number FP6-IST- 511368.

more efficient hierarchical model-predictive control approach for supermarket refrigeration systems. As in Sontag et al. (2007, 2008), the switching of the valves of the display cases is not optimized directly, but simple low-level controllers are employed that regulate the temperatures in the display cases. The high-level approach presented in this paper differs from the previous version in two respects: the compressors are not switched by the high-level controller anymore since the corresponding control goal for the suction pressure is merely safety-related, and it is not necessary to achieve optimality in this part of the system. A simple discrete low-level controller is employed instead. Furthermore, the optimization approach of the high-level controller is significantly different. It is based on the observation that the considered supermarket refrigeration system belongs to an important subclass of hybrid systems, the *Discretely Controlled Continuous Systems* (DCCS). These systems have been investigated in academia for many years, see e.g. Matveev and Savkin (2000); Dankowicz and Piiroinen (2002); Schild and Lunze (2008). The property of DCCS that is exploited in this paper is that simple yet accurate finite-dimensional models of the system behavior (the so-called *embedded maps* of the system) can be approximated linearly around simulated reference trajectories based on the sensitivities of the continuous subsystems. These low-dimensional embedded maps are then used as linear constraints in the optimizing high-level control system, and the complex dynamic optimization problem can be recast as a sequence of algebraic optimization problems. In combination with a decomposition of the large-scale supermarket system into (virtually) independent subcomponents which are approximated separately, this new approach allows for a very efficient computation of optimal switching times which enables the real-time control even of large-scale supermarket systems.

2. THE SUPERMARKET REFRIGERATION SYSTEM

Fig. 1 shows a schematic representation of a supermarket refrigeration system. It consists of five major parts: a liquid manifold, several display cases, a compressor rack, a suction manifold, and a condenser. Liquid refrigerant is supplied to the display cases from the liquid manifold through inlet valves (see Fig. 2). Within each display case, cold air circulates and forms an air curtain in front of the edible goods. Thermal energy is transferred from the goods to the air curtain ($\dot{Q}_{goods-air}$) and, since the temperature of the surrounding air is larger than that of the air curtain, the curtain also absorbs heat from the surroundings ($\dot{Q}_{airload}$). The absorbed thermal energy is transported to the evaporator ($\dot{Q}_{air-wall}$) in which the refrigerant evaporates and thus takes on the thermal energy ($\dot{Q}_e$). The vapor accumulates in the suction manifold and is fed to the condenser via the compressors which increase the pressure of the refrigerant vapor. Since the evaporation temperature of the refrigerant increases with the pressure, the energy from the display cases can be removed in the condenser at room temperature. Finally, the liquefied refrigerant is fed back to the display cases.

The hybrid model of the supermarket refrigeration system used in this work was proposed in Larsen et al. (2007). It may contain an arbitrary number of display cases n_{dc}. The state of each display case $i \in \{1, \ldots, n_{dc}\}$ is described by four differential state variables: the temperature of the goods ($T_{g,i}$), the temperature of the evaporator wall ($T_{w,i}$), the temperature of the air inside the case ($T_{air,i}$), and the mass of liquid refrigerant within the evaporator of the display case ($m_{ref,i}$). Thus, the vector of continuous state variables of the i-th display case is given by $\mathbf{x}_{dc,i} = [T_{g,i}, T_{w,i}, T_{air,i}, m_{ref,i}]^T$. Since the dynamics of the condenser unit is not modeled, the overall continuous state vector of the model can be written as $\mathbf{x} = [\mathbf{x}_{dc,1}^T, \ldots, \mathbf{x}_{dc,n_{dc}}^T, P_{suc}]^T$, where P_{suc} is the pressure in the suction manifold. Each display case is equipped with an expansion valve for the refrigerant, and the discrete input vector is given by $\mathbf{v} = [v_1, \ldots, v_{n_{dc}}, v_c]^T$. Here, $v_1, \ldots, v_{n_{dc}} \in \{1, 0\}$ are binary variables representing the state of the inlet valves (open/closed), and $v_c \in \Xi_c$ (given in %) determines the relative capacity of the compressors that are currently running within the compressor rack [1]. The set Ξ_c contains all discrete capacity levels that can be realized by switching the compressors on or off. In this paper, a system with six compressors of equal capacity is investigated. For this system, Ξ_c is defined as:

$$\Xi_c := \{0\%, 16.7\%, 33.3\%, 50\%, 66.7\%, 83.3\%, 100\%\}. \quad (1)$$

[1] Thus, the values 0% (100%) always indicate that all compressors are off (on), independently of the number of compressors.

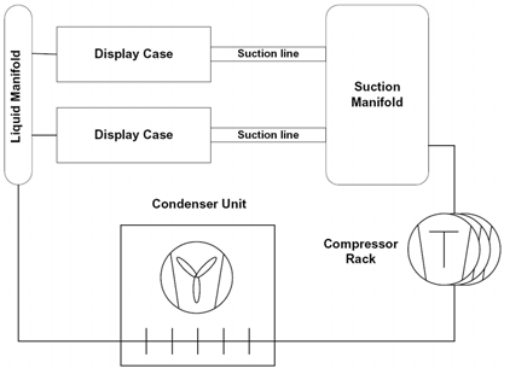


Fig. 1. A simplified scheme of a supermarket refrigeration system with two display cases.

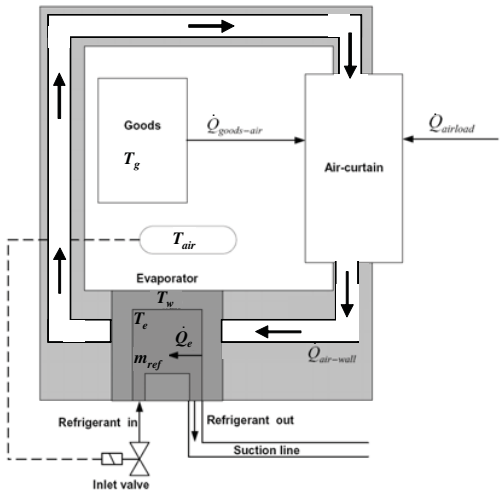


Fig. 2. Cross section of a display case.

The continuous dynamics is modeled by a lumped-parameter ODE system [2] under the assumption that all display cases are of equal design. Each display case exhibits two different continuous dynamics, depending on the setting of the corresponding expansion valve, i.e.

$$\frac{d\mathbf{x}_{dc,i}}{dt} = \begin{cases} \mathbf{f}_{i,vo}(\mathbf{x}_{dc,i}) \text{ if } v_i = 1, \text{ (a)} \\ \mathbf{f}_{i,vc}(\mathbf{x}_{dc,i}) \text{ if } v_i = 0. \text{ (b)} \end{cases} \quad (2)$$

The vector functions $\mathbf{f}_{i,vo}$ and $\mathbf{f}_{i,vc}$ only differ in the dynamic equation that determines the mass of refrigerant in the evaporator of a display case $m_{ref,i}$ according to

$$\frac{dm_{ref,i}}{dt} = \begin{cases} \dfrac{m_{ref,max} - m_{ref,i}}{\tau_{fill}} \text{ if } v_i = 1, \text{ (a)} \\ -\dfrac{\dot{Q}_{e,i}}{\Delta h_{lg}} \qquad \text{if } v_i = 0. \text{ (b)} \end{cases} \quad (3)$$

Here, the maximum mass of refrigerant each display case can accommodate is represented by $m_{ref,max}$, $\dot{Q}_{e,i}$ is defined in Eq. 9, the specific enthaply of evaporation of the remaining liquefied refrigerant in the evaporator is given by Δh_{lg}, and τ_{fill} is a time constant. The display case is filled with refrigerant as long as the inlet valve is open (Eq. 3.a), and after the inlet valve has been closed, the remaining refrigerant evaporates according to Eq. 3.b. The temperature dynamics within the i-th display case does not change with the valve setting and is given by:

$$\frac{dT_{g,i}}{dt} = -\frac{\dot{Q}_{goods-air,i}}{m_{goods} \cdot cp_{goods}}, \quad (4)$$

$$\frac{dT_{w,i}}{dt} = \frac{\dot{Q}_{air-wall,i} - \dot{Q}_{e,i}}{m_{wall} \cdot cp_{wall}}, \quad (5)$$

$$\frac{dT_{air,i}}{dt} = \frac{\dot{Q}_{goods-air,i} + \dot{Q}_{airload} - \dot{Q}_{air-wall,i}}{m_{air} \cdot cp_{air}}, \quad (6)$$

with

$$\dot{Q}_{goods-air,i} = UA_{goods-air} \cdot (T_{g,i} - T_{air,i}), \quad (7)$$

$$\dot{Q}_{air-wall,i} = UA_{air-wall} \cdot (T_{air,i} - T_{w,i}), \quad (8)$$

$$\dot{Q}_{e,i} = UA_{wall-ref}(m_{ref,i}) \cdot (T_{w,i} - T_e(P_{suc})), \quad (9)$$

$$UA_{wall-ref}(m_{ref,i}) = UA_{wall-refmax} \cdot \frac{m_{ref,i}}{m_{ref,max}}. \quad (10)$$

Here, m_{goods}, m_{wall}, m_{air}, cp_{goods}, cp_{wall}, cp_{air}, $UA_{goods-air}$, $UA_{air-wall}$, and $UA_{wall-refmax}$ are constant model parameters, and T_e is the evaporation temperature of the refrigerant which is a nonlinear function of P_{suc}. The dynamics of the suction pressure is given by

$$\frac{dP_{suc}}{dt} = \frac{\dot{m}_{in-suc} + \dot{m}_{ref-const} - \dot{V}_c \cdot \rho_{suc}}{V_{suc} \cdot \frac{d\rho_{suc}}{dP_{suc}}}, \quad (11)$$

with

$$\dot{m}_{in-suc} = \sum_{i=1}^{n_{dc}} \frac{\dot{Q}_{e,i}}{\Delta h_{lg}}, \quad \dot{V}_c = \frac{v_c \cdot \eta_{vol} \cdot V_d}{100}. \quad (12)$$

Here, the total mass flow of refrigerant from all display cases into the suction manifold is given by $\dot{m}_{in-suc}$, and $\dot{m}_{ref-const}$ is a measurable external disturbance that represents an additional flow of refrigerant from other unmodeled cooling facilities into the suction manifold. $\dot{V}_c$ is the volume flow from the suction manifold, and ρ_{suc} and $\frac{d\rho_{suc}}{dP_{suc}}$ are nonlinear refrigerant-dependent functions modeling the density of the vapor in the suction manifold and the derivative of ρ_{suc} w.r.t. the suction pressure, and η_{vol} and V_d are constant model parameters.

The controlled variables of the system are the pressure inside the suction manifold (P_{suc}) and the temperatures of the air inside the display cases ($T_{air,i}$). As the system never reaches a steady state since the different continuous dynamics of the display cases have distinct equilibrium points, the control goal is not to track setpoints, but to maintain the controlled variables within specified bounds $\underline{T}_{air,i} \leq T_{air,i} \leq \overline{T}_{air,i}$ and $\underline{P}_{suc} \leq P_{suc} \leq \overline{P}_{suc}$.

3. THE CONTROL STRATEGY

A scheme of the hierarchical control strategy is shown in Fig. 3. This strategy is very similar to the hierarchical approach that was presented in Sonntag et al. (2008), with one important difference: while in Sonntag et al. (2008), the parameters of the low-level controllers as well as the settings of the compressors are adapted by the high-level optimizer, the controlled subsystems are separated into two categories in the new approach: The first category consists of subsystems for which the desired control functionality is only safety-related, i.e. the control goal is to keep a process variable within an admissible region, and a quantitative measure of optimality is not necessary. Among the subsystems of the supermarket refrigeration system, the suction manifold and the compressor rack fall into this category and are not considered in the high-level predictive controller. The second category consists of subsystems for which quantitative optimality measures can be defined. In the supermarket system, the display cases belong to this category since here the control goal, the temporal desynchronization of the air temperatures, can be formulated in a quantitative way that is amenable to minimization. This temporal desynchronization ensures that the variations in the suction pressure and, thus, the necessity for compressor switching are minimized.

3.1 The Low-Level Control System

The switching strategy for the valves of display case i is shown in Fig. 4. The valve of the corresponding display case remains closed as long as the air temperature remains below the switching threshold δ_s. After the refrigerant has evaporated, the air temperature starts to rise. Once the air temperature crosses δ_s from below, v_i is opened for a constant period of time t_{v_i}, and the air temperature will decrease again. The time period t_{v_i} is a continuous parameter that is assigned by the high-level controller for

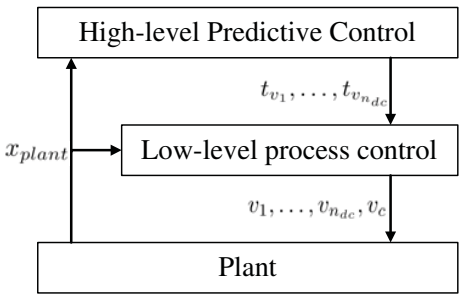


Fig. 3. Scheme of the control strategy.

[2] See also Larsen et al. (2007).

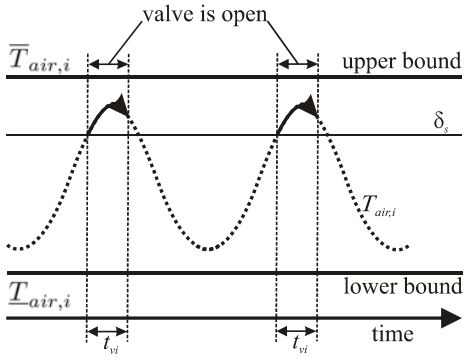


Fig. 4. Switching strategy for the expansion valves for an exemplary evolution of the air temperature of display case i.

each display case ($t_{v_1}, \ldots, t_{v_{n_{dc}}} \in \mathbb{R}^{\geq 0}$ in Fig. 3). The value of δ_s was determined in simulation studies assuming that $T_{air,i}$ will always decrease shortly after the valve v_i is opened (which can be deduced from the continuous model dynamics and parameters).

The low-level pressure controller switches off a compressor if $P_{suc} \leq \underline{P}_{suc}$, and it switches on an additional compressor if $P_{suc} \geq \overline{P}_{suc} - 0.1\ bar$. To avoid excessive switching of the compressors over a short time period, a compressor can only be switched 10 seconds after the previous compressor switch at the earliest. To compensate for fast changes of the external disturbances, an additional controller is employed that monitors the stationary continuous compressor capacity v_{cs} that is needed to keep all process variables within the admissible region over long time periods. From the ODE system, v_{cs} can be computed as:

$$v_{cs} = 100\ \% \cdot \left(\frac{\dot{m}_{in-suc} + \dot{m}_{ref-const}}{\rho_{suc} \cdot \eta_{vol} \cdot V_d} \right). \qquad (13)$$

If v_{cs} changes by more than $\delta_c = \frac{1}{2} \cdot \frac{100\ \%}{n_c}$ (n_c is the number of compressors) over a time period of 30 seconds, the controller switches the compressors to the discrete capacity level that is closest to v_{cs}.

3.2 Desynchronizing High-level Control

The high-level controller operates on a moving time horizon. In every iteration, a fixed time interval of t_p seconds is available for the adaptation of the valve opening times $t_{v_1}, \ldots, t_{v_{n_{dc}}}$ of the low-level temperature controllers. The algorithm is based on the assumption that all display cases can be regarded separately, i.e. that the interactions between the display cases are negligible. To confirm this assumption, a sensitivity analysis was performed for the system in which the cross-correlations between the system variables were computed by the solution of the matrix-valued linear sensitivity equation

$$\frac{d\mathbf{S}}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \cdot \mathbf{S}, \quad \mathbf{S}(t_0) = \mathbf{I}. \qquad (14)$$

Here, $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ is the Jacobian of the dynamic vector equation of the system. It was found that the effect of changes of the state variables $\mathbf{x}_{dc,i}$ of a display case i and the variables $\mathbf{x}_{dc,j}$ of other display cases with $j \neq i$ as well as the effect of changes of the suction pressure P_{suc} on $\mathbf{x}_{dc,i}$ are several orders of magnitude smaller than the effects between the internal variables of the display case in all operating regimes that are relevant for nominal process operation. Thus, the display cases can be considered as independent subsystems for control design purposes.

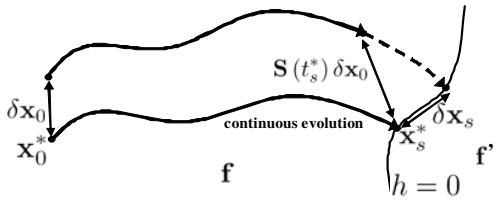


Fig. 5. On the linear approximation of the embedded map.

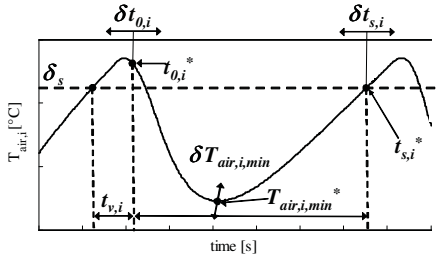


Fig. 6. Application of the approximation procedure to a single display case.

The desynchronization algorithm is based on the computation of an abstract algebraic model (an embedded map) for each display case which is locally valid around a simulated reference trajectory. Such a model maps small deviations $\delta\mathbf{x}_0$ of an initial continuous state $\mathbf{x}_0^*$ to the corresponding deviations $\delta\mathbf{x}_s$ of a reference state $\mathbf{x}_s^*$ that lies on a switching threshold which is defined by the zero-level set of a switching function h (see Fig. 5). The linear approximation is computed by (Parker and Chua, 1989):

$$\delta\mathbf{x}_s = \underbrace{\left[I - \frac{\mathbf{f}(\mathbf{x}_s^*) \cdot \frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}_s^*)}{\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}_s^*) \cdot \mathbf{f}(\mathbf{x}_s^*)} \right]}_{(1)} \underbrace{\mathbf{S}\,(t_s^*)}_{(2)} \delta\mathbf{x}_0 \qquad (15)$$

Here, $\mathbf{f}(\mathbf{x}_s^*)$ is the evaluation of the dynamic equations at $\mathbf{x}_s^*$, $\frac{\partial h}{\partial \mathbf{x}}(\mathbf{x}_s^*)$ is the gradient of h evaluated at $\mathbf{x}_s^*$, and $\mathbf{S}\,(t_s^*)$ is the solution of Eq. 14 at t_s^* with $\mathbf{S}(t_0^*) = \mathbf{I}$. Term (2) of Eq. 15 maps the evolution of the initial deviation $\delta\mathbf{x}_0$ along the reference trajectory while the term (1) projects the resulting state along $\mathbf{f}$ onto the zero-level set of h.

Fig. 6 depicts how this mapping is adapted for each display case i. Under the assumption that the air temperature $T_{air,i}$ exhibits periodic behavior after the first period for constant disturbances and for unchanged $t_{v,i}$, the remainder of the time evolution of $T_{air,i}$ after $t_{s,i}^*$ can be neglected for control purposes. The state and sensitivity trajectories that are needed for the approximation of the embedded maps are computed simultaneously by the simulation of the hybrid model. Since the display cases are assumed to be independent, the sensitivity equations are evaluated separately for each display case which drastically reduces the number of sensitivities that must be computed.

The main idea is to formulate the resulting linear model of the controlled system as an algebraic optimization problem in which the finite-state embedded maps of the display cases appear as linear constraints [3] . Depending on the type of the cost function that is chosen, the overall problem can then be stated as a linear, a quadratic, or even a nonlinear programming problem. The control goal for the supermarket system is to desynchronize the air temperatures which corresponds to a uniform temporal distribution of the "time points of impact" $t^*_{s,i}$ of the air temperatures on the switching threshold δ_s. Thus, we are only interested in how a deviation $\delta t_{0,i}$ of the nominal valve closing time $t^*_{0,i}$ changes the next "time point of impact" $t^*_{s,i}$ of the air temperature $T_{air,i}$ on the switching threshold δ_s. Hence, a mapping $\delta t_{s,i} = H \cdot \delta t_{0,i}$ must be derived.

For simplicity, a new state vector $\mathbf{x}_{em,i} = [\mathbf{x}_{dc,i}, t]$ is defined for each display case that includes the time as an additional state. Since the switching function $h_i = T_{air,i} - \delta_s$ represents a constant switching threshold for each display case, its gradient is given by $\partial h_i / \partial \mathbf{x_{em,i}} = [0, 0, 1, 0, 0]$. Substituting this equation into Eq. 15, evaluating the right-hand side (with $\mathbf{f} = \mathbf{f_{i,vc}}$, see Eq. 3), and considering only the last row of the resulting matrix yields

$$\delta\, t_{s,i} = -\frac{1}{\frac{dT_{air,i}}{dt}(t^*_{s,i})} \cdot \mathbf{s}(t^*_{s,i}) \cdot \delta \mathbf{x}_{0,i} \tag{16}$$

with

$$\mathbf{s}(t^*_{s,i}) = \begin{bmatrix} s_{T_{air,i},T_{g,i}}(t^*_{s,i}) \\ s_{T_{air,i},T_{w,i}}(t^*_{s,i}) \\ s_{T_{air,i},T_{air,i}}(t^*_{s,i}) \\ s_{T_{air,i},m_{ref,i}}(t^*_{s,i}) \\ -1 \end{bmatrix}^T \tag{17}$$

In this equation, $s_{x_1,x_2}(t^*_{s,i})$ corresponds to the solution of the sensitivity equation at time $t^*_{s,i}$ that represents the effect of a change of x_2 on x_1. The final step is to express $\delta \mathbf{x}_{0,i}$ in terms of $\delta t_{0,i}$ which is achieved by linear interpolation: in addition to the simulation that is performed to determine the reference trajectory, a second simulation for the maximally allowed variation $\delta t_{0,i,max}$ from $t^*_{0,i}$ yields the maximal deviation of the state variables $\delta \mathbf{x}_{0,i,max}$. Here, $\delta t_{0,i,max}$ is a constant design parameter that is determined a priori by simulation such that the error of the linear approximation is negligible. Now, $\delta \mathbf{x}_{0,i}$ can be related linearly to $\delta t_{0,i}$ by $\delta\ \mathbf{x}_{0,i} = \frac{\delta \mathbf{x}_{0,i,max}}{\delta t_{0,i,max}} \cdot \delta\ t_{0,i}$. To ensure that the air temperatures do not violate the lower temperature bounds $\underline{T}_{air,i}$, additional linear constraints are derived that represent the variation $\delta T_{air,i,min}$ of the minimal air temperature $T_{air,i,min}$ of the reference trajectory with a variation of $t^*_{0,i}$ (see Fig. 6).

Since the embedded maps are only valid in a neighborhood $\epsilon := [\delta t^*_{0,i} - \delta t_{0,i,max}, \delta t^*_{0,i} + \delta t_{0,i,max}]$ around the reference trajectory, an algorithm is used that iterates between the generation of reference trajectories and optimization until convergence to the optimal solution is achieved. In each iteration j, the following steps are performed:

(1) **Determination of the impact order:** A reference trajectory and the order of the impact points $t^*_{s,i}$ are determined by simulation with the optimal values $t^*_{0,j-1}$ from the previous iteration. This order remains fixed in iteration j.

(2) **Computation of optimal impact points:** The air temperature of the display case with the earliest impact point $t^*_{s,\bullet}$ is driven to the lower bound $T_{air,i,min}$ using the linear approximations. Considering the resulting time trajectory over two periods yields lower (impact time point after the first period) and upper (impact time point after the second period) reference values t_{min} and t_{max} for the impact time points of all other air temperatures. The optimal impact points $t^*_{s,i}$ for all other air temperatures are distributed equidistantly in the range $[t_{min}, t_{max}]$.

(3) **Computation of optimal parameters:** Since the optimal impact points are known for all air temperatures, the corresponding valve closing times $t^*_{0,i}$ are computed by inverting the embedded maps ($\delta t^*_{0,i} = H^{-1} \cdot \delta t^*_{s,i}$). If all $t^*_{0,i}$ are within ϵ, the algorithm **terminates** since the optimal values have been found. If one or more $t^*_{0,i}$ are outside ϵ, go to step (4).

(4) **Recomputation of the reference trajectory:** A new reference trajectory is computed using values for the valve opening times that are determined from the results of step (3). Here, all values that are outside the neighborhood ϵ are replaced by the upper (if they are larger than the maximum value in ϵ) or the lower (if they are smaller than the minimum value in ϵ) limits of ϵ. Then, the algorithm returns to step (3).

4. APPLICATION RESULTS

The optimization algorithm was implemented in *Matlab* and was tested with a large-scale supermarket refrigeration system with 10 display cases and 6 compressors. Fig. 7 shows the optimization results for a day-night scenario. From 0 to 7200 seconds, the system is in day-time operation, and after 7200 seconds, a night-time operation is assumed. During the day, the masses of the goods are varied to model the removal by customers and the replenishment by the supermarket staff, as shown in Fig. 7 (d). Tab. 1 shows the parameter values that were used in the simulation. The controller is capable of keeping all process variables within the bounds and desynchronizes the air temperatures very quickly, as is shown in the lower part of 7 (a). Furthermore, the low-level compressor controller detects the drastic change in the external disturbances at 7200 seconds and switches off four compressors to counteract the sudden change in the stationary compressor capacity. A comparison to previously obtained results for smaller systems shows that the control scheme significantly reduces the frequency of the compressor switching. As expected, the very major contribution to the run-time of the algorithm is the computational effort for the simulation of the nonlinear model. During nominal operation with only slowly varying disturbances, the algorithm only needs to execute very few simulations since the valve opening times are already close to the optimal values. Thus, an MPC iteration only takes a few seconds in this case. In the worst case, i.e. when the temperatures are completely synchronized, the computation time to achieve a complete

[3] Note that for the supermarket system, the optimal solution can be computed analytically from the embedded maps, as is described below. For other systems, however, optimization may be necessary.

desynchronization was less than 100 seconds on a standard PC. With some optimization of the prototype implementation, it seems realistic to achieve worst-case computation times in the region of 20 seconds.

Table 1. Parameter values used in the simulation studies.

n_{dc}	n_c	δ_s
10	6	4.6 °C °C
t_p	$\underline{T}_{air,1}$ - $\underline{T}_{air,10}$	$\overline{T}_{air,1}$ - $\overline{T}_{air,10}$
100 s	2 °C	5 °C
$t_{0,i,max}$	$\underline{P}_{suc}$	$\overline{P}_{suc}$ (day/night)
1 s	1 bar	1.7 bar / 1.9 bar
$\dot{Q}_{airload}$ (day)	$\dot{Q}_{airload}$ (night)	
3000 W	1800 W	
$\dot{m}_{ref-const}$ (day)	$\dot{m}_{ref-const}$ (night)	
0.2 $\frac{kg}{s}$	0.0 $\frac{kg}{s}$	

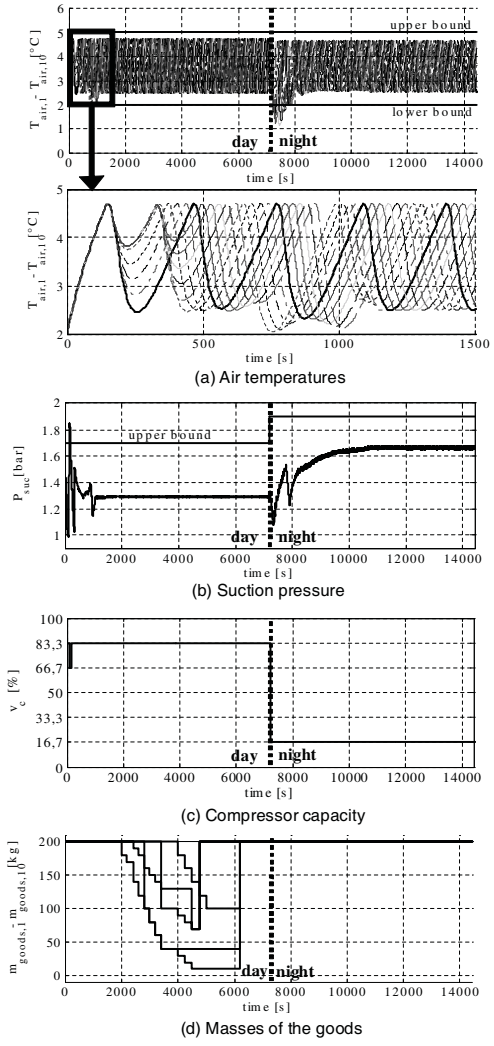


Fig. 7. Simulation results for a supermarket refrigeration system with 10 display cases and 6 compressors.

5. CONCLUSIONS

In this paper, a new hierarchical approach for the real-time capable control of large-scale supermarket refrigeration systems with hybrid dynamics is presented. Simple low-level temperature controllers are employed, and the high-level control task is the optimal adjustment of the parameters of these controllers to achieve a desynchronization of the air temperatures in the display cases which reduces the wear of the process equipment. Efficient desynchronizing control is achieved using a combination of model decomposition and approximation of the system dynamics by simple algebraic models. The main advantages of this approach over existing control techniques for supermarket refrigeration systems are that (a) the system is not linearized a priori. Any nonlinear characteristics of the system that are encoded in the reference trajectories are considered implicitly in the high-level control scheme, and (b) the computational performance is improved considerably since the original mixed-integer nonlinear dynamic optimization problem is replaced by a sequence of low-dimensional analytic problems that can be solved efficiently. Future work will concentrate on the extension of the developed technique to more general hybrid systems.

REFERENCES

Bemporad, A. and Morari, M. (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35, 407–427.

Dankowicz, H. and Piiroinen, P.T. (2002). Exploiting discontinuities for stabilization of recurrent motions. *Dynamical Systems*, 17(4), 317–342.

Larsen, L.F.S., Geyer, T., and Morari, M. (2005). Hybrid MPC in supermarket refrigeration systems. In *Proc. 16th IFAC World Congress*. Th-E12-TO/5.

Larsen, L.F.S., Zamanabadi, R.I., Wisniewski, R., and Sonntag, C. (2007). Supermarket refrigeration systems - a benchmark for the optimal control of hybrid systems. Technical report for the Network of Excellence HYCON. Http://tinyurl.com/23nrkc.

Matveev, A.S. and Savkin, A.V. (2000). *Qualitative Theory of Hybrid Dynamical System*. Birkhäuser.

Parker, T.S. and Chua, L.O. (1989). *Practical Numerical Algorithms for Chaotic Systems*. Springer.

Sarabia, D., Capraro, F., Larsen, L.F.S., and De Prada, C. (2009). Hybrid NMPC of supermarket display cases. *Control Engineering Practice*, 17(4), 428–441.

Schild, A. and Lunze, J. (2008). Switching surface design for periodically operated discretely controlled continuous systems. In *Hybrid Systems: Computation and Control*, volume 4981 of *LNCS*, 471–485.

Sonntag, C., Devanathan, A., and Engell, S. (2008). Hybrid NMPC of a supermarket refrigeration system using sequential optimization. In *Proc. 17th IFAC World Congress*, 13901–13906.

Sonntag, C., Devanathan, A., Engell, S., and Stursberg, O. (2007). Hybrid nonlinear model-predictive control of a supermarket refrigeration system. In *Proc. IEEE Multi-Conference on Systems and Control (MSC/CCA)*, 1432–1437.

Wisniewski, R. and Larsen, L.F.S. (2008). Method for analysis of synchronization applied to supermarket refrigeration system. In *Proc. 17th IFAC World Congress*, 3665–3670.

PWA Modelling and Co-ordinated Continuous and Logical Control of a Laboratory Scale Plant with Hybrid Dynamics

J. Hlava

Faculty of Mechatronics, Technical University of Liberec
461 17 Liberec, Studentská 2, Czech Republic,
(e-mail: jaroslav.hlava@tul.cz)

Abstract: Many process plants are nonlinear and together with this they include a combination of continuous valued and logical control inputs and subsystems. This paper attempts to explore the potential of hybrid model predictive control (MPC) to cope with both of these problems. It uses a laboratory scale plant that was designed for experiments with hybrid systems. This plant has both continuous and logical control inputs and it is considerably nonlinear. An approximate hybrid model of the plant in the form of a piecewise affine (PWA) system is developed and evaluated in the first part of the paper. After that a hybrid MPC based on PWA model is applied to the control of the plant. While designing hybrid MPC and evaluating its performance, there is a special focus on the following question. Logical and continuous control systems are usually designed separately. This may result in unforeseen interactions between logical and continuous control and in the deterioration of the control performance. However, hybrid MPC is based on hybrid model that captures both logical and continuous dynamics in one unified framework. Hence it can reasonably be expected that hybrid MPC can avoid undesirable interactions and possibly also make use of these interactions in a positive way (e.g. to speed up the control response using logical inputs). Control results obtained with hybrid MPC are indeed fairly good and they show clear improvement over the results achieved with separate design of logical and continuous control.

Keywords: Hybrid systems, model predictive control, piecewise affine systems

1. INTRODUCTION

Model predictive control (MPC) of hybrid systems has recently attracted a considerable research attention. This attention is reflected in the growing number of publications on hybrid MPC. Monographs such as (Christophersen, 2007), (Borrelli, 2003) and survey paper (Morari & Baric, 2006) can be quoted as important examples representing a vast and constantly growing body of literature. The application area of hybrid MPC is twofold. First, many process plants comprise continuous-valued as well as logical/discrete-valued control inputs and components. Such plants are naturally modelled as hybrid systems and this requires the use of control approaches for hybrid systems. Second, non-linearities that are ubiquitous in the models of process plants can often be well approximated by a special class of hybrid systems called piecewise affine (PWA) systems. The result is again a plant model in the form of a hybrid system.

This paper is focused on both of the above mentioned aspects of the hybrid MPC. It uses a case study of a laboratory scale plant. This plant exhibits hybrid phenomena that are found in many process control applications. The plant includes both continuous valued and logical control inputs and its dynamic behaviour abruptly changes at certain operating points. Most continuous components of the plant are nonlinear. This nonlinear behaviour must be approximated by a PWA model. This approximation is necessary for the design of hybrid MPC controller and it is a non-trivial task. PWA approximation, selection of individual affine models, their validity regions and comparison with the responses of the original nonlinear plant are described in detail. Finally a hybrid model is obtained whose hybrid features are both due to the hybrid nature of the plant itself and due to the PWA approximation of plant nonlinearities. Further, the attention is turned to MPC control of this plant. A special emphasis is laid on the ability of hybrid MPC to achieve an integrated design of logical and continuous control.

Typically, logical control is responsible for safety related and limiting functions such as preventing the process variables from leaving safe operation limits, starting and shutdown of process equipment. Logical controllers are also used to manipulate logical control inputs such as on/off valves. On the other hand, the regulatory and supervisory control is performed by continuous controllers. Common design practice relies on separate design of logic and continuous control. As non-trivial and not easily predictable interactions often arise between continuous and logical parts of the control system, this practice may result in a poor control performance. On the other hand, hybrid model describes both continuous and logical (or more generally discrete-valued) parts of the whole system, including continuos/logical interactions, the hybrid controller designed on the basis of this model can be expected to control the whole plant in a co-ordinated manner and avoid the deteriorating effects of interactions between separately designed logical and continuous control systems. However, it is well known that expectations though well founded in the theory and practical reality may be two different worlds. For this reason, this

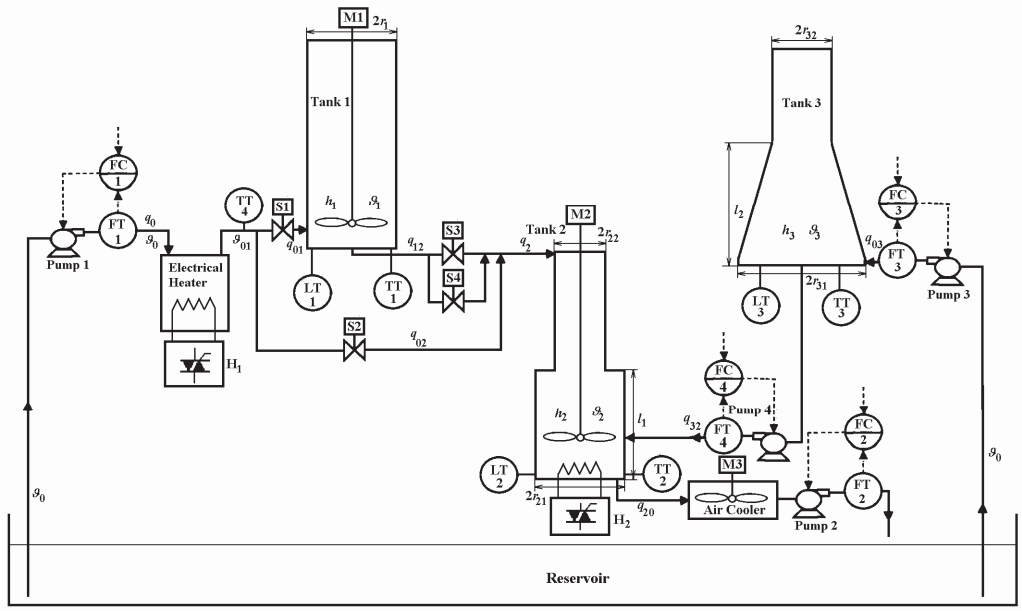


Fig. 1. Structure of the laboratory scale plant, FT, LT, TT are flow, level and temperature transmitters FC – flow controller, S– solenoid valve, M – motor, r_1= 5.64 cm, r_{21}= 5.8 cm, r_{22}= 3 cm, r_{31}= 6 cm, r_{32}= 2.9 cm, tank height l_{max}=80 cm, l_1= l_2=40 cm

paper in its final part attempts to make an experimental comparison of a separate design of logic and continuous control on the one hand and co-ordinated design based on hybrid model on the other hand.

2. EXPERIMENTAL PLANT

A detailed description of the experimental plant has recently been given by the author in (Hlava & Šulc, 2008). As the full text of this paper is available from the IFAC-PapersOnLine website, the description of this plant in the present paper can be short. Plant structure is shown in Fig. 1. Basic components are three water tanks. Tanks 2 and 3 have special shapes that introduce changes in dynamics. The tanks are thermally insulated to make the heat losses negligible. Water from the reservoir mounted under the plant is drawn by Pump 1 and Pump 3 to the respective tanks. The delivery rates can be continuously changed. The flow rates are measured using turbine flow-meters. To compensate for pump non-linearity, it is beneficial to use slave flow rate controllers.

The flow from Pump 3 is fed directly to Tank 3. The flow from Pump 1 goes through heater and it is further controlled by solenoid valve S1. The power output of the heater can be changed continuously and ϑ_{01} can be made to follow a specified function of time. Another continuously controlled heater is mounted on the bottom of Tank 2. The temperatures are measured with Pt1000 sensors at the points shown in Fig. 1. In addition to the pumps, whose delivery flow rates can be changed continuously, the plant includes another way of manipulating the flow: solenoid valves. These discrete valued actuators control the flow from Tank 1 to Tank 2 (valves S3, S4). The flow is changed in three steps: no valve open, one open, both valves open. Tank 1 can be bypassed by closing S1 and opening S2. The air-water heat exchanger with cooling fan at the output from Tank 2 keeps the water temperature in the reservoir roughly constant. Water levels are measured using pressure sensors.

The plant is controlled from a PC using two data acquisition boards (11 analog inputs, 6 analog outputs, 6 digital outputs) and interface hardware (power amplifiers, solid state relays, signal conditioning devices). The basic software tool for identification and control experiments is the Real Time Toolbox. It allows an easy connection of Matlab/Simulink environment with the real world. Alternatively WinCon-8000 industrial control system produced by ICP DAS can be used. This control system makes it possible to experiment with the implementation of advanced control algorithms using real industrial hardware. The changeover from PC to WinCon-8000 and vice versa is simple: two connectors with analog and digital inputs/outputs have to be reconnected.

3. CONTROL OBJECTIVE

Many control scenarios can be defined with this plant. Some examples are given in (Hlava & Šulc, 2008). The scenario considered in this paper uses Tank 1 and 2. This scenario is inspired in part by (Slupphaug *et al.*, 1997) and it can be formulated as follows. Tank 1 serves as a buffer that receives water from an upstream process. Water flow rate and temperature are disturbances. The main control objective is to deliver the water to a downstream process at a desired temperature (temperature ϑ_2), while the flow demand of the downstream process is variable and hence it also acts as a disturbance. Power output of heater H_2 is a continuous manipulated variable. Valves S3 and S4 are used as a discrete valued manipulated variable to control the flow from Tank 1 to Tank 2 in three steps. Valve S1 is used to close water flow to Tank 1, if tank overflow is to be avoided. There is no valve at the output of Tank 2, but an effect equivalent to closing an output valve is achieved by switching off Pump 2.

The main control objective necessarily includes several auxiliary objectives. Tank levels must be kept within specified limits, and overflow as well as emptying of the tanks must be avoided. It is also necessary to avoid the necessity to close valve S1 in order to prevent Tank 1 overflow. In a real control situation, closing S1 would mean that water from the upstream process cannot flow to the buffer but must be re-routed to the environment. Similarly it is necessary to avoid the necessity to switch off Pump 2 in order to prevent Tank 2 from underflow. The standard way to satisfy these auxiliary objectives would be to use separately designed control logic.

4. MATHEMATICAL MODEL OF THE PLANT

Plant model is derived using mass and energy balances. The reader is referred to (Hlava & Šulc, 2008) for details. In this paper just the part of the plant model will be given that is relevant to the specified control objective (i.e. excluding Tank 3). Assuming liquid incompressibility and constant heat capacity c, negligible heat losses and ideal mixing, the following model is obtained

$$\dot{h}_1(t) = (1/A_1)\left(q_0(t)\sigma_0(t) - 0.1k_v\sigma_1(t)\sqrt{gh_1(t)}\right) \tag{1}$$

$$\dot{h}_2(t) = \begin{cases} (1/A_{21})\left(0.1k_v\sigma_1(t)\sqrt{gh_1(t)} - q_{20}(t)\sigma_2(t)\right) & h_2(t) \le l_1 \\ (1/A_{22})\left(0.1k_v\sigma_1(t)\sqrt{gh_1(t)} - q_{20}(t)\sigma_2(t)\right) & h_2(t) > l_1 \end{cases} \tag{2}$$

$$\dot{\vartheta}_1(t) = q_0(t)\sigma_0(t)\left(\vartheta_{01}(t) - \vartheta_1(t)\right)/A_1h_1(t) \tag{3}$$

$$\dot{\vartheta}_2(t) = \begin{cases} \dfrac{\left(0.1k_v\sigma_1(t)\sqrt{gh_1(t)}\left(\vartheta_1(t) - \vartheta_2(t)\right) + \dfrac{H(t)}{\rho c}\right)}{A_{21}h_2(t)} & h_2(t) \le l_1 \\ \dfrac{\left(0.1k_v\sigma_1(t)\sqrt{gh_1(t)}\left(\vartheta_1(t) - \vartheta_2(t)\right) + \dfrac{H(t)}{\rho c}\right)}{A_{21}l_1 + A_{22}(h_2(t) - l_1)} & h_2(t) > l_1 \end{cases} \tag{4}$$

where $A_i = \pi r_i^2$, discrete valued input σ_0 assumes values 0,1 (S1 closed, S1 open), σ_1 assumes values 0,1,2 (no valve open, S3 open, S3 and S4 open), σ_2 assumes values 0,1 (Pump 2 off, Pump 2 running with flow rate q_{20} depending on the flow demand of the downstream process), H is power output of heater H_2, k_v is flow coefficient of valves S3 and S4.

5. APPROXIMATE PLANT MODEL IN A PWA FORM

Plant model (1)-(4) includes continuous and discrete valued inputs, dynamics switching depending on operating point in (2), (4) and non-linear elements. It must first be approximated by a PWA model. The general form of a discrete-time PWA system is given by

$$\begin{aligned} \boldsymbol{x}(k+1) &= \boldsymbol{M}_i\boldsymbol{x}(k) + \boldsymbol{N}_i\boldsymbol{u}(k) + \boldsymbol{f}_i \\ \boldsymbol{y}(k) &= \boldsymbol{C}_i\boldsymbol{x}(k) + \boldsymbol{D}_i\boldsymbol{u}(k) + \boldsymbol{g}_i \end{aligned} \tag{5}$$

where each dynamics i=1,2..N_D is active in a polyhedral partition D that is defined by guard lines described by

$$\boldsymbol{G}_i^x\boldsymbol{x}(k) + \boldsymbol{G}_i^u\boldsymbol{u}(k) \le \boldsymbol{G}_i^c \tag{6}$$

That means, the dynamics i represented by matrices and vectors [$\boldsymbol{M_i}$, $\boldsymbol{N_i}$, $\boldsymbol{f_i}$, $\boldsymbol{C_i}$, $\boldsymbol{D_i}$, $\boldsymbol{g_i}$] is active in the region of state-input space which satisfies constraints (6). Unlike some other approaches to hybrid MPC that use probabilistic Bayesian approach to combine weighted local linearized models (e.g. Nandola & Bhartiya, 2008), the approach considered in this paper is deterministic and local models are just switched depending on the region in state-input space.

It has been noted already in the well known seminal paper on PWA systems (Sonntag, 1981) that nonlinear systems can be globally approximated arbitrarily close by PWA systems and this claim has often been repeated. This claim is certainly true. However, it is also true that although there are several methods of experimental identification of PWA models (see Paoletti *et al.*, 2007 for an overview), there is no general systematic procedure to find a PWA approximation of a given non-linear system described by analytical state equations. The route to the PWA approximation is always closely connected with a particular system to be approximated. In this section, model described by (1)- (4) will be considered. Its approximation by a set of affine models of the form (5), (6) can proceed as follows.

An obvious source of partial models is logical control inputs σ_0, σ_1, σ_2. The best way to handle these inputs is to associate one partial model with each combination of their values. This results in 12 partial models. Equations (2) and (4) include dynamics switching at water level l_1. That means the number of 12 must be doubled and 24 partial models are obtained as an absolute minimum for modeling this plant. These 24 partial models must further be linearized. To achieve an acceptable precision, each model is approximated by a set of linearized models. The linearization is done in two steps.

1. Obtain linearizations around general operating points characterized by a vector of input and state variables (σ_{0P}, σ_{1P}, σ_{2P}, q_{0P}, q_{20P}, ϑ_{0P}, H_p, h_{1P}, h_{2P}, ϑ_{1P}, ϑ_{2P}). If possible, steady state operating points should be preferred.
2. Find a suitable set of operating points together with adequate partitioning of state-input space that will be well representative of the dynamics of the original system.

This procedure can be most simply illustrated for (1). If S1 is open (σ_{0P}=1) the respective steady state characteristics is

$$q_{0P} = \frac{1}{\sigma_{0P}}0.1k_v\sigma_{1P}\sqrt{gh_{1P}}\ , \tag{7}$$

and (1) can be linearized around a steady state operating point (h_{1P}, q_{0P}, σ_{0P}, σ_{1P}). This linearization is given by

$$\dot{h}_1(t) = \frac{\sigma_{0P}}{A_1}(q_0(t) - q_{0P}) - \frac{0.1k_v\sigma_{1P}}{2A_1}\sqrt{\frac{g}{h_{1P}}}\,(h_1(t) - h_{1P}) \tag{8}$$

Substituting for q_{0P} from (7), linearization can be modified to

$$\dot{h}_1(t) = \frac{\sigma_{0P}}{A_1}q_0(t) - \frac{0.1k_v\sigma_{1P}\sqrt{g}}{2A_1\sqrt{h_{1P}}}\,h_1(t) - \frac{0.1k_v\sigma_{1P}\sqrt{gh_{1P}}}{2A_1} \tag{9}$$

It can be seen that (9) holds even if σ_{1P}=0. What is less obvious is the case of σ_{0P}=0. Equation (7) cannot be used and (1) is autonomous system that has just zero steady state. However, it can be linearized around a non-steady state

$$\dot{h}_{1P}(t) = -0.1k_v\sigma_{1P}\sqrt{gh_{1P}}\,/A_1 \tag{10}$$

and this linearization has the same form as (9). Thus (9) is general linearized approximation of (1). As actual values of the variables and not deviations from operating point are used, (9) is affine and not linear.

The next step is to find a suitable set of representative operating points. The simplest approach would be to divide the whole range of h_1 into several intervals of identical length and to take midpoints of these intervals as selected nominal operating points. A better way is to modify (9) to the form

$$\frac{2A_1\sqrt{h_{1P}}}{0.1k_v\sigma_{1P}\sqrt{g}}\dot{h}_1(t) + h_1(t) = \frac{2\sigma_{0P}\sqrt{h_{1P}}}{0.1k_v\sigma_{1P}\sqrt{g}}q_0(t) - h_{1P} \tag{11}$$

Dynamics of (11) can be characterised by time constant

$$\tau = 2A_1\sqrt{h_{1P}}\,/\left(0.1k_v\sigma_{1P}\sqrt{g}\right) \tag{12}$$

and the partitioning of the whole range of h_1 can be done in such a way as to keep the ratio of maximum and minimum time constant within each interval the same. The nominal operating point in each interval is then again selected so that the nominal time constant would have the same ratio to the minimum and maximum value of time constant within this interval. If the whole range is divided into four intervals, this reasoning leads to the following

$$\begin{aligned}
&h_{1m1} = h_{1min}^{\frac{3}{4}}h_{1max}^{\frac{1}{4}}; \quad h_{1m2} = h_{1min}^{\frac{2}{4}}h_{1max}^{\frac{2}{4}}; \quad h_{1m3} = h_{1min}^{\frac{1}{4}}h_{1max}^{\frac{3}{4}}\\
&h_{1P1} = \sqrt{h_{1min}h_{1m1}}; \quad h_{1P2} = \sqrt{h_{1m1}h_{1m2}}; \quad h_{1P3} = \sqrt{h_{1m2}h_{1m3}}\\
&h_{1P4} = \sqrt{h_{1m3}h_{1max}}
\end{aligned} \tag{13}$$

where $h_{1\min}$ and $h_{1\max}$ are the minimum and maximum values of h_1 respectively, h_{1m1}, h_{1m2}, h_{1m3} are limiting points of subintervals, h_{1P1}, h_{1P2}, h_{1P3}, h_{1P4} are nominal operating points.

Other parts in plant model can be approximated by PWA systems in a similar way. The dynamics of water level in Tank 2 as described by (2) depends mainly on h_1, while h_2 just governs switching between two partial models. Linearization of (1) as described by (9) actually means that square root was replaced by the linear part of Taylor series. Using the same method for (2) results in

$$\dot{h}_2(t) = \begin{cases} \dfrac{1}{A_{21}}\left(\dfrac{0.1k_v\sigma_{1P}\sqrt{g}}{2\sqrt{h_{1p}}}h_1(t) + \dfrac{0.1k_v\sigma_{1P}\sqrt{gh_{1p}}}{2} - q_{20}(t)\sigma_{2P}\right) & h_2(t) \le l_1 \\ \dfrac{1}{A_{22}}\left(\dfrac{0.1k_v\sigma_{1P}\sqrt{g}}{2\sqrt{h_{1p}}}h_1(t) + \dfrac{0.1k_v\sigma_{1P}\sqrt{gh_{1p}}}{2} - q_{20}(t)\sigma_{2P}\right) & h_2(t) \le l_1 \end{cases} \tag{14}$$

Steady state characteristics of (3) are unity gain for any nonzero h_1. Its linearization is therefore also very simple as all terms that include the difference ϑ_{0P}-ϑ_{1P} equal zero.

$$\dot{\vartheta}_1(t) = \left(q_{0P}/A_1h_{1P}\right)\left(\vartheta_{01}(t) - \vartheta_1(t)\right) \tag{15}$$

Using (7) this equation can be modified to

$$\dot{\vartheta}_1(t) = \left(0.1k_v\sigma_{1P}/A_1\right)\sqrt{\left(g/h_{1P}\right)}\left(\vartheta_{01}(t) - \vartheta_1(t)\right) \tag{16}$$

This equation has time constant

$$\tau = A_1\sqrt{h_{1P}}\,/\left(0.1k_v\sigma_{1P}\sqrt{g}\right) \tag{17}$$

Apart from multiplicative constant, this is the same expression as (12). Thus, the same partitioning of the range h_1 is obtained.

Equation (4) poses a more difficult problem. However, the effect of power output of the heater $H(t)$ on the dynamics of ϑ_2 is linear and the equation can be linearized around a point where $H(t)$=0. Then the steady state relation between ϑ_2 and ϑ_1 is again unity gain and the linearized equation is

$$\dot{\vartheta}_2(t) = \begin{cases} \dfrac{\left(0.1k_v\sigma_{1P}\sqrt{gh_{1P}}\left(\vartheta_1(t) - \vartheta_2(t)\right) + \dfrac{H(t)}{\rho c}\right)}{A_{21}h_{2P}} & h_2(t) \le l_1 \\ \dfrac{\left(0.1k_v\sigma_{1P}\sqrt{gh_{1P}}\left(\vartheta_1(t) - \vartheta_2(t)\right) + \dfrac{H(t)}{\rho c}\right)}{A_{21}l_1 + A_{22}(h_{2P} - l_1)} & h_2(t) > l_1 \end{cases} \tag{18}$$

Selection of nominal operating points is partly given by dynamics switch at level l_1. To achieve good approximation subranges $\langle h_{2\min}, l_1\rangle$ and $(l_1, h_{2\max}\rangle$ are further partitioned in a similar way as it was done with Tank 1. In this paper, the partitioning of both subranges into four intervals is used. Total number of partial models is then 384 (=2*3*2*4*8). The result is a continuous time PWA model with four states and seven inputs

$$\begin{aligned}
&\dot{\mathbf{x}}(t) = \mathbf{A}_\mathbf{i}\mathbf{x}(t) + \mathbf{B}_\mathbf{i}\mathbf{u}(t) + \mathbf{o}_\mathbf{i} \quad i = 1,...,384\\
&\mathbf{x}(t) = \left[h_1(t)\ \ h_2(t)\ \ \vartheta_1(t)\ \ \vartheta_2(t)\right]^T\\
&\mathbf{u}(t) = \left[\sigma_0(t)\ \ \sigma_1(t)\ \ \sigma_2(t)\ \ H(t)\ \ q_0(t)\ \ \vartheta_{01}(t)\ \ q_{20}(t)\right]^T
\end{aligned} \tag{19}$$

Most elements of matrices in (19) are zeros. The expressions for the few nonzero elements can be written according to (9)-(18). Partial models are valid in regions that are defined by specified values of logical inputs σ_0, σ_1, σ_2 and minimum and maximum limits on state variables. These specifications are formulated in the form of guard lines (6). Due to the high dimensions and great number of variants the expressions for matrices $\boldsymbol{G}_i^x$, $\boldsymbol{G}_i^u$, $\boldsymbol{G}_i^c$ cannot be given here.

System (19) must finally be converted to discrete time. This is done by discretizing each partial model separately assuming zero order hold at the inputs. Control Systems Toolbox for Matlab has no routine for discretization of affine systems. However it can be easily derived that the formulae for $\boldsymbol{N}_\mathrm{i}$ and $\boldsymbol{f}_\mathrm{i}$ in (5) have the same form

$$\boldsymbol{N}_i = e^{\boldsymbol{A}_iT_v}\int_0^{T_v} e^{-\boldsymbol{A}_i\tau}\boldsymbol{B}_i d\tau;\ \boldsymbol{f}_i = e^{\boldsymbol{A}_iT_v}\int_0^{T_v} e^{-\boldsymbol{A}_i\tau}\boldsymbol{o}_i d\tau \tag{20}$$

Hence the computation of discretized model is possible by using c2d command first with arguments ($\boldsymbol{A}_\mathrm{i}$, $\boldsymbol{B}_\mathrm{i}$, $\boldsymbol{C}_\mathrm{i}$, $\boldsymbol{D}_\mathrm{i}$) to obtain $\boldsymbol{N}_\mathrm{i}$ and then with arguments ($\boldsymbol{A}_\mathrm{i}$, $\boldsymbol{o}_\mathrm{i}$, $\boldsymbol{C}_\mathrm{i}$, 0) to obtain $\boldsymbol{f}_\mathrm{i}$.

The comparison of responses of the original model (1)–(4) and its PWA approximation is shown in Fig. 2. To evaluate the PWA model in a wide range of changes, the following step changes are used: q_0 changes at t=1000 s from 1 to 1.2 l/min, σ_1 changes at t=2000 s from 1 to 0 and at t= 2100 s back to 1, q_{20} changes at t=2400 s from 1 to 1.4 l/min and at t=3000 s to 1.2 l/min, ϑ_0 changes at t=3000 s from 40 to 50°C, H changes at t=4500 s from 0 to 500 W. Initial conditions are: h_{10}=0.6 m, h_{20}=0.12 m, ϑ_{10}=30°C, ϑ_{20}=55°C, sampling period T_s=1 s. The responses of PWA system were simulated using the Simulink block included in Multi-Parametric (MPT) Toolbox (Kvasnica *et al.*, 2004). There is generally a good agreement between the responses of the original system and its PWA approximation. Any comparison using a specified set of input signals has naturally a limited value because the agreement depends also on how close is the actual response to the selected set of representative operating points. However, it can be said that in most cases the maximum peak error does not exceed 1 cm or 1°C and normally the difference is in the range of tenths of cm and °C most of the time. It is also possible to decrease the number of partial models. Figure 2 was obtained with 384 partial models. If the range of h_2 is divided in just four sub-ranges, the number of models is reduced to 192 and the precision of PWA approximation remains good. However, any further reduction of the number of partial models results in a marked decrease of approximation precision.

6. CONTROL DESIGN AND EXPERIMENTS

Standard procedure to design a control system satisfying the objectives specified above is divided into two separate tasks: design of logical control and design of continuous control (in this paper the term continuous relates to control where the variables are continuous-valued regardless of whether the controller is designed in continuous or discrete time setting). Logic part of the control system is described by a set of simple rules. Normal and desirable state of the logical control inputs is σ_0=1; σ_1=1, σ_2=1. That means, water from the upstream process flows to the buffer (Tank 1) and further to the supply (Tank 2) and the supply is able to meet the demand of the downstream process while water levels of both tanks remain within acceptable ranges $w_{1\min} \le h_1 \le w_{1\max}$, $w_{2\min} \le h_2 \le w_{2\max}$. If water levels deviate from these ranges, the following rules apply

A. If $h_1 < w_{1\min}$, set σ_1 to zero to avoid Tank 1 emptying
B. If $h_1 > w_{1\max}$, set σ_0 to zero to avoid Tank 1 overflow, if also $h_2 < w_{1\min}$, set σ_1 to 2 to accelerate the recovery of both water levels to normal ranges
C. If $h_2 < w_{2\min}$, set σ_2 to zero to avoid Tank 2 emptying
D. If $h_2 > w_{2\max}$, set σ_1 to zero to avoid Tank 2 overflow

Continuous control system is designed as SISO control loop, where $H(t)$ is manipulated variable and temperature ϑ_2 is a controlled variable. It can be seen from (4) that logical control inputs act as disturbances and they may have adverse effects on control performance.

The situation changes when logical and continuous control is designed in a unified way. System described by (1)-(4) is treated as a hybrid system and model predictive controller can be designed based on its PWA approximation. Hybrid MPC controller is designed using MPT Toolbox. To compare performance of the systems using separate design and hybrid model predictive control, the following control experiment was performed. Starting from the state q_0=q_{20}=1 l/min, h_1=0.5 m, h_2=0.3 m, ϑ_1=50°C, ϑ_2=40°C, the setpoint was increased from 40°C to 60°C. Separate design used logic rules defined in the beginning of this section. The continuous control algorithm was MPC with linear performance function and control horizon 2. The unified design used MPC with the same performance function and control horizon, however this MPC algorithm could make use of the logical control inputs. The responses are shown in the following figures.

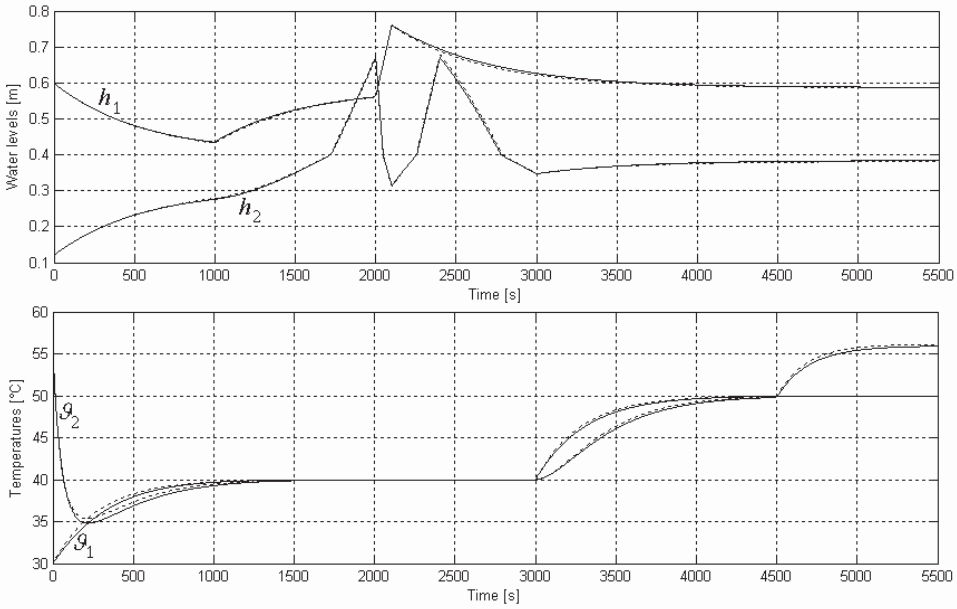


Fig. 2 Comparison of responses of the original plant model and its PWA approximation (in all responses original plant models is plotted with solid line and PWA approximation with dotted line).

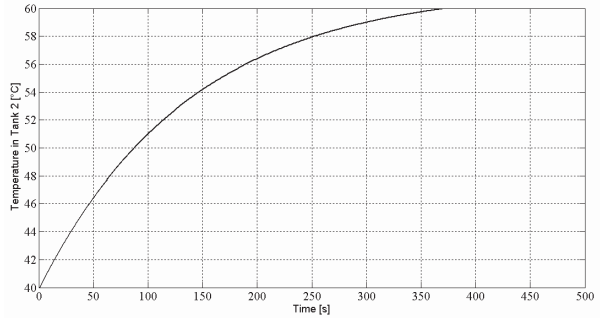


Fig. 3 Setpoint response Separate design

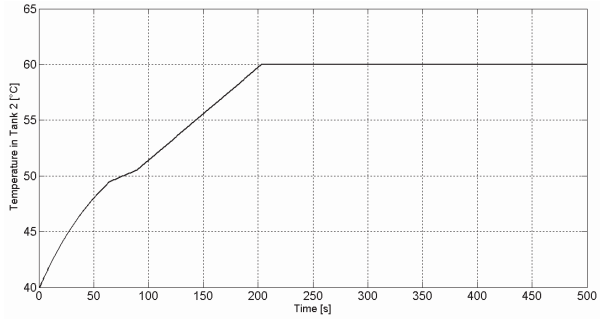


Fig. 4 Setpoint response – Unified design

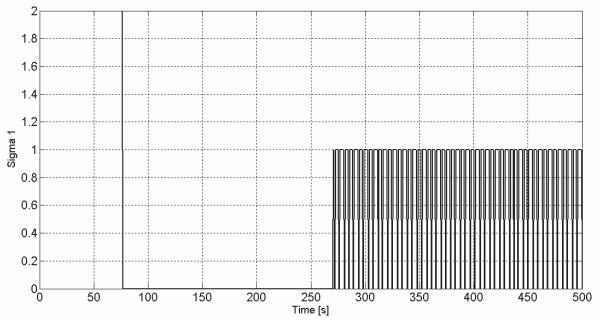


Fig. 5 Logical manipulated variable σ_1

It can be seen that unified design achieves better results. The control time is much shorter. Fig. 5 shows that this improvement is due to the ability of the unified design to make use of logical manipulated variables. Unlike separate design, logical control inputs can be used not only to keep the water levels within specified limits but also to accelerate control responses. In the beginning σ_1 is used to accelerate the setpoint response by increasing the inflow of warmer water to Tank 2. When controlled variable reaches setpoint, σ_1 is used just to keep the water levels within specified range.

7. CONCLUSION

This paper was focused on the possibilities that hybrid model predictive control can offer for unified design of logical and continuous control. As the character of this paper is mainly experimental, its results are connected with the particular laboratory plant considered and they cannot be regarded as completely general. In spite of that, some conclusions can be made. It could be seen that the application of hybrid MPC is not particularly easy. The development of the PWA model that is necessary for hybrid MPC takes up a substantial part of the paper. The complexity of the PWA model expressed by several hundreds of partial systems is also quite high even in the case of this laboratory scale plant that has still quite simple structure compared with real industrial process plants.

On the other hand, it has been shown that hybrid MPC can make use of the information how the controlled variable is affected by logical control inputs to improve control responses. The setpoint response was improved by adding the effect of opening valve S4 to the effect of increasing heater power output. Thus, the control results achieved with hybrid MPC were better than the results obtained with separate design of logical and continuous control.

Given paper length allowed to present one selected control experiment. Other control scenarios can be devised and tested with similar results and there is still a large open space for further experiments with this plant focused on exploring the possibilities offered by hybrid MPC for co-ordinated design of logical and continuous control. A particular attention will also be paid to the real time implementation of explicit MPC using industrial control system WinCon-8000 that can be used with this plant as an alternative to the academic experimental setting using Matlab/Simulink, Real Time Toolbox and PC data acquisition boards.

ACKNOWLEDGEMENT:

This research has been supported by the Czech Science Foundation within project 101/07/1667.

REFERENCES

Borrelli, F. (2003), *Constrained Optimal Control of Linear and Hybrid System,* Springer Verlag.

Christophersen, F. (2007), *Optimal Control of Constrained Piecewise Affine Systems*, Springer Verlag.

Hlava, J., & Šulc B. (2008), Advanced Modelling and Control using a Laboratory Plant with Hybrid Processes, *Proceedings of the 17th IFAC World Congress*, Seoul Korea July 2008, pp. 14636-14641.

Kvasnica, M., Grieder, P., Baotic M. & Morari M.(2004), Multi-Parametric Toolbox (MPT), *HSCC 2004 (Hybrid Systems: Computation and Control)*, pp.448-462, Springer Verlag.

Nandola, N. N., & Bhartiya, S. (2008), A multiple model approach for predictive control of nonlinear hybrid systems, *Journal of Process Control*, vol. 18, No. 2, pp.131-148.

Morari M., Baric M. (2006), Recent developments in the control of constrained hybrid systems, *Computers & Chemical Engineering*, vol. 30, pp. 1619–1631.

Paoletti, S., Juloski, A.Lj., Ferrari-Trecate G., & Vidal R. (2007), Identification of hybrid systems: a tutorial. *European Journal of Control*, vol. 13, pp.242-260.

Slupphaug, O., Vada J. & Foss B. A. (1997), MPC in Systems with Continuous and Discrete Control Inputs, *Proceedings of ACC 97*, Albuquerque USA June 1997.

Sontag, E. (1981), Nonlinear regulation: The piecewise linear approach, *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 346-358.

Nonlinear and Adaptive Control

Oral Session

Thermodynamic approach for Lyapunov based control

H. Hoang * **F. Couenne** * **C. Jallut** * **Y. Le Gorrec** **

* *LAGEP, University of Lyon, University of Lyon 1, UMR CNRS 5007, Villeurbanne, France (e-mails: {hoang;jallut;couenne}@lagep.univ-lyon1.fr)*
** *FEMTO-ST / AS2M, ENSMM Besançon, Besançon, France (e-mail: yann.le.gorrec@ens2m.fr)*

Abstract: This paper focuses on non linear control of non isothermal Continuous Stirred Tank Reactors (CSTRs). The model of the CSTR is thermodynamically consistent in order to apply the control strategy based on the concavity of the entropy function and the use of thermodynamic availability as Lyapunov function. More precisely the stabilization problem of continuous chemical reactors is addressed operated at an unstable open loop equilibrium point. The chosen control variable is the jacket temperature. In this paper we propose a state feedback strategy to insure asymptotic stability with physically admissible control variable solicitations. Theoretical developments are illustrated on a first order chemical reaction.

Keywords: Lyapunov based control, Irreversible thermodynamics, Non isothermal CSTR, Multiple steady states, Entropy.

1. INTRODUCTION

Continuous Stirred Tank Reactors (CSTR) have been widely studied in the literature with respect to process control design (Luyben (1990); Alvarez (1999); Hua (2000); Guo (2001); Hoang (2008)). Numerous strategies have been developed to control such non linear systems. Let us cite for example: feedback linearization (Viel (1997)) for control under constraints, nonlinear PI control (Alvarez (1999)), classical Lyapunov based control (Antonellia (2003)), nonlinear adaptive control (Guo (2001)) and more recently thermodynamical Lyapunov based control (Hoang (2008)).

Besides these control problems, observation/estimation strategies have been developed in the case of under sensored CSTRs (Gibon-Fargeot (2000); Dochain (2009)). Usually, the reactor temperature is the only on-line available measurement. Then the purpose is to estimate the missing state variables that are used in the control strategy.

In this paper we focus on the control purposes only and we assume that concentrations and temperature are measured. This control synthesis is based on thermodynamic concepts defined in Callen (1985) and more recently in (Ruszkowski (2005); Ydstie (1997)) and (Hoang (2008)). More precisely, we propose a Lyapunov based approach for the stabilization of CSTR about unstable steady state as in (Hoang (2008)). This is done thanks to the Lyapunov function issued from thermodynamics consideration: the availability function $\mathcal{A}$ (Ruszkowski (2005)).

In Hoang (2008), we proposed feedback laws involving inlet and jacket temperatures as well as inlet flows. These feedback laws were obtained by imposing that the time derivative of the availability $\mathcal{A}$ remains negative, insuring consequently the global asymptotic stability. However, no care was given on the amplitude of the controls. Moreover the temperature of the reactor had to be inverted and the feedback laws had in some case some oscillatory behaviors about the critical point.

The main contribution of this paper with respect to previous work (Hoang (2008)) is the redesign of the exponential asymptotic controller in order to prevent excessive control demand and oscillation problems. In this way the obtained controller is practically more efficient. The price to pay is that global asymptotic stability is obtained on some validity domain only.

This paper is organized as follows: in section 2, we remind thermodynamical concepts and variables necessary to construct thermodynamic availability. This latter function is the Lyapunov candidate of the method. In section 3 the dynamic model of the considered CSTR is presented and analyzed. Section 4 is devoted to the design of the state feedback insuring asymptotic stability. Simulation results are given in section 5. It is shown that the resulting control leads to admissible manipulated control variables.

2. THERMODYNAMIC BASIS FOR AN AVAILABILITY FUNCTION

Irreversible thermodynamics concept will play a leading role in the methodology used for the design of the Lyapunov function (Ruszkowski (2005); Hoang (2008)). In this section we review the main ideas concerning this thermodynamical approach and the construction of the candidate Lyapunov function: the availability function in the case of an homogeneous phase.

In equilibrium thermodynamics, the system variables are divided into extensive and intensive variables, depending on whether their values depend on the "size" of the system or not. The internal energy of a homogeneous system is then expressed in terms of products of pairings of energy conjugate variables such as pressure P/ volume V, temperature T/ entropy S and chemical potential μ_i/ mole number n_i for each species i of the mixture.

The fundamental relation of thermodynamics expresses the entropy S of a given phase as a function of the so called extensive variables $Z = (U, V, n_i)$ by the Gibbs equation:

$$dS = \frac{1}{T}dU + \frac{P}{T}dV + \sum_{i=1}^{n_c} \frac{-\mu_i}{T} dn_i. \quad (1)$$

It can also be written as:

$$dS = w^T dZ \quad (2)$$

with $w = (\frac{1}{T}, \frac{P}{T}, \frac{-\mu_i}{T})$.

Since the entropy S is an extensive variable, it is a homogenous function of degree 1 of Z (Callen (1985)). From Euler's theorem we get:

$$S(Z) = w^T Z \quad (3)$$

Equation (2) can also be applied in irreversible thermodynamics as soon as the local state equilibrium is assumed: it postulates that the present state of the homogeneous system in any evolution can be characterized by the same variables as at equilibrium and is independent on the rate of evolution. So (2) can also be applied at any time.

Moreover, it is well known that balance equations can be established for $Z = (U, V, n_i)$ as well as for the entropy S but this latter is not conservative: in irreversible thermodynamics there is a source term σ which is always positive from the the second law of thermodynamics. This term represents the irreversible entropy production: the energy $T\sigma$ associated to this term represents the energy lost from material, space or thermal domains and that will never more contribute to some physical works. As a consequence of (2), the entropy balance can alternatively be written as:

$$\frac{dS}{dt} = w^T \frac{dZ}{dt} \quad (4)$$

Finally let us notice that for homogeneous thermodynamical systems (one phase only), the entropy function $S(Z)$ is necessarily strictly concave (see Callen (1985)) as shown in Fig. 1.

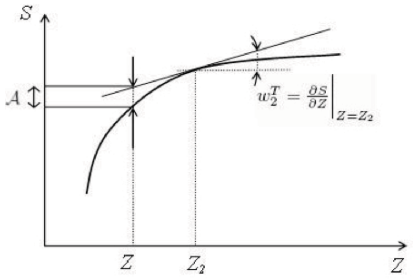


Fig. 1. Entropy and availability functions w. r. to Z.

From these observations, it can be shown (see Ydstie (1997)) that the non negative function:

$$\mathcal{A}(Z) = S_2 + w_2^T(Z - Z_2) - S(Z) \geq 0 \quad (5)$$

where Z_2 is some fixed reference point (for example the desired set point for control), is a measure of the distance between entropy $S(Z)$ and its tangent plane passing through Z_2. It is geometrically presented in Fig. 1. The slope of the tangent plane is related to intensive vector $w(Z)$ calculated at $Z = Z_2$.

As soon as we consider homogeneous mixture, S remains concave and then $\mathcal{A}$ remains also non negative. As a consequence, A is a natural Lyapunov candidate. It remains to build a feedback law to insure:

$$\frac{d\mathcal{A}}{dt} \leq 0. \quad (6)$$

3. CASE STUDY: A NON ISOTHERMAL CSTR MODEL

3.1 Assumptions of the model

We consider a jacketed homogeneous CSTR with the following first-order chemical reaction: $A \rightarrow B$. The temperature of the jacket T_w is supposed to be uniform and is used for the control purpose. The dynamics of the CSTR is deduced from volume, material and energy balances.

The following assumptions are made:

- The fluid is incompressible and the reaction mixture is supposed to be ideal.
- The two species are supposed to have the same partial molar volume v.
- At the inlet of the reactor, the pure component A is fed at temperature T_e.
- The reaction volume V is supposed to be constant.
- The heat flow exchanged with the jacket is represented by $\dot{Q} = \lambda(T_w - T)$.
- The kinetics of the liquid phase reaction is modelled thanks to the Arrhenius law. The reaction rate r_v is given by $k_0 \exp(\frac{-k_1}{T})\frac{n_A}{V}$.

In Tables (1,2) are given the notations and numerical values that will be used for modelling and simulation. Finally let us notice that constant volume assumption implies that the total number of moles n_T is constant since the two species have the same partial molar volume.

Notation	unit	
F_{Ae}	mol/s	Inlet molar flow rate of A
F_A	mol/s	Outlet molar flow rate of A
F_B	mol/s	Outlet molar flow rate of B
F	mol/s	Total outlet molar flow rate
h_{Ae}	J/mol	Inlet molar enthalpy of A
h_i	J/mol	Molar enthalpy of species i $(i = A, B)$
H	J	Total enthalpy of the mixture
n_A	mol	Mole number of species A
n_B	mol	Mole number of species B
T	K	Temperature in the CSTR
n_T	mol	Total mole number
r_v	$mol/m^3/s$	Reaction rate
U	J	Internal energy
$x_i = \frac{n_i}{n_T}$		Molar fraction of species i, $i = A, B$

Table 1. Notation of the variables of the model.

	Numerical value	
C_{pA}	75.24 $(J/K/mol)$	Heat capacity of species A
C_{pB}	60 $(J/K/mol)$	Heat capacity of species B
h_{Aref}	0 (J/mol)	Reference enthalpy of A
h_{Bref}	-4575 (J/mol)	Reference enthalpy of B
k_0	0.12 10^{10} $(1/s)$	Kinetics constant
k_1	8.7 10^3 (K)	Parameter in Arrhenius law
P	10^5 (Pa)	Pressure
T_{ref}	300 (K)	Reference temperature
v	0.0005 (m^3/mol)	Molar volume
V	0.001 (m^3)	Reaction volume
λ	0.05808 (W/K)	Heat transfer coefficient
s_{Aref}	210.4 $(J/K/mol)$	Reference entropy of A
s_{Bref}	180.2 $(J/K/mol)$	Reference entropy of B

Table 2. Parameters of the CSTR.

Moreover the constant volume assumption constrains the total outlet molar flow rate F.

3.2 CSTR modelling

The material balances are given by:

$$\begin{cases} \dfrac{dn_A}{dt} = F_{Ae} - F_A - r_{\mathrm{v}} V \\ \dfrac{dn_B}{dt} = -F_B + r_{\mathrm{v}} V \end{cases} \tag{7}$$

and the energy balance by:

$$\frac{dU}{dt} = \dot{Q} - P\frac{dV}{dt} + F_{Ae}h_{Ae} - (F_A h_A + F_B h_B) \tag{8}$$

Remark 1. Since we suppose ideality of the mixture, the enthalpy of species A_i, $i = A, B$ in the mixture can be expressed as: $h_i(T) = c_{pAi}(T - T_{ref}) + h_{iref}$. Let us furthermore note that, as the species are involved in a chemical reaction, the reference molar enthalpies are chosen with regard to the enthalpy of formation of species.

Finally the volume balance leads to:

$$\frac{dV}{dt} = 0 \tag{9}$$

Since molar volume of species are assumed to be equal, it implies that $F = F_{Ae}$ and $F_A = x_A\, F_{Ae}$ and $F_B = x_B\, F_{Ae}$

The internal energy balance can be written in term of temperature. This is done by using the expression of the enthalpy of the system $H = \sum_{i=A,B} n_i h_i$ and by noticing that under our assumptions $\frac{dU}{dt} = \frac{dH}{dt}$. We finally obtain:

$$C_p\frac{dT}{dt} = \Big(-\Delta H\Big) r_{\mathrm{v}} V + F_{Ae}C_{pA}(T_e - T) + \lambda(T_w - T) \tag{10}$$

where $\Delta H = (h_B - h_A)$ is the enthalpy of the reaction and $C_p = C_{pA}n_A + C_{pB}n_B$ is the total heat capacity.

The dynamics of states variables (H, n_A) ((8) and (7)) or (T, n_A) ((10) and (7)) give two equivalent representations of the CSTR. These representations will be used for late purpose.

3.3 Analysis of the steady states

For this purpose, manipulated variables are chosen as:

$$F_{Ae} = 0.0183\ (mol/s),\ T_e = 310\ (K)\ T_w = 300\ (K) \tag{11}$$

Steady states are calculated by setting (7) and (10) equal to zero.

By introducing the expression of the steady state mole number of n_A in the temperature equation, the steady state temperatures are the values that satisfy $P_e(T) = 0$ with:

$$P_e(T) = \frac{h_A - h_B}{C_p} k_0 \exp(\frac{-k_1}{T}) \frac{F_{Ae}}{\Big(\frac{F_{Ae}}{n_T} + k_0 \exp(\frac{-k_1}{T})\Big)} + \frac{F_{Ae}C_{pA}}{C_p}(T_e - T) + \frac{\lambda}{C_p}(T_w - T) \tag{12}$$

These values are represented in Fig. 2(a). It shows that the system has three steady state operating points: P_1, P_2 and P_3.

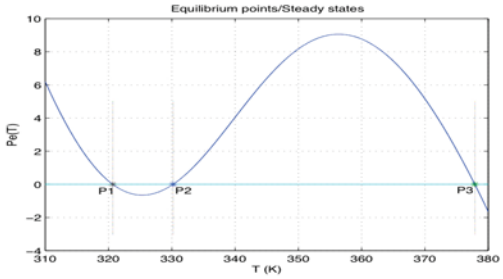


Fig. 2. Steady states

The numerical values of these steady states and the eigenvalues of the linearized system about these points are given in Table 3.

Points	Values	Eigenvalues
P_1: $[n_A\ T]$	[1.6449 320.6704]	[−0.0090 −0.0024]
P_2: $[n_A\ T]$	[1.3583 330.1997]	[−0.0090 0.0027]
P_3: $[n_A\ T]$	[0.1416 377.8795]	[−0.0802 −0.0100]

Table 3. Steady state points and eigenvalues

From Table 3, one can see that steady state operating points P_1 and P_3 are stable, whereas the steady state operating point P_2 is not stable since one of its eigenvalues is positive.

Control Problem: we are interested to operate the reactor at $T = 330.1997$ corresponding to the unstable steady state operating point P_2 and at fixed F_{Ae} and T_e. As a consequence a control feedback law on T_w is necessary.

4. CONTROLLER SYNTHESIS

In this paper we propose a feedback law that is less conservative than the one proposed in Hoang (2008) and that still insures asymptotic stability in some admissible domain. We first give some preliminary results necessary for the controller synthesis.

Proposition 1 shows that n_A belongs to an invariant domain $[0, n_T]$.

Proposition 1. If $n_A(0) \in [0, n_T]$ then $n_A(t) \in [0, n_T]\ \forall t$

Proof. It is straightforward looking at (7) since $\frac{dn_A}{dt}\Big|_{n_A=0} = F_{Ae} > 0$ and $\frac{dn_A}{dt}\Big|_{n_A=n_T} = -k_0 \exp(-\frac{k_1}{T}) n_T < 0$ □

Moreover we notice that the sign of $\frac{dn_A}{dt}$ is the same as that of $G(T) = \frac{F_{Ae}}{\frac{F_{Ae}}{n_T} + k_0 \exp(\frac{-k_1}{T})}$.

In order to stabilize the closed loop system about (n_{a_2}, T_2), we propose the following feedback law for T_w.

Proposition 2. At fixed T_e and F_{Ae}, the system defined by ((7) and (8)) with the non linear feedback law (13) for T_w:

$$T_w = \frac{1}{\lambda}\Big(K_1\widetilde{v}_1 - \mathfrak{F}F_{Ae} + \frac{f}{-\widetilde{v}_1}\frac{dn_A}{dt}\Big) + T \tag{13}$$

where:

$$\widetilde{v}_1 = \left[\frac{1}{T} - \frac{1}{T_2}\right] \tag{14}$$

$$\mathfrak{F}(T_e, T, n_A, n_B) = \Big(h_{Ae} - (x_A h_A + x_B h_B)\Big) \tag{15}$$

and

$$\begin{aligned} f(T) = &\Big((C_{pA} - C_{pB})T_{ref} - (h_{Aref} - h_{Bref})\Big)\widetilde{v}_1 \\ &+ \Big(C_{pA} - C_{pB}\Big)\ln\Big(\frac{T}{T_2}\Big) \end{aligned} \tag{16}$$

is stable and asymptotically converges to the desired operating point $P_2 = (T_2, n_{A2})$ for any initial condition (T_0, n_{A0}) contained in some validity domain for which the constant K_1 is chosen positive.

Proof. K_1 insures the continuity of T_w at $t = 0$: $T_w(0) = T_0$ or,

$$\left[K_1\widetilde{v}_1 - \mathfrak{F}F_{Ae} + \frac{f}{-\widetilde{v}_1}\frac{dn_A}{dt}\right]_{t=0} = 0 \tag{17}$$

The proof of the proposition 2 contains two parts:

1. Determination of the validity domain of initial conditions: developing (17) and using the material balance (7) and since $n_B = n_T - n_A$, we have at $t = 0$:

$$K_1\widetilde{v}_1 = F_{Ae}h_{Ae} - F_{Ae}h_B + \frac{f}{\widetilde{v}_1}F_{Ae} - n_A D(T) \tag{18}$$

with $D(T) = \left[\frac{F_{Ae}}{n_T}(h_A - h_B) + \Big(\frac{F_{Ae}}{n_T}F_{Ae} + k_0\exp(\frac{-k_1}{T})\Big)\frac{f}{\widetilde{v}_1}\right]$.

For positive K_1, (18) is positive if $\widetilde{v}_1 > 0$. So the right hand side of the equality has the sign of $\widetilde{v}_1$.

In a same way, we obtain :

$$\begin{cases} n_{A0} < F(T_0) \text{ if } T_0 > T_2 \\ n_{A0} > F(T_0) \text{ if } T_0 < T_2 \end{cases} \tag{19}$$

with $F(T) = \frac{(h_{Ae} - h_B + \frac{f}{\widetilde{v}_1})}{\frac{1}{n_T}(h_A - h_B) + \frac{1}{G(T)}\frac{f}{\widetilde{v}_1}}$.

The domain of validity is given in Fig. 3.

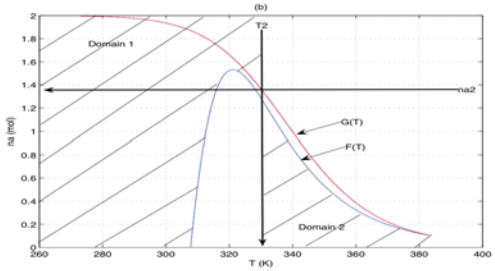


Fig. 3. Domain of validity of initial conditions

2. Stability and convergence to the desired point (T_2, n_{A2}): Let us consider the function $\mathcal{A}$ (5). The time derivative of such function can be written:

$$\frac{d\mathcal{A}}{dt} = -\widetilde{v}_1\frac{dU}{dt} - \widetilde{v}_2\frac{dn_A}{dt} \tag{20}$$

with $\widetilde{v}_2 = -\big(\frac{\mu_A}{T} - \frac{\mu_B}{T}\big) + \Big(\frac{\mu_{A2}}{T_2} - \frac{\mu_{B2}}{T_2}\Big)$. From the energy balance (8), (20) can be written:

$$\frac{d\mathcal{A}}{dt} = -\widetilde{v}_1\Big[F_{Ae}\mathfrak{F} + \lambda(T_w - T)\Big] - \widetilde{v}_2\frac{dn_A}{dt} \tag{21}$$

where $\mathfrak{F}$ is defined in (14). Furthermore, using the constitutive equation

$$\mu_A(T, P, x_A) = \mu_A^0(T) + RT\ln(\frac{n_A}{n_A + n_B}) \tag{22}$$

where $\mu_A^0(T) = C_{pA}(T - T_{ref}) + h_{Aref} - T\Big(C_{pA}ln(\frac{T}{T_{ref}}) + s_{Aref}\Big)$ one can write $\widetilde{v}_2$ on the following form:

$$\widetilde{v}_2 = f(T) + g(n_A) \tag{23}$$

where $f(T)$ is defined in (16) and $g(n_A) = R\ln\Big(\frac{n_{A2}}{n_A}\frac{n_B}{n_{B2}}\Big)$. Then (21) becomes :

$$\frac{d\mathcal{A}}{dt} = -\widetilde{v}_1\Big[F_{Ae}\mathfrak{F} + \lambda(T_w - T)\Big] - (f + g)\frac{dn_A}{dt} \tag{24}$$

We propose the following feedback law :

$$T_w = \frac{1}{\lambda}\Big(K_1\widetilde{v}_1 - \mathfrak{F}F_{Ae} + \frac{f}{-\widetilde{v}_1}\frac{dn_A}{dt}\Big) + T \tag{25}$$

for systems with initial conditions $(T_w(0) = T(0))$ such that $K_1 > 0$. Using this feedback law, $\frac{d\mathcal{A}}{dt}$ becomes:

$$\frac{d\mathcal{A}}{dt} = -K_1\widetilde{v}_1^2 - g\frac{dn_A}{dt} \tag{26}$$

The idea is to not constrain the system by imposing $\frac{d\mathcal{A}}{dt} < 0\ \forall t$ as in Hoang (2008).

We are now going to show that depending on the initial conditions from the domain of validity (associated with condition $K_1 > 0$), $-g\frac{dn_A}{dt}$ is either negative $\forall t$ or becomes negative and converges to 0.

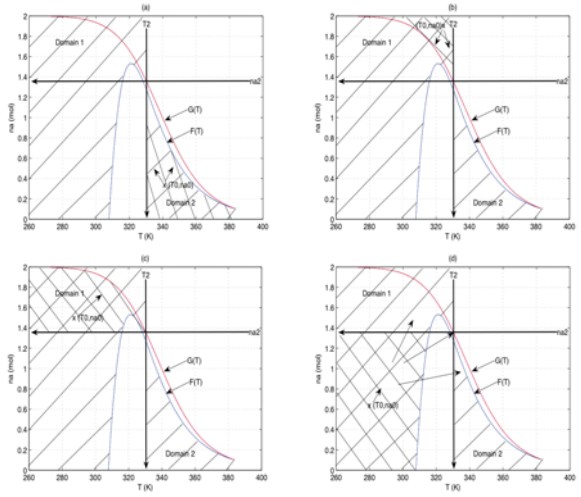


Fig. 4. Admissible initial conditions in the domain of validity.

Remark 2. A simple analysis permits to conclude that $g(n_A)$ is positive as soon as $n_A \leq n_{A2}$.

In all cases in using (19), lemma 1 and remark 2 we will show the negativeness of $-g\frac{dn_A}{dt}$.

With initial conditions such as shown in Fig. 4(a) and using additionally the remarks 4.5 and 4.6 of appendix A), we have:

$$\frac{d\mathcal{A}}{dt} = -K_1\widetilde{v}_1^2 - g\frac{dn_A}{dt} \leq 0, \ \forall t \tag{27}$$

With initial conditions such as shown in Fig. 4(b), using the remarks 4.3 and 4.4 we obtain the same inequality (27).

The trajectory of (T, n_A) issued from initial domain as shown in Fig. 4(c) is trapped in the domain of Fig. 4(b). This is obtained thanks to remarks 4.2 and 4.3 and 4.4.

Finally for initial conditions as shown in Fig. 4(d), there are two possible scenarios : one is that the trajectory of (T, n_A) is trapped in the domain of figure 4(a) or 4(c) then 4(b). The result then follows from remarks 4.2 and 4.3 and 4.4. The other scenario is that the trajectory of (T, n_A) is not trapped in these domains and then $\mathcal{A}$ always decreases and converges to 0.

Finally, from all the admissible initial conditions and after some time, $\mathcal{A}$ plays the role of a Lyapunov function.

Remark 3. The feedback law T_w (13) is well defined for $T = T_2$ since $\lim_{T\to T_2} \frac{f}{v_1} = \Big((C_{pA} - C_{pB})T_{ref} - (h_{Aref} - h_{Bref})\Big) + \Big(C_{pA} - C_{pB}\Big)(-T_2)$.

5. SIMULATION

The purpose of this section is to illustrate the good performances obtained from the aforementionned control strategy and the admissibility of the resulting control variables. The open and closed loop simulations are carried out respect to four different initial conditions chosen in the initial domain of validity of the control law. These initial conditions correspond to the four different scenarios depicted in Fig. 4 in view of studying the convergence properties of the control law and the control variable solicitation. The four initial conditions are:

(C1): $(T(0) = 340, n_{A0} = 0.6)$ belongs to Fig. 4(a).
(C2): $(T(0) = 325, n_{A0} = 1.8)$ belongs to Fig. 4(b).
(C3): $(T(0) = 300, n_{A0} = 1.6)$ belongs to Fig. 4(c).
(C4): $(T(0) = 300, n_{A0} = 0.6)$ belongs to Fig. 4(d).

5.1 Open loop simulation

First of all let us consider open loop simulations with inputs defined by (11) and initial conditions (C1) to (C4). Simulations are given in Figure (5).

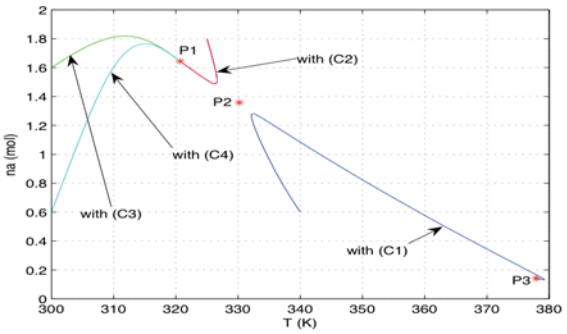


Fig. 5. The representation of the open loop phase plan

5.2 Closed loop system

The open loop system is closed with the feedback law T_w constructed with the state variables n_A and T.

The trajectories issued from the initial points (C1) to (C4) are given in Fig. 6. We notice that for all the initial conditions the system converges to the desired operating point P2.

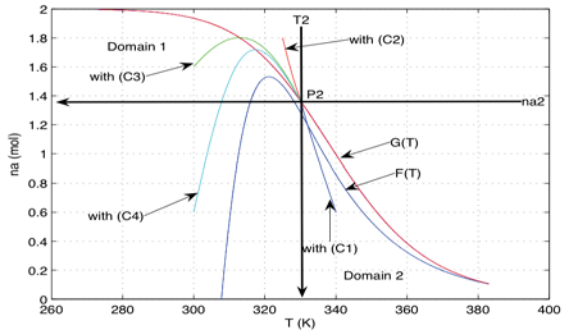


Fig. 6. Closed loop trajectories in phase plane.

Fig. 7 shows the control variable T_w. Its values are admissible and its evolution is slow enough.

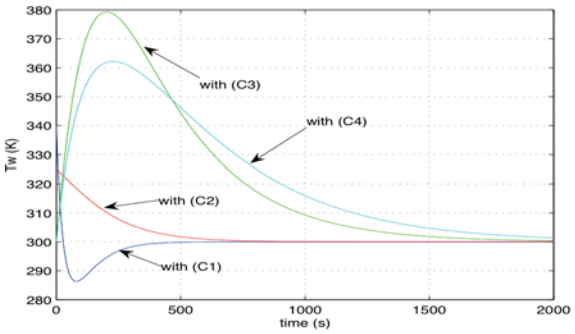


Fig. 7. The feedback law T_w

Fig.8 shows the time trajectory of $\mathcal{A}$ for the different initial conditions. For initial conditions (C1) and (C2),

the availability $\mathcal{A}$ can be assimilated to Lyapunov function from the beginning of the reaction. For initial conditions (C3) and (C4), $\frac{d\mathcal{A}}{dt}$ is forced to be negative only after a certain time from which $\mathcal{A}$ plays the role of Lyapunov function, and converges to 0.

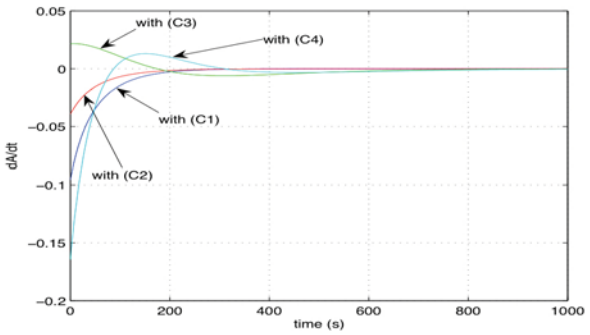


Fig. 8. The dynamics of $\frac{d\mathcal{A}}{dt}$

6. CONCLUSION

In this paper, we have shown how to stabilize a CSTR about the desired operating point by means of Lyapunov-based method. The Lyapunov function is the availability function $\mathcal{A}$. $\mathcal{A}$ is derived from thermodynamic considerations. The stabilization is ensured in some domain of validity issued from the condition of positivity of the design parameter K_1 and the continuity of the feedback law T_w.

The simulation results showed that convergence objective is satisfied and that the state feedback law is physically implementable since jacket temperature remains in some physical domain with admissible rate of variation.

Nevertheless,in the proposed control strategy the closed loop dynamic is imposed by the initial conditions (with K_1). This is the reason why we are now studying for dynamic controllers with additional freedom degrees. It remains also to compare our result with previous results as given in Viel (1997) for example) in term of performance and robustness.

REFERENCES

Alvarez-Ramirez, J. and Femat, R. (1999). Robust PI stabilization of a class of chemical reactors. *Systems & Control Letters*, 38(4-5), 219-225.

Antonellia, R., and Astolfi, A. (2003). Continuous stirred tank reactors: easy to stabilise? *Automatica*, 39, 1817-1827.

Callen, H.B. (1985). *Thermodynamics and an introduction to thermostatics.* John Wiley & Sons Inc, New York, 2nd edition.

Dochain, D., Couenne, F.,and Jallut, C. (2009) Enthalpy Based Modelling and Design of Asymptotic Observers for Chemical Reactors. *accepted to International Journal of Control.*

Gibon-Fargeot, A.M., Celle-Couenne, F. , and Hammouri, H. (2000). Cascade estimation design for cstr models. *Computers & Chemical Engineering*, 24(11), 2355-2366.

Guo, B., Jiang, A., Hua, X., and Jutan, A. (2001). Nonlinear adaptive control for multivariable chemical processes. *Chemical Engineering Science*, 56, 67816791.

Hoang, H., Couenne, F.,Jallut, C., and Le Gorrec, Y. (2008). Lyapunov based control for non isothermal continuous stirred tank reactor. *Proceedings of the 17th World Congress of the IFAC*, July 6-11, 2008, Seoul, Korea.

Hua, X., and Jutan A. (2000). Nonlinear Inferential Cascade Control of Exothermic Fixed-bed Reactors. *AICHE Journal*, 46, 980-996, 2000.

Luyben, W.L. (1990). *Process Modeling, Simulation, and Control for Chemical Engineers.* McGraw-Hill, Singapore.

Ruszkowski, M., Garcia-Osorio, V., and Ydstie, B.E. (2005) Passivity based control of transport reaction systems. *AIChE Journal*, 51, 3147-3166.

Viel, F., Jadot, F., and Bastin, G. (1997). Global stabilization of exothermic chemical reactors under input constraints. *Automatica*, 33(8), 1437-1448.

Ydstie, B.E., and Alonso, A.A. (1997). Process systems and passivity via the Clausius-Planck inequality. *Systems Control Letters*, 30(5), 253-264.

Appendix A

Lemma 1. The energy balance (8) with feedback law (13) gives rise to:

1.
$$(\sum_i n_i C_{pi})\frac{dT}{dt} = K_1\widetilde{v}_1 + L(T)\frac{dn_A}{dt} \qquad (A.1)$$
with $L(T) = \left(-\frac{f}{v_1} - (h_A - h_B)\right)$.

2. With assumptions presented in section 3.1, we have: $L(T) > 0$ if $T < T_2$ and $lim_{T\to T_2} L(T) = 0$.

Remark 4. The following remarks hold:

(1) From Proposition 1, C_p is bounded and positive.
(2) Lemma 1 insures that if $T < T_2$ and $\frac{dn_A}{dt} > 0$ then $\frac{dT}{dt} > 0$ since $C_p\frac{dT}{dt} = K_1\underbrace{\widetilde{v}_1}_{>0} + \underbrace{L(T)\frac{dn_A}{dt}}_{>0}$.
(3) When $\frac{dn_A}{dt} = 0$ (n_A reaches $G(T)$) and $T < T_2$ then $(\sum_i n_i C_{pi})\frac{dT}{dt} = K_1\underbrace{\left(\frac{1}{T} - \frac{1}{T_2}\right)}_{>0}$ and $\frac{dT}{dt}$ remains positive.
(4) When $\frac{dn_A}{dt} < 0$ and $T = T_2$, then $C_p\frac{dT}{dt} = 0$ and T stays equal to T_2.
(5) When $\frac{dn_A}{dt} = 0$ (n_A reaches $G(T)$) and $T > T_2$, then $(\sum_i n_i C_{pi})\frac{dT}{dt} = K_1\underbrace{\left(\frac{1}{T} - \frac{1}{T_2}\right)}_{<0}$ and T decreases.
(6) When $\frac{dn_A}{dt} > 0$ and $T = T_2$, then $(\sum_i n_i C_{pi})\frac{dT}{dt} = 0$ and T remains equal to T_2.
(7) When $\frac{dn_A}{dt} = 0$ and $T = T_2$, the system reaches the desired point and stays on.

Boundary geometric control of co-current heat exchanger

Ahmed MAIDI [*,**] **Moussa DIAF** [*] **Jean Pierre CORRIOU** [**]

[*] *Université Mouloud MAMMERI, Faculté de Génie Electrique et d'Informatique, Département Automatique, Tizi-Ouzou, Algérie*
[**] *Laboratoire des Sciences de Génie Chimique, CNRS-ENSIC, 1 rue Grandville BP 20451, 54001 Nancy Cedex, France*

Abstract: A control strategy is proposed to control the internal fluid temperature at the outlet of a co-current heat exchanger by manipulating the inlet external fluid temperature. The dynamic model of the heat exchanger is given by two partial differential equations. Based on nonlinear geometric control, a state-feedback law that ensures a desired performance of a measured output defined as spatial average temperature of the internal fluid is derived. Then, in order to control the outlet internal fluid temperature, a control strategy is proposed where an external controller is introduced to provide the set point of the considered measured output by taking as input the error between the outlet internal fluid temperature and its desired set point. The validity of the proposed control design and strategy is examined in simulation by considering the tracking and perturbation rejection problems. *Copyright©2009 IFAC*.

Keywords: distributed parameter system, partial differential equation, geometric control, characteristic index,PI controller, co-current heat exchanger.

1. INTRODUCTION

As a thermal device, heat exchangers are widely used in process industries both for cooling and heating operations. The dynamic behavior of the heat exchanger is modeled by a set of partial differential equations (PDE) that describe the spatio-temporal variation of the temperatures. Thus, the need to find the best operating conditions for the heat exchangers and to improve their effectiveness lead to take into account their distributed nature. In this context, good performances can be attained using more efficient control strategy based on the direct use of the distributed parameter model rather than a reduced or a lumped model (Ray, 1989; Christofides, 2001).

Heat exchangers can be classified into two major types according to their flow arrangement: co-current and counter-current heat exchangers. For the first one, the two fluids travel in the same direction. By contrast, for the second one, the fluids move in opposite directions.

In the control problem of tube heat exchangers, the variable which is manipulated, theoretically, is the thermal power at the inlet of the outer tube, i.e. grossly the product of a flow rate and a difference of temperature. In practice, to control the outlet temperature of a heat exchanger, two possible strategies which are not equivalent exist. The first one is to use the inlet temperature of the external fluid, while the second is to manipulate its flow rate.

When the flow rate is considered as a manipulated variable, if it becomes too low, the flow regime in the outer tube can be laminar instead of turbulent, which affects the parameters of the models, in particular the heat transfer coefficient (Xuan and Roetzel, 1993). So the tuning of the controller should vary with the flow rate, which is a difficult task (Abdelghani-Idrissi et al., 2001; Arbaoui et al., 2007). In addition, by manipulating the flow rate, a minimum bound is to set on this input. With the temperature as a manipulated input, it is possible to work at a constant large flow rate and the hydrodynamic regime is invariable. Physically, manipulating the temperature is almost possible if this latter is the outlet of a process with fast dynamics like plate heat exchangers. Potential flow rate variations will be assumed as a disturbance that affects the system and needs to be rejected by the designed controller.

Control of counter-current heat exchanger has attracted much attention, and several strategies are proposed based either on the PDE model or ODE model (see e.g. Maidi et al. (2008a) for more references) compared to the co-current heat exchanger for which few methods have been proposed in the literature. Derese and Noldus (1980) addressed the problem of controlling of the co-current heat exchanger using dynamical lumped parameter controllers designed based on technical frequency domain specifications. Based on the conjugate gradient method (CGM) of minimization, Huang and Yeh (2003) proposed an algorithm for determining an optimal external distributed heat-flux of a steady state co-current heat exchanger.

In this paper, a control strategy is proposed to control the outlet internal fluid temperature of a co-current heat exchanger by manipulating the inlet external fluid temperature. The designed approach is based on the use of the PDE model that describes the dynamic behavior. The idea is to design a state-feedback control that allows controlling the average temperature of the internal tube of the heat exchanger, assumed as the measured output. As it will be demonstrated, the direct design of a control law by considering the outlet temperature as the controlled variable is a difficult task due to the fact that the process is infinite-dimensional. Then, in order to control the outlet fluid temperature, a control strategy is proposed where a PI controller is introduced to provide the set point of the measured

⋆ Corresponding author: corriou@ensic.inpl-nancy.fr

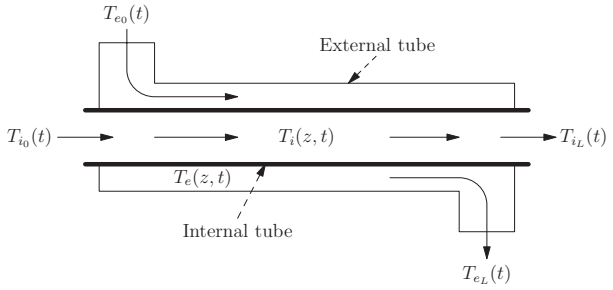


Fig. 1. co-current heat exchanger.

output (spatial average temperature). The design of the state-feedback control law makes use of the concept of characteristic index (Christofides and Daoutidis, 1996), which characterize the spatiotemporal interactions between the controlled and manipulated variables.

The paper is organized as follows. In section 2, the studied co-current heat exchanger is presented and its dynamic model given as a system of two PDEs. Section 3 concerns the formulation of the control problem and the design methodology. Section 4 is dedicated to simulation results concerning tracking and perturbation rejection problems. Finally, a conclusion ends the article.

2. CO-CURRENT HEAT EXCHANGER DYNAMIC MODEL

2.1 Description of the heat exchanger

The process studied in this work corresponds to a tubular co-current heat exchanger (Fig. 1). A fluid of constant density ρ_i and of heat capacity C_{p_i} flows through the internal tube of a heat exchanger, of length L, with a constant velocity v_i. This fluid enters at temperature T_{i_0} and exchanges heat with the an external fluid or non condensating vapor fluid, of constant density ρ_e and of heat capacity C_{p_e}, which flows in the same direction in the jacket with a velocity v_e. This fluid enters at temperature T_{e_0}. At the outlet of the exchanger, the internal fluid leaves at temperature T_{i_L}. In the present study, the internal and external cross sections S_i and S_e of the heat exchanger are supposed to be uniform and the surface area used for the heat transfer per unit length is $\mathcal{A}$. Both temperatures T_i of the internal fluid and T_e of the external fluid depend on time and spatial position along the tube.

The energy balance of the heat exchanger, after classical simplifying hypotheses (Ray and Ogunnaike, 1994), gives the following partial differential equation for the internal tube (internal fluid)

$$\frac{\partial T_i(z,t)}{\partial t} = -v_i \frac{\partial T_i(z,t)}{\partial z} + h_i \left[T_e(z,t) - T_i(z,t)\right] \qquad (1)$$

and the following partial differential equation for the jacket (external fluid)

$$\frac{\partial T_e(z,t)}{\partial t} = -v_e \frac{\partial T_e(z,t)}{\partial z} + h_e \left[T_i(z,t) - T_e(z,t)\right] \qquad (2)$$

where $h_i = \dfrac{U_i\,\mathcal{A}}{\rho_i\,S_i\,C_{p_i}}, \quad h_e = \dfrac{U_e\,\mathcal{A}}{\rho_e\,S_e\,C_{p_e}}.$

T_i and T_e are the temperatures of the internal and external fluids, respectively, h_i and h_e are the heat transfer coefficients, v_i and v_e are the velocities, U_i and U_e are the overall heat transfer coefficients, $\mathcal{A}$ is the surface area devoted to heat transfer.

Each PDE requires an initial condition and a boundary condition to be fully defined. The studied heat exchanger is of co-current type. For Eq. (1) describing the temperature of the internal fluid, the boundary condition is usually specified at $z = 0$ as the temperature of the fluid entering the tube is in general known and measurable. Thus, at $z = 0$, it gives

$$T_i(0,t) = T_{i_0}(t) \qquad (3)$$

and most often the initial condition is some given temperature profile at $t = 0$

$$T_i(z,0) = T_i^*(z) \qquad (4)$$

Similarly, for Eq. (2), describing the distribution of temperature of the external fluid in the jacket, the boundary condition is the temperature of the entering fluid T_{e_0}, specified at $z = 0$, consequently

$$T_e(0,t) = T_{e_0}(t) \qquad (5)$$

while the initial condition is some given temperature profile at $t = 0$

$$T_e(z,0) = T_e^*(z) \qquad (6)$$

Eqs. (1)-(6) constitute the dynamic model of the co-current heat exchanger.

3. CONTROL OF THE CO-CURRENT HEAT EXCHANGER

3.1 Control problem formulation

As indicated above, to control the outlet internal temperature T_{i_L}, two manipulated variables are possible, either the inlet external fluid temperature T_{e_0} or the flow rate represented by the velocity v_e. In this work, the temperature T_{e_0}, corresponding to the boundary condition (5), is taken as a manipulated variable to easily control the outlet internal fluid temperature T_{i_L} since the hydrodynamic regime remains invariable. Now, due to Eq. (2), it is noticeable that by manipulating the boundary condition of the jacket, given by Eq. (5), a variation of the temperature of the external fluid T_e along the jacket results. Thus, by denoting as u the control variable and y the controlled variable, the model of the heat exchanger (1)-(6) takes the following form

$$\frac{\partial T_i(z,t)}{\partial t} = -v_i \frac{\partial T_i(z,t)}{\partial z} + h_i \left[T_e(z,t) - T_i(z,t)\right] \qquad (7)$$

$$\frac{\partial T_e(z,t)}{\partial t} = -v_e \frac{\partial T_e(z,t)}{\partial z} + h_e \left[T_i(z,t) - T_e(z,t)\right] \qquad (8)$$

$$T_i(0,t) = T_{i_0}(t) \qquad (9)$$

$$T_e(0,t) = T_{e_0}(t) = u(t) \qquad (10)$$

$$T_i(z,0) = T_i^*(z) \qquad (11)$$

$$T_e(z,0) = T_e^*(z) \qquad (12)$$

$$y(t) = \mathcal{C}\big(T_i(z,t)\big) = \int_0^L \delta(z-L)\,T_i(z,t)\,dz \qquad (13)$$

where $\mathcal{C}(\,.\,)$ is a bounded linear operator.

3.2 Design approach

Recently the nonlinear geometric control has proved to be very successful as a control approach of the linear and quasi-linear DSP (Christofides and Daoutidis, 1996; Gundepudi and Friedly, 1998; Christofides, 2001; Wu and Liou, 2001; Maidi et al., 2008a,b). The most important advantage of geometric control is that the control law can be designed using directly the PDE model, which leads to distributed control that increases the performances (Christofides, 2001). Thus, this theoretical approach will be used to derive a boundary control law for the co-current heat exchanger.

The manipulated variable $u(t)$ appears as an inhomogeneous part in the boundary condition (10), so in order to obtain the expression of the control law, we propose to insert the manipulated variable $u(t)$ through the use of Dirac delta function in the state equation (7) as follows

$$\frac{\partial T_e(z,t)}{\partial t} = -v_e \frac{\partial T_e(z,t)}{\partial z} + h_e \left[T_i(z,t) - T_e(z,t)\right] + v_e\, \delta(z)\, u(t) \tag{14}$$

so that the model will be affine with respect to the input $u(t)$.

Under these conditions, the boundary condition (10) becomes homogeneous,

$$T_e(0,t) = 0 \tag{15}$$

Now, as the open-loop system $u(t)$-$y(t)$ is infinite dimensional, the characteristic index σ does not exist. This can be easily verified by calculating the successive derivatives of the output (13) with respect to time. To overcome this problem, we propose to consider another measured output given as the average of the external fluid temperature, i.e.

$$y_m(t) = \mathcal{C}_m\big(T_i(z,t)\big) = \int_0^L c_m(z)\, T_i(z,t)\, dz \tag{16}$$

where $\mathcal{C}_m(\,.\,)$ is a bounded linear operator and $c_m(z)$ is a smooth positive function ($c_m(z) > 0$).

In this case, the derivative of the measured output (16) with respect to time yields

$$\begin{aligned}\frac{dy_m(t)}{dt} &= \int_0^L c_m(z)\, \frac{\partial T_i(z,t)}{\partial t}\, dz \\ &= \int_0^L c_m(z) \Bigg(- v_i \frac{\partial T_i(z,t)}{\partial z} + h_i \left[T_e(z,t) - T_i(z,t)\right] \Bigg) dz \end{aligned} \tag{17}$$

the characteristic index is greater than one. Performing one more differentiation, we obtain:

$$\frac{d^2 y_m(t)}{dt^2} = \int_0^L c_m(z) \Bigg(- v_i \frac{\partial}{\partial t}\left(\frac{\partial T_i(z,t)}{\partial z}\right) + h_i \left[\frac{\partial T_e(z,t)}{\partial t} - \frac{\partial T_i(z,t)}{\partial t}\right] \Bigg) dz \tag{18}$$

By substituting the term $\dfrac{\partial T_e(z,t)}{\partial t}$ by its expression given by (14) and after arrangement, equation (18) takes the form

$$\frac{d^2 y_m(t)}{dt^2} = I_1 + h_i\, v_e \Bigg[\underbrace{\int_0^L c_m(z) \delta(z)\, dz}_{I_2} \Bigg] u(t) \tag{19}$$

where I_1 is the remaining term of the integral in equation (18). According to equation (19), it is clear that the input appears linearly.

Now, in order to have the control law $u(t)$ well-defined, the integral term I_2 must be different from zero. This condition ensures that the characteristic index of the measured output $y_m(t)$ with respect to the manipulated input $u(t)$ is equal to 2. The calculus of I_2 gives

$$I_2 = \int_0^L c_m(z)\, \delta(z)\, dz = c_m(z)|_{z=0} \tag{20}$$

The condition on the characteristic index being equal to 2 is related to the choice of the function $c_m(z)$, i.e. the value of $c_m(z)$ should not be zero at $z = 0$

$$I_2 = c_m(0) \neq 0 \tag{21}$$

Thus, by choosing a function $c_m(z) \geq 0$ that satisfies the condition (21), the characteristic index will be $\sigma = 2$. In summary, the modification of equation (14) by introduction of the manipulated input and the consideration of the new output (16) have ensured the existence of the characteristic index.

As $\sigma = 2$, this suggests requesting the following input-output response of the closed-loop system

$$\tau_2 \frac{d^2 y_m(t)}{dt^2} + \tau_1 \frac{dy_m(t)}{dt} + y_m(t) = v(t) \tag{22}$$

Substituting (19) into equation (22), we obtain the following state-feedback control law

$$u(t) = \frac{1}{h_i\, v_e\, \tau_2\, I_2} \Big[v(t) - y_m(t) - \tau_1\, \dot{y}_m(t) - \tau_2\, I_1 \Big] \tag{23}$$

where τ_1, τ_2 are adjustable controller parameters chosen to guarantee the input-output stability and to enforce the desired performance specifications for the output $y_m(t)$ (Christofides, 2001), and $v(t)$ is an external input.

The control robustness dealing with problems of model and parameter uncertainty and unmodeled dynamics, is provided in (23) through application of the linear control theory to the resulting linear [input $v(t)$-output $y_m(t)$] linear system to define the external input $v(t)$ by a robust controller. In this work, in order to ensure the robustness, i.e. to handle uncertainties and unmodeled dynamics, the external input $v(t)$ is defined by means of a PI controller (Kravaris and Kantor, 1990) as follows

$$v(t) = K_{c_m} \left[\big(y_m^d(t) - y_m(t)\big) + \frac{1}{\tau_{I_m}} \int_0^t \big(y_m^d(\xi) - y_m(\xi)\big)\, d\xi \right] \tag{24}$$

where K_{c_m}, τ_{I_m} are respectively the proportional gain, integral time constant of the PI controller, respectively. $y_m^d(t)$ is the set-point of the measured variable $y_m(t)$.

Thus, the transfer function of the closed loop system is the following

$$\frac{Y_m(s)}{Y_m^d(s)} = \frac{K_{c_m}\left(\tau_{I_m}+1\right)}{\tau_{I_m}\,\tau_2\,s^3+\tau_{I_m}\,\tau_1\,s^2+\left(\tau_{I_m}+K_{c_m}\,\tau_1\right)+K_{c_m}} \quad (25)$$

The scalar parameters K_{c_m}, τ_{I_m} and τ are tuned in order for the denominator to approach a polynomial minimizing an ITAE criterion (Corriou, 2004) and it can be verified that the following polynomial is Hurwitz (the poles have a negative real part) to ensure the closed loop stability related to the roots of the characteristic equation

$$\tau_2\tau_{I_m}\,s^3+\tau_1\tau_{I_m}\,s^2+\left(\tau_{I_m}+\tau_1 K_{c_m}\right)s+K_{c_m}=0 \quad (26)$$

At this point, it is clear that the control law (23) ensures the desired performances of the introduced measured output $y_m(t)$ rather the controlled output $y(t)$. Actually, the output $y_m(t)$ is introduced only in order to avoid the problem of non-existence of the characteristic index. In order to solve the formulated boundary control problem, i.e. controlling the output $y(t)$, we propose to keep the control law (23) derived for the measured output (16) with $c(z)$ satisfying the condition (21). Then, define the set point of the measured output $y_m(t)$, denoted by $y_m^d(t)$, by means of a PI controller taking as input the error $e(t) = y^d(t) - y(t)$, where $y^d(t)$ is the corresponding set point of the controlled variable $y(t)$. Note that another control technique can be adopted to provide the set point $y_m^d(t)$. The proposed global control strategy is summarized in Fig. 2.

The control law (23) requires that the complete state $T_i(z,t)$ must be available especially to evaluate the integral term I_1 and the measured output $y_m(t)$. From a practical point of view, this is impossible since the state $T_i(z,t)$ is infinite. Ray (1989) discusses some way that can provide the complete state of a distributed parameter system. The design of Kalman filter that estimates the whole state variables vector in the case of a counter-current heat exchanger has been studied by Maidi et al. (2008a). In this work, it is considered that the vector of state variables is fully available to clearly show the effectiveness and the contribution of the proposed control strategy.

Note that the choice of the function $c_m(z)$ is not unique. Nevertheless, the relation (21) shows that the function $c_m(z)$ is involved in the evaluation of the integral term I_1 and in calculating the measured output $y_m(t)$ and its derivative $\dot{y}_m(t)$, so it is suggested to choose a simple function for example $c_m(z) = L - z$. From a practical point of view, these calculations can be provided simply by a computer by processing the data measurements $T_i(z,t)$ and $T_e(z,t)$.

4. SIMULATION RESULTS

In this section, the performance of the proposed control strategy will be illustrated through application examples. For simulation purpose of the closed-loop system, the method of lines (Wouwer et al., 2004) is used by considering a number of discretization points $N = 100$. The control is held constant over the sampling period equal to $0.02\,\mathrm{s}$ in all simulation runs. The integral term I_1, the measured output $y_m(t)$ and its derivative $\dot{y}_m(t)$ involved in the control law (23) are evaluated numerically using the trapezoidal method. The terms involving differentiation according to the space variable z are evaluated by means of finite differences.

The heat exchanger parameters (Friedly, 1972) are $v_e = 2\,\mathrm{m.s^{-1}}$, $v_i = 1\,\mathrm{m.s^{-1}}$, $h_e = 1\,\mathrm{s^{-1}}$, $h_i = 1\,\mathrm{s^{-1}}$ and $L = 1\,\mathrm{m}$. For the internal PI controller, the tuning parameters obtained following the tuning procedure described at the end of the section 3.2 are $K_{c_m} = 0.0240$, $\tau_{I_m} = 0.0469\,\mathrm{s}$. The tuning parameters K_c and τ_I of the external PI controller that provide the set point $y_m^d(t)$ have been achieved by trial and error and observation of the obtained performance, so that the retained parameters are $K_c = 0.02$ and $T_i = 0.3\,\mathrm{s}$.

The initial conditions $T_i(z,0)$ and $T_e(z,0)$ are the steady state profiles (Fig. 3) defined by $T_{i_0}(t) = 25°\mathrm{C}$ and $T_{e_0}(t) = 50°\mathrm{C}$.

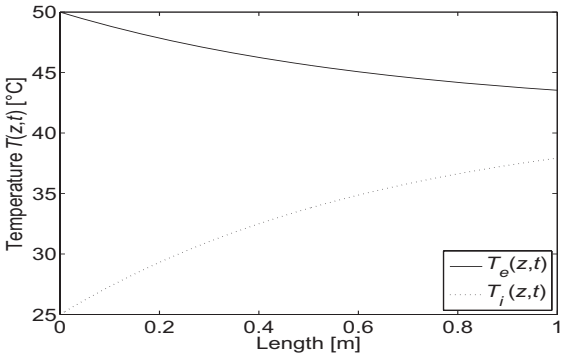


Fig. 3. Profiles of the temperatures of the internal and external fluids at steady state.

4.1 Tracking problem

In the first simulation run, the reference input tracking capabilities of the controller are studied. Thus, two step set points have been specified at times $t = 1\,\mathrm{s}$ and $t = 30\,\mathrm{s}$ corresponding respectively to $y^d(t) = 60\,°\mathrm{C}$ and $y^d(t) = 30\,°\mathrm{C}$. On Fig. 4, it is clear that the output $y(t)$ (Fig. 4b) follows perfectly the imposed set point whereas the control moves of $u(t)$ are physically acceptable (Fig. 4c). In addition, the spatial profiles of temperature obtained at time $t = 60\,\mathrm{s}$ is also realistic (Fig. 4d).

4.2 Disturbance rejection

The second performed test concerns the problem of disturbance rejection. The performances of the control strategy are thus evaluated with respect to changes of the internal fluid temperature at the inlet of the heat exchanger which is a disturbance for the process. For that reason, a step of -10% of the temperature of the entering internal fluid (at $z = 0$) is imposed as a disturbance at time $t = 30\,\mathrm{s}$, after having imposed a step set point at time $t = 1\,\mathrm{s}$ corresponding to $y^d(t) = 60\,°\mathrm{C}$. From Fig. 5, it is clear that the controller behaves adequately to reject the disturbance effect and achieve perfectly the set point tracking (Fig. 5b). The dynamic behavior of the manipulated variable $u(t)$ (Fig. 5c) remains also physically admissible. Again, the profiles of temperatures at $t = 60\,\mathrm{s}$, after successively the step set point and the step disturbance, are typical of the behavior of a co-current heat exchanger (Fig. 5d).

5. CONCLUSION

In this paper, the geometric control of a co-current heat exchanger is investigated, and a control strategy is proposed to

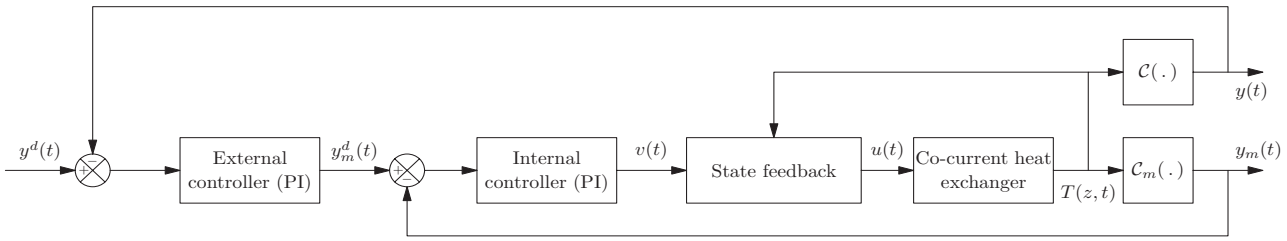


Fig. 2. Global control strategy of a co-current heat exchanger ($T(z,t) = [T_e(z,t)\, T_i(z,t)]^T$).

control the outlet internal fluid temperature. The main idea consists in inserting the manipulated variable, i.e. the inlet external fluid temperature, in the state equations of the heat exchanger by means of a Dirac function. Furthermore, the spatial average temperature of the internal fluid has been introduced, as measured output, in order to ensure the existence of the characteristic index. Then, to achieve a desired performance of the outlet internal fluid temperature, a control strategy is proposed where a PI external controller is introduced to provide the set point of the introduced measured output by taking as input the error between the outlet internal fluid temperature and its desired set point. The effectiveness of the proposed design and control strategy is demonstrated through numerical experiments. The simulation results show that the control strategy behaves correctly and ensures a satisfactory tracking and disturbance rejection.

REFERENCES

M. A. Abdelghani-Idrissi, F. Bagui, and L. Estel. Analytical and experimental response time to flow rate step along a counter flow double pipe heat exchanger. *International Journal of Heat and Mass Transfer*, 44:3721–3730, 2001.

M. A. Arbaoui, L. Verniéres-Hassimi, D. Seguin, and M. A. Abdelghani-Idrissi. Counter-current tubular heat exchanger: modeling and adaptive predictive functional control. *Applied thermal Engineering*, 27:2332–2338, 2007.

P. D. Christofides. *Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes*. Birkhäuser, Boston, 2001.

P. D. Christofides and P. Daoutidis. Feedback control of hyperbolic PDE systems. *AIChE Journal*, 42:3063 – 3086, 1996.

J. P. Corriou. *Process control - Theory and applications*. Springer, London, 2004.

I. Derese and E. Noldus. Control of parallel current and counter-current heat exchangers. *Analysis and Optimization of Systems*, 28:497–515, 1980.

J. C. Friedly. *Dynamic Behaviour of Processes*. Prentice-Hall, New Jersey, 1972.

P. K. Gundepudi and J. C. Friedly. Velocity control of hyperbolic partial differential equation systems with single characteristic variable. *Chemical Engineering Science*, 53(24): 4055 – 4072, 1998.

C.-H. Huang and C.-Y. Yeh. An optimal control algorithm for entrance co-current flow problems. *International Journal of heat and Mass Transfer*, 46:1013–1027, 2003.

C. Kravaris and J. C. Kantor. Geometric methods for nonlinear process control. 2. controller synthesis. *Industrial & Engineering Chemistry Research*, 29:2310–2323, 1990.

A. Maidi, M. Diaf, and J.-P. Corriou. Boundary geometric control of a counter-current heat exchanger. *Journal of Process Control*, In Press, 2008a.

A. Maidi, M. Diaf, and J.-P. Corriou. Distributed geometric control of wave equation. *17th IFAC World Congress*, July 6-11, Seoul, Korea, 2008b.

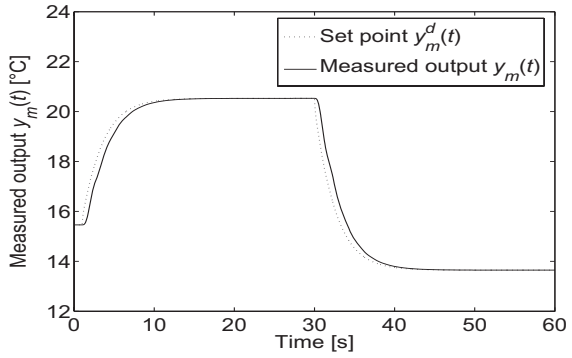
W. H. Ray. *Advanced Process Control*. Butterworths, Boston, 1989.

W.H. Ray and B.A. Ogunnaike. *Process Dynamics, Modeling and Control*. Oxford University Press, 1994.

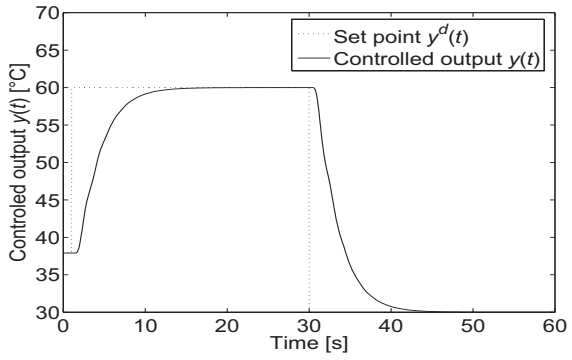
A. V. Wouwer, P. Saucez, and W. E. Schiesser. Simulation of distributed parameter systems using a matlab-based method of lines toolbox: Chemical engineering applications. *Industrial and Engineering Chemical Research*, 43(14):3469–3477, 2004.

W. Wu and C. Liou. Output regulation of nonisothermal plug-flow reactors with inlet perturbations. *Computers and Chemical Engineering*, 25:433–443, 2001.

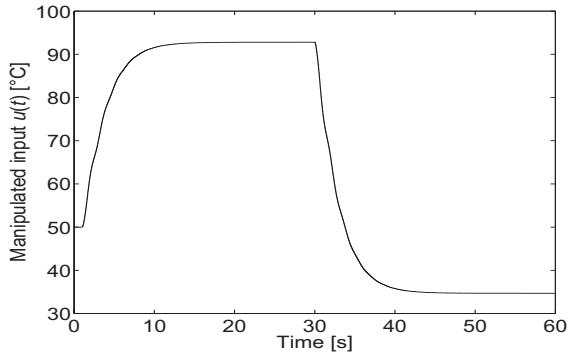
Y. Xuan and W. Roetzel. Dynamics of shell-and-tube heat exchangers to arbitrary temperature and step flow variations. *AIChE Journal*, 39:413–421, 1993.



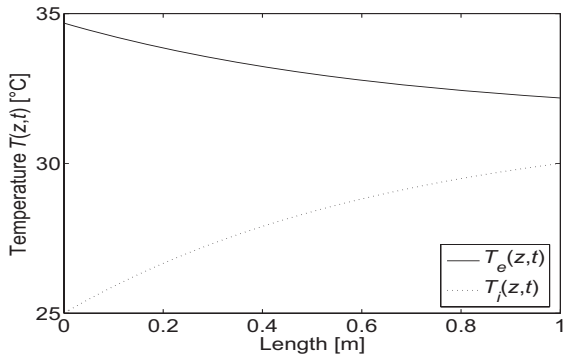
(a) Measured output $y_m(t)$.



(b) Controlled output $y(t)$.

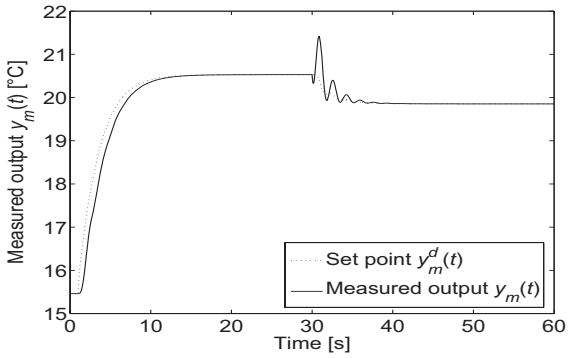


(c) Manipulated variable $u(t) = T_e(0,t)$.

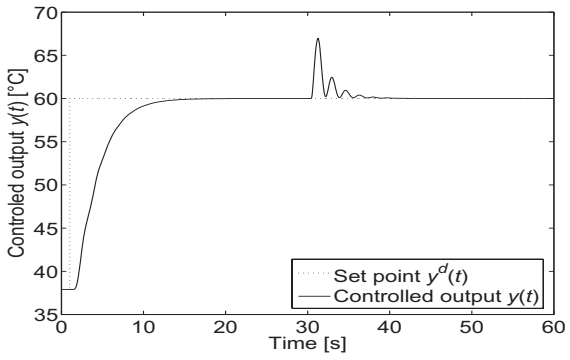


(d) Spatial profiles $t = 60$ s.

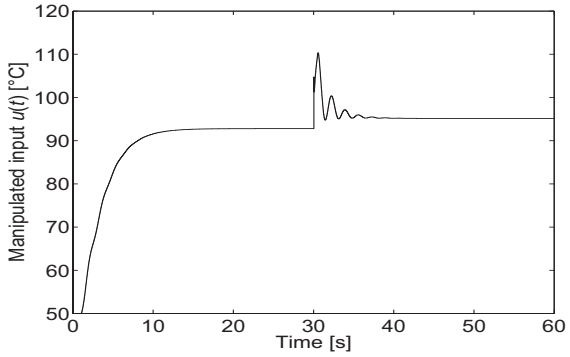
Fig. 4. Set point tracking.



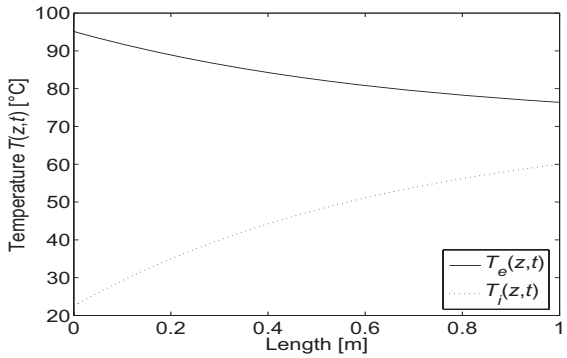
(a) Measured output $y_m(t)$.



(b) Controlled output $y(t)$.



(c) Manipulated variable $u(t) = T_e(0,t)$



(d) Spatial profile at $t = 60$ s.

Fig. 5. Disturbance rejection.

Feedback Controller Design for the Four-Tank Process using Dissipative Hamiltonian Realization

Nicolas Hudon * **Martin Guay** *

* *Department of Chemical Engineering, Queen's University, Kingston, ON, Canada, K7L 3N6 (e-mail:* martin.guay@chee.queensu.ca*)*

Abstract: This paper considers the problem of stabilizing the quadruple-tank process using an approximate dissipative Hamiltonian realization. The proposed approach consists in canceling by feedback the deviation of the system from a Hamiltonian system. First, we obtain a characteristic one-form for the system by taking the interior product of a non vanishing two-form with respect to the controlled vector field. We then construct a homotopy operator on a star-shaped region centered at a desired equilibrium point. The dynamics of the system is then decomposed into an exact part and an anti-exact one. The exact part is generated by a potential, hence stability of this part is guaranteed using the generating potential as a Lyapunov function. The stabilizing feedback controller is designed by canceling the anti-exact part of the characteristic one-form. Application of the resulting controller is illustrated by numerical simulations.

Keywords: Feedback Regulation, Approximate Dissipative Hamiltonian Realization, Stability.

1. INTRODUCTION

Application of generalized Hamiltonian systems are an important approach for stability studies and controller design of nonlinear control systems (van der Schaft, 2000) and several physical problems were studied using this class of dynamical system representations. One example in chemical engineering was given recently by Otero-Muras et al. (2008) who studied the stability of a reaction network using its dissipative Hamiltonian representation. However, one limitation associated with the study of non-mechanical nonlinear systems using dissipative Hamiltonian is to derive a suitable Hamiltonian function for the problem. As discussed in (Johnsen and Allgöwer, 2007) and (Ortega et al., 1999), applications of Interconnection and Damping Assignment Passivity-Based Control (IDA-PBC) techniques is difficult since the concept of "energy" is usually ill-defined for process control applications, for example when mass balances are considered. In (Cheng et al., 2005), it was shown that a nonlinear system of the form

$$\dot{x} = f(x) + G(x)u, \tag{1}$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, and $G(x)$ full rank, is transformable to a stable Port-Controlled Hamiltonian (PCH) system

$$\dot{x} = F(x)\nabla H(x), \quad F(x) = [J(x) - R(x)] \tag{2}$$

if there exists a feedback $\beta : \mathbb{R}^n \to \mathbb{R}^m$ such that the *matching equation*

$$f(x) + G(x)\beta(x) = F(x)\nabla H(x) \tag{3}$$

holds. In particular, for a fixed structure matrix $F(x)$ and a free Hamiltonian function $H(x)$, the problem leads to a set of PDEs parameterized by the structure matrix and the feedback controller $\beta(x)$. Relaxing the need for exact matching, a non-exact matching IDA-PBC approach was recently developed and applied successfully to chemical reactor process stabilization (Ramírez et al., 2009).

In this paper, we will address the problem of stabilizing controllers design using approximate dissipative Hamiltonian realization. In (Cheng et al., 2000), conditions for approximate Hamiltonian realizations were given in terms of a normal form. Sufficient conditions and a constructive algorithm for a generalized Hamiltonian realization for time-invariant nonlinear systems were presented in (Wang et al., 2003). In particular, the method proposed in (Wang et al., 2003) proposed a vector field decomposition along the gradient direction $\nabla H(x)$ and the tangential direction of the energy surfaces of $H(x)$, for a regular positive-definite function $H(x)$. Following the work in (Maschke et al., 2000) which related port-controlled Hamiltonian systems to the construction of Lyapunov functions, it was shown in (Wang et al., 2007) how k-th degree approximate dissipative Hamiltonian systems can be used to solve the realization problem and how associated k-th degree approximate Lyapunov functions can be used to study the stability of such systems.

In the following, we propose to use the tools of exterior calculus to construct the Hamiltonian function and design a stabilizing controller. It is shown that a stabilizing controller can be developed by canceling the anti-exact part of the dynamics (this dynamics acts tangentially to the dynamics generated by the potential). More precisely, assuming a controller structure, we obtain a characteristic one-form for the system by taking the interior product of a non vanishing two-form with respect to the vector field. A homotopy operator centered at a desired equilibrium

point for the system is used to obtain an exact one form, generated by a Hamiltonian function, and an anti-exact form that generates the tangential dynamics. We design the controller in such a way that the anti-exact form vanishes. The stability argument presented in (Hudon et al., 2008) uses local equivalence between the exact part of the dynamics and a pre-defined Hamiltonian dissipative realization, viewed as a reference system to develop a local change of coordinates to derive the desired local dissipative potential for the system.

The paper is organized as follows. Section 2 presents the four-tank system as a motivating example. In Section 3, mathematical background is presented, recalling the elements required for the development of the radial homotopy operator that is used in the sequel. The application of this operator to discriminate the exact and anti-exact parts of the dynamics and the development of the stabilizing controller are presented in Section 4. Numerical applications to the four-tank system are given in Section 5. Conclusions and future areas of investigation are outlined in Section 6.

2. QUADRUPLE-TANK PROCESS EXAMPLE

To motivate the present paper, we use the four-tank system studied in details in (Johansson, 2000). More recently, Johnsen and Allgöwer (2007) developed an IDA-PBC controller for the system by introducing error dynamics and solving the matching equations assuming a perturbed Hamiltonian function for the closed-loop dynamics.

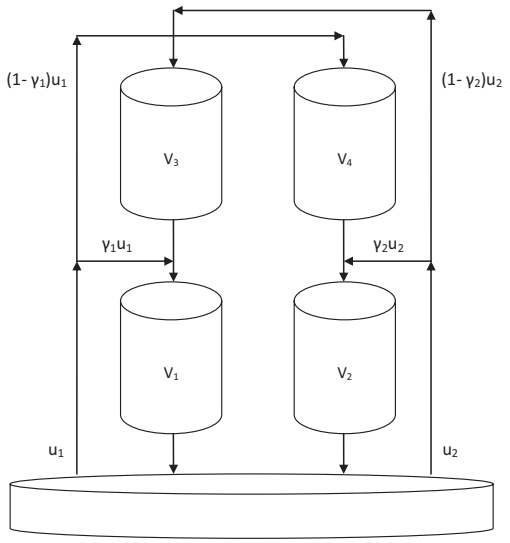


Fig. 1. Four-Tank System

The dynamic model for the four-tank system is given as a control affine nonlinear system of the form

$$\dot{x} = f(x) + G(x)u \tag{4}$$

where $x \in \mathbb{R}^4$ are the levels in the respective tanks and $u \in \mathbb{R}^2$ are the manipulated flows. Using the model proposed in (Johnsen and Allgöwer, 2007), $f(x)$ and $G(x)$ are given by

$$f(x) = \begin{pmatrix} \frac{-a_1}{A_1}\sqrt{2gx_1} + \frac{a_3}{A_1}\sqrt{2gx_3} \\ \frac{-a_2}{A_2}\sqrt{2gx_2} + \frac{a_4}{A_2}\sqrt{2gx_4} \\ \frac{-a_3}{A_3}\sqrt{2gx_3} \\ \frac{-a_4}{A_4}\sqrt{2gx_4} \end{pmatrix}, \tag{5}$$

$$G(x) = \begin{pmatrix} \frac{\gamma_1}{A_1} & 0 \\ 0 & \frac{\gamma_2}{A_2} \\ 0 & \frac{1-\gamma_2}{A_3} \\ \frac{1-\gamma_1}{A_4} & 0 \end{pmatrix}. \tag{6}$$

The parameters A_i represent the cross sections of the respective tanks $i = 1, \ldots, 4$, such that the volumes are given by $V_i = A_i x_i$. The parameters a_i are the cross section of the outlet holes. The gravitational acceleration is denoted by g. The parameters $\gamma_1,\ \gamma_2 \in [0,\ 1]$ are the valve parameters that determined how much of the flows u_i are re-directed in bottom tanks $i = 1, 2$. If the levels of tanks 1 and 2 are the only measured states, it was shown in (Johansson, 2000) that the condition for stable zero dynamics is that $\gamma_1 + \gamma_2 \neq 1$.

To stabilize the system at a desired admissible steady-state, (x^*, u^*), we propose a controller of the form

$$u_1(t) = k_{11}(x) \cdot x_1(t) + k_{12}(x) \cdot x_2(t) \tag{7}$$

$$u_2(t) = k_{21}(x) \cdot x_1(t) + k_{22}(x) \cdot x_2(t). \tag{8}$$

At this point, we assume that all tanks levels are measured. In Section 5, we will discuss how this requirement can be relaxed in the case where only x_1 and x_2 are measured.

3. EXTERIOR CALCULUS AND HOMOTOPY OPERATOR

In this section, we show how to construct a homotopy operator $\mathbb{H}$, *i.e.*, a linear operator on differential forms ω, that satisfies the identity

$$\omega = d(\mathbb{H}\omega) + \mathbb{H}d\omega. \tag{9}$$

We first recall some notions of exterior calculus on $\mathbb{R}^n$ (Edelen, 1985). We denote a smooth vector field $X \in \Gamma^\infty(\mathbb{R}^n)$ as a smooth map

$$X : \mathbb{R}^n \to T\mathbb{R}^n, \qquad X|_x = \sum_{i=1}^{n} v^i(x)\partial_{x_i}|_x. \tag{10}$$

The cotangent (dual) space $T_x^*\mathbb{R}^n$ is the set of all linear functionals on $T_x\mathbb{R}^n$,

$$T_x^*\mathbb{R}^n = \{\omega|_x : T_x\mathbb{R}^n \to \mathbb{R}\} \tag{11}$$

where each $\omega|_x$ is linear, *i.e.*

$$(a\omega_1|_x + b\omega_2|_x)(X_x) = a\omega_1|_x(X|_x) + b\omega_2|_x(X|_x). \quad (12)$$

The standard basis of $T_x^*\mathbb{R}^n$ is given by $\{dx_1, \ldots, dx_n\}$, where $dx_i(\partial_{x_j}) = \delta_j^i$, δ_j^i being the Kronecker delta. An element $\omega|_x$ in the cotangent space $T_x^*\mathbb{R}^n$ can be written as

$$\omega|_x = \sum_{i=1}^{n} \omega_i dx_i, \quad \omega_i \in \mathbb{R}. \quad (13)$$

In the sequel, differential one-forms will be used. We write

$$\omega = \sum_{i=1}^{n} \omega_i(x) dx_i, \quad (14)$$

where ω_i are smooth functions on $\mathbb{R}^n$. The exterior (wedge) product $\wedge$ is defined on $\Omega^1(\mathbb{R}^n) \times \Omega^1(\mathbb{R}^n)$ by the requirements

$$dx_i \wedge dx_j = -dx_j \wedge dx_i$$
$$dx_i \wedge f(x)dx_j = f(x)dx_i \wedge dx_j$$

for all smooth functions $f(x)$ and

$$\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \gamma, \quad (15)$$

for all $\alpha, \beta, \gamma \in T^*\mathbb{R}^n$. If $\alpha \in \Lambda^k(\mathbb{R}^n)$, then we write $\deg \alpha = k$. Notice that $\Lambda^1(\mathbb{R}^n) = T^*\mathbb{R}^n$ and $\Lambda^0(\mathbb{R}^n) = \mathcal{C}^\infty(\mathbb{R}^n)$.

The differential operator d is the unique operator on $\Lambda(\mathbb{R}^n) = \bigoplus_{k=0}^n \Lambda^k(\mathbb{R}^n)$ with the following properties:

$$d : \Lambda^k(\mathbb{R}^n) \to \Lambda^{k+1}(\mathbb{R}^n), \quad 0 \le k \le n-1, \quad (16)$$

1. $d(\alpha + \beta) = d\alpha + d\beta$.
2. $d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^{\deg \alpha} \alpha \wedge d\beta$.
3. $df = \frac{\partial f_i}{\partial x_i} dx_i$, $\forall f(x) \in \Lambda^0(\mathbb{R}^n)$.
4. $d \circ d\alpha = 0$.

A k-form α is said to be closed if $d\alpha = 0$. It is said to be exact if there exists a $(k-1)$-form β such that $d\beta = \alpha$.

The interior product $\lrcorner$ is a map

$$\lrcorner : \Gamma^\infty(\mathbb{R}^n) \times \Lambda^k(\mathbb{R}^n) \to \Lambda^{k-1}(\mathbb{R}^n) \quad (17)$$

with the following properties $\forall X \in \Gamma^\infty(\mathbb{R}^n)$ and $\forall f \in \Lambda^0(\mathbb{R}^n)$:

1. $X \lrcorner f = 0$.
2. $X \lrcorner \omega = \omega(X), \forall \omega \in \Lambda^1(\mathbb{R}^n)$.
3. $X \lrcorner (\alpha + \beta) = X \lrcorner \alpha + X \lrcorner \beta, \forall \alpha, \beta \in \Lambda^k(\mathbb{R}^n)$, $k = 1, \ldots, n$.
4. $X \lrcorner (\alpha \wedge \beta) = (X \lrcorner \alpha) \wedge \beta + (-1)^{\deg(\alpha)} \alpha \wedge (X \lrcorner \beta), \forall \alpha, \beta \in \Lambda(\mathbb{R}^n)$.

The first step in the construction of a homotopy operator is to define a star-shaped domain on $\mathbb{R}^n$ (Edelen, 1985; Banaszuk and Hauser, 1996). An open subset S of $\mathbb{R}^n$ is said to be star-shaped with respect to a point $p^0 = (x_1^0, \ldots, x_n^0) \in S$ if the following conditions hold:

- S is contained in a coordinate neighborhood U of p^0.
- The coordinate functions of U assign coordinates $(x_1^0, \ldots, x_n^0)$ to p^0.
- If p is any point in S with coordinates $(x_1, \ldots, x_n)$ assigned by functions of U, then the set of points $(x^0 + \lambda(x - x^0))$ belongs to S, $\forall \lambda \in [0, 1]$.

A star-shaped region S has a natural associated vector field $\mathfrak{X}$, defined by

$$\mathfrak{X}(x) = [x_i^0 + \lambda(x_i - x_i^0)]\partial_{x_i}, \quad \forall x \in S. \quad (18)$$

In this paper, we will consider the case where the star-shaped domain is centered at the desired equilibrium point x_i^*.

For a differential form ω of degree k on a star-shaped region S centered at an equilibrium point, the homotopy operator will be defined, in coordinates, as

$$(\mathbb{H}\omega)(x) = \int_0^1 \mathfrak{X}(x) \lrcorner\, \omega(\lambda x) \lambda^{k-1} d\lambda, \quad (19)$$

where $\omega(\lambda x)$ denotes the differential form evaluated on the star-shaped domain in the local coordinates defined above.

The properties of the homotopy operator are as follows (Edelen, 1985):

H1. $\mathbb{H}$ maps $\Lambda^k(S)$ into $\Lambda^{k-1}(S)$ for $k \ge 1$ and maps $\Lambda^0(S)$ identically to zero.
H2. $d\mathbb{H} + \mathbb{H}d =$ identity for $k \ge 1$ and $(\mathbb{H}df)(x) = f(x) - f(x_0)$ for $k = 0$.
H3. $(\mathbb{H}\mathbb{H}\omega)(x_i) = 0$, $(\mathbb{H}\omega)(x_i^0) = 0$.
H4. $\mathfrak{X} \lrcorner \mathbb{H} = 0$, $\mathbb{H}\mathfrak{X} \lrcorner = 0$.

The first part of the right hand side of (9), $d(\mathbb{H}\omega)$, is obviously a closed form, since $d \circ d(\mathbb{H}\omega) = 0$. By property (H1), for $\omega \in \Lambda^k(S)$, we have $(\mathbb{H}\omega) \in \Lambda^{k-1}(S)$, $d(\mathbb{H}\omega)$ is also exact on S. We denote the exact part of ω by $\omega_e = d(\mathbb{H}\omega)$ and the anti-exact part by $\omega_a = \mathbb{H}d\omega$. It is possible to show that ω vanishes on $\mathbb{R}^n$ if and only if ω_e and ω_a vanish together (Edelen, 1985).

In the sequel, we will apply the homotopy operator on one-forms. Since in our applications, ω_e is an exact one-form, $(\mathbb{H}\omega)$ computed by homotopy is a dissipative potential. A non dissipative potential is associated with the anti-exact part, but on the star-shaped domain S, ω_a does not contribute to the dissipative part of the system. In other words, ω_a belongs to the kernel of $\mathbb{H}$, which can be seen by applying property (H3) from above to the definition of ω_a. In the next section, we will show how stabilization of the desired equilibrium will be ensured by canceling the dynamics associated with ω_a using feedback.

4. FEEDBACK CONTROLLER DESIGN

4.1 Potential Computation

We now present the central element to the proposed construction, namely using the homotopy operator to discriminate the exact and the anti-exact parts associated to a given autonomous system and then computing a feedback controller to cancel the anti-exact part of the dynamics.

Let the vector field $X|_x = \sum_{i=1}^n f_i(x)\partial_{x_i}$, $i = 1, \ldots, n$ be known. We assume that X is of class $\mathcal{C}^k$ with $k \ge 2$. It is also assumed that X has an equilibrium point, in the

present case, an admissible steady-state for the four-tank process. First, we define a non vanishing closed two-form $\Omega = \sum_{1\leq i<j\leq n} dx_i \wedge dx_j$ on $\mathbb{R}^n$.

Taking the interior product of Ω with respect to the vector field X, we compute a one-form ω as follows

$$\omega = X \lrcorner\, \Omega \tag{20}$$
$$= \sum_{1\leq i<j\leq n} (f_i dx_j - f_j dx_i)\,. \tag{21}$$

Given a star-shaped region centered at the origin, with associated vector field $\mathfrak{X}(x) = x_i \partial_{x_i}$, we have

$$(\mathbb{H}\omega)(x) = \int_0^1 (\mathfrak{X} \lrcorner\, \omega(\lambda x))\, d\lambda. \tag{22}$$

Letting $\tilde{f}_i$ denote the values of the components of f after integration with respect to λ, we have

$$(\mathbb{H}\omega)(x) = \sum_{1\leq i<j\leq n} \left(\tilde{f}_i \cdot x_j - \tilde{f}_j \cdot x_i\right) := \tilde{F}(x). \tag{23}$$

Taking the exterior derivative, we have

$$\omega_e = \sum_{i=1}^{n} \frac{\partial \tilde{F}}{\partial x_i} dx_i. \tag{24}$$

The anti-exact form is then given by

$$\omega_a = \omega - \omega_e$$
$$= \sum_{1\leq i<j\leq n} \left(f_i - \frac{\partial \tilde{F}}{\partial x_j}\right) dx_j - \left(f_j + \frac{\partial \tilde{F}}{\partial x_i}\right) dx_i. \tag{25}$$

Remark 1. As a special case, if one defines Ω to be the canonical symplectic two-form and if X_H is the vector field generated by a known Hamiltonian H, ω obtained by the interior product $X_H \lrcorner\, \Omega$ is closed, and we can show that $\omega = \omega_e = -dH$.

For the quadruple-tank system, using our knowledge of the coupling between the tanks, we define the non-vanishing two-forms as

$$\Omega = dx_1 \wedge dx_3 + dx_2 \wedge dx_4. \tag{26}$$

The characteristic one-form for the system is thus given by

$$\omega = -f_3 dx_1 - f_4 dx_2 + f_1 dx_3 + f_2 dx_4. \tag{27}$$

On a star-shaped region centered at the desired equilibrium point $(x_1^*, x_2^*, x_3^*, x_4^*)$, we have

$$\mathfrak{X} = x_i^* + \lambda(x_i - x_i^*). \tag{28}$$

A net result on our notation for the sequel is that on the star-shaped domain, x denotes deviation variables from the center x^*. In (Hudon et al., 2008), the exact part

$$\omega_e = \sum_{i=1}^{4} f_{e,i}(x) dx_i \tag{29}$$

was used to compute a dissipative potential by equivalence to a normal form of dissipative Hamiltonian realization. In the present paper, we are interested in canceling the anti-exact part ω_a by feedback to ensure stability of the closed-loop dynamics.

4.2 Anti-exact Dynamics Cancelation

As mentioned before, the anti-exact part does not influence the value of the computed dissipative potential, at least on the star-shaped domain where the homotopy operator is defined. However, in order to prove stability, the anti-exact part must also vanishes at the equilibrium point of the system (and only there). In the considered example, we will show that a desired equilibrium can be rendered attractive provided that $\omega_a(x^*) = 0$.

The controlled vector field for the four-tank system is given as in Johnsen and Allgöwer (2007) by Equations (4-6). We fixed the controller to be

$$u_1(t) = k_{11}(x) \cdot x_1(t) + k_{12}(x) \cdot x_2(t) \tag{30}$$
$$u_2(t) = k_{21}(x) \cdot x_1(t) + k_{22}(x) \cdot x_2(t). \tag{31}$$

From Section 4.1, we are left we an anti-exact part of the form:

$$\omega_a = \sum_{i=1}^{4} f_{a,i}(x) dx_i. \tag{32}$$

It is desired to make this form closed by using the elements $k_{i,j}(x)$ of the proposed controller. A one-form is closed if

$$\frac{\partial f_{a,i}}{\partial x_j} = \frac{\partial f_{a,j}}{\partial x_i}. \tag{33}$$

For the four-tank system, it leads us to 5 equations with 4 unknown:

$$k_{22}\frac{\gamma_2 - 1}{A_3} - k_{11}\frac{\gamma_1 - 1}{A_4} = 0 \tag{34}$$
$$\frac{1}{2}\frac{a_3 A_1 \sqrt{2g}}{A_3\sqrt{x_3}} + \frac{1}{2}\frac{a_1\sqrt{2g}}{\sqrt{x_1}} - \gamma_1 k_{11} = 0 \tag{35}$$
$$-\frac{k_{21}\gamma_2}{A_2} = 0 \tag{36}$$
$$-\frac{k_{12}\gamma_1}{A_1} = 0 \tag{37}$$
$$\frac{1}{2}\frac{a_4 A_2 \sqrt{2g}}{A_4\sqrt{x_4}} + \frac{1}{2}\frac{a_2\sqrt{2g}}{\sqrt{x_2}} - \gamma_2 k_{22} = 0. \tag{38}$$

From Equations (36-37), we have that $k_{12} = k_{21} = 0$. From Equations (35) and (38), we have that

$$k_{11}(x) = -\kappa_1 \frac{A_3\gamma_1\sqrt{2x_1x_3}}{a_3A_1\sqrt{gx_1} + a_1A_3\sqrt{gx_3}} \tag{39}$$
$$k_{22}(x) = -\kappa_2 \frac{A_4\gamma_2\sqrt{2x_2x_4}}{a_4A_2\sqrt{gx_2} + a_2A_4\sqrt{gx_4}} \tag{40}$$

where the gains κ_1 and κ_2 are used to guarantee the first equality (34).

The stability argument for the closed loop system uses the Barbashin-Krasovskiĭ, hence the requirement that ω_a vanishes only at the desired equilibrium point. In fact, the condition that $\omega = \omega_a + \omega_e$ be closed along with the requirement that ω vanishes at the desired equilibrium is essentially a convexity condition of a generating potential. In that sense, decomposition of the dynamics using a characteristic one-form is related to the stability requirements

for IDA-PBC as presented in (Ortega et al., 1999) and (Ortega et al., 2002). In the next section, we will illustrate the application of the proposed stabilizing controller.

5. NUMERICAL SIMULATION RESULTS

We now present some numerical applications of the feedback controllers derived in the previous section. Simulation parameters are taken from (Johnsen and Allgöwer, 2007) and are presented in Table 1. We will look at 3 different cases parameterized by the values of γ_1 and γ_2.

Table 1. System Parameters (Johnsen and Allgöwer (2007))

	A_i (cm^2)	a_i (cm^2)
$i = 1, 2$	50.3	0.233
$i = 3, 4$	28.3	0.127

First, we look at the case where $\omega_1 = \omega_2 = 0.6$. For these values, an admissible steady-state x^* is computed to be approximately $x^* = [9,\ 9,\ 4.8,\ 4.8]^T$, and we initialize the system at $x = [4,\ 7,\ 6.8,\ 6.8]^T$. Figures 2 and 3 show that the controller (even with small gains) drives the trajectory to the desired equilibrium and the controller to the consistent steady-state value u^*. Hence by canceling the anti-exact part, the center of the star-shaped domain is attractive.

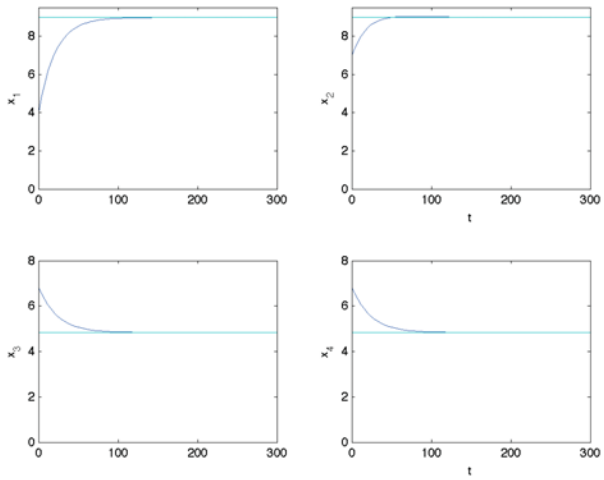


Fig. 2. Full state stabilization of case $\gamma_1 = 0.6,\ \gamma_2 = 0.6$

We now consider the case where $\gamma_1 = \gamma_2 = 0.5$. An admissible steady-state x^* is computed to be approximately $x^* = [10.9,\ 10.9,\ 9.17,\ 9.17]^T$. Initializing the simulation at $x = [5.9,\ 9.9,\ 11.2,\ 11.2]^T$, the proposed controller drives the system to the desired equilibrium (Figures Figures 4 and 5). This case is interesting since, as mentioned in Section 2, if we had considered only output feedback, the zero dynamics for the system are unstable at those values.

To consider output feedback for the case $\gamma_1 = \gamma_2 = 0.5$, we replace $x_3(t)$ and $x_4(t)$ in the controller expressions (39-40) by their desired steady-state values x_3^* and x_4^*. In this particular case, since the zero dynamics is unstable,

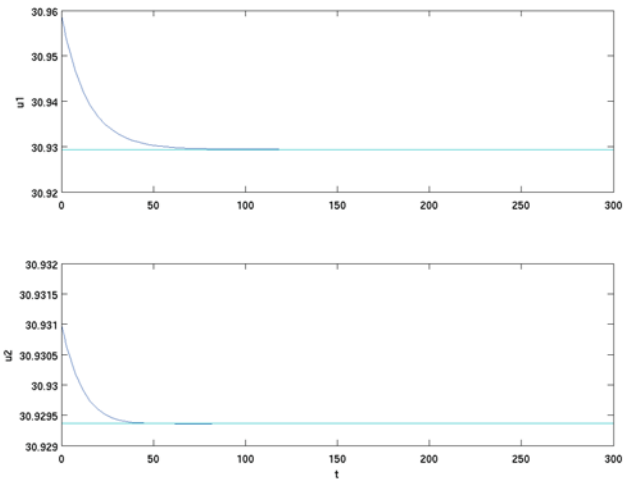


Fig. 3. Control variables values for case $\gamma_1 = 0.6,\ \gamma_2 = 0.6$

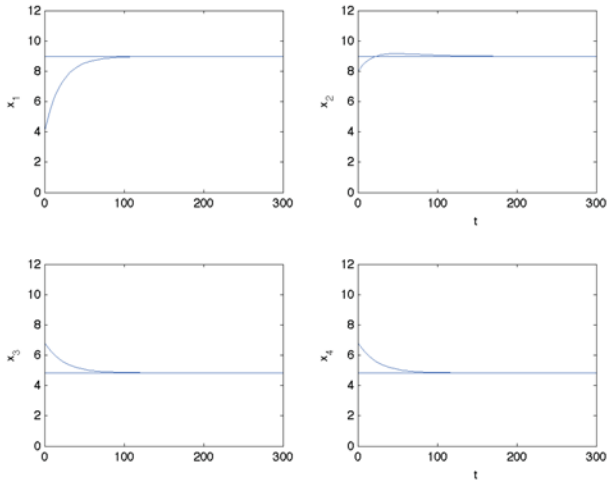


Fig. 4. Full state feedback stabilization of case $\gamma_1 = 0.5,\ \gamma_2 = 0.5$

we use the design parameters κ_1 and κ_2 to make the dynamics of the system associated with the anti-exact part dominated by the gradient term. We seek to reach the same equilibrium point as above $x^* = [10.9,\ 10.9,\ 9.17,\ 9.17]^T$ from two different initial states: $[5.9,\ 9.9,\ 7.2,\ 7.2]^T$ and $[15.9,\ 11.9,\ 11.2,\ 11.2]^T$. Results are presented in Figure 6. As argued in (Ramírez et al., 2009) for a related design approach, the stabilization results present here still hold locally since the proposed controller design procedure does not involve inversion of the dynamics.

6. CONCLUSION

In this paper, a procedure to construct stabilizing controllers using local dissipative Hamiltonian realization for nonlinear dynamical systems was presented. The proposed approach can be seen as an extension of the approximate feedback linearization approach proposed by Banaszuk

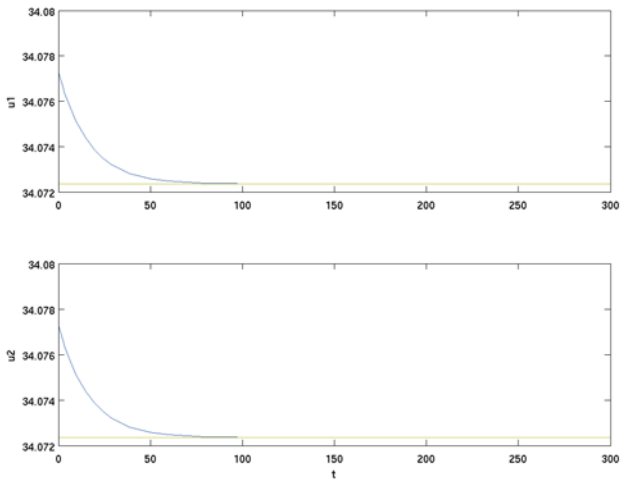


Fig. 5. Control variables values for case $\gamma_1 = 0.5,\ \gamma_2 = 0.5$

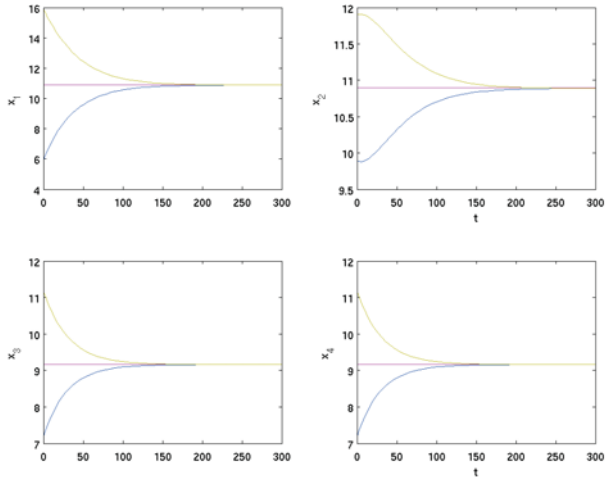


Fig. 6. Output feedback stabilization of case $\gamma_1 = 0.5,\ \gamma_2 = 0.5$

and Hauser (1996). Taking the interior product of a non vanishing two-form with respect to the vector field defining the system, we first obtained a (possibly) non-closed characteristic one-form for the system. Constructing a locally defined homotopy operator on a star-shaped domain centered at the desired equilibrium point, we presented how to decompose locally the obtained form into an exact and an anti-exact one-forms. From (Hudon et al., 2008), we know that the exact part is associated to a dissipative (stable) potential. The obtained anti-exact form is associated to a non dissipative potential which generated tangential dynamics that do not contribute to the value of the dissipative potential on the star-shaped domain. However, using a pre-defined feedback controller to make this error one-form exact, it was shown, using the four-tank system example, that the procedure enables us to construct a stabilizing control. Future research will focus on the limitations of the technique, especially cases where the controller information does not appear in the expression of the anti-exact form, for example the nonisothermal CSTR system presented in (Ramírez et al., 2009).

REFERENCES

A. Banaszuk and J. Hauser. Approximate feedback linearization. *SIAM Journal on Control and Optimization*, 34(5):1533–1554, 1996.

D. Cheng, S Spurgeon, and J. Xiang. On the development of generalized Hamiltonian realizations. In *Proceedings of the 39th IEEE Conference on Decision and Control*, pages 5125–5130, 2000.

D. Cheng, A. Astolfi, and R. Ortega. On feedback equivalence to port-controlled Hamiltonian systems. *Systems and Control Letters*, 54:911–917, 2005.

D.G.B. Edelen. *Applied Exterior Calculus*. Wiley, New York, NY, 1985.

N. Hudon, K. Höffner, and M. Guay. Equivalence to dissipative Hamiltonian realization. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 3163–3168, 2008.

K.H. Johansson. The quadruple-tank process: A multivariable laboratory process with an adjustable zero. *IEEE Transactions on Control Systems Technology*, 8(3):456–465, 2000.

J. Johnsen and F. Allgöwer. Interconnection and damping assignment passivity-based of a four-tank system. In F. Bullo and K. Fugimoto, editors, *Lagrangian and Hamiltonian Methods for Nonlinear Control 2006*, volume 366 of *Lectures Notes in Control and Information Science*, pages 111–122, Berlin, 2007. Springer-Verlag.

B. Maschke, R. Ortega, and A.J. van der Schaft. Energy-based Lyapunov functions for forced Hamiltonian systems with dissipation. *IEEE Transactions on Automatic Control*, 45(8):1498–1502, 2000.

R. Ortega, A. Astolfi, G. Bastin, and H/ Rodrigues-Cortes. Output feedback control of food-chain systems. In H. Nijmeijer and T.I. Fossen, editors, *New Directions in Nonlinear Observer design*, volume 244 of *Lectures Notes in Control and Information Science*, pages 291–310, London, 1999. Springer-Verlag.

R. Ortega, A.J. van der Schaft, B. Maschke, and G. Escobar. Interconnection and damping assignment passivity-based control of port-controlled Hamiltonian systems. *Automatica*, 38:585–596, 2002.

I. Otero-Muras, G. Szederkényi, A.A. Alonso, and K.M. Hangos. Local dissipative Hamiltonian description of reversible reaction networks. *Systems and Control Letters*, 57:554–560, 2008.

H. Ramírez, D. Sbarbaro, and R. Ortega. On the control of non-linear processes: An IDA-PBC approach. *Journal of Process Control*, 19:405–414, 2009.

A.J. van der Schaft. *L_2-Gain and Passivity Techniques in Nonlinear Control*. Springer-Verlag, London, 2nd edition, 2000.

Y. Wang, C. Li, and D. Cheng. Generalized Hamiltonian realization of time-invariant nonlinear systems. *Automatica*, 39:1437–1443, 2003.

Y. Wang, D. Cheng, and S.S. Ge. Approximate dissipative Hamiltonian realizations and construction of local Lyapunov functions. *Systems and Control Letters*, 56: 141–149, 2007.

Robust Nonlinear Model Predictive Control using Volterra Models and the Structured Singular Value (μ)

Rosendo Díaz-Mendoza, Hector Budman

Department of Chemical Engineering, University of Waterloo, Waterloo, ON, Canada, N2L 3G1

Abstract: A methodology is proposed for designing robust nonlinear model predictive controllers based on a Volterra series model with uncertain coefficients. The objective function of the on-line optimization is formulated in terms of a Structured Singular Value (μ). The proposed formulation considers a penalty on the manipulated variables actions and manipulated variables and terminal condition constraints.

Keywords: Model predictive control, nonlinear control, robust control, structured singular value.

1. INTRODUCTION

Linear model predictive control is a widely accepted control strategy in the chemical industry. Many theoretical studies and industrial applications of linear MPC have been reported elsewhere (Qin and Badgwell, 2003). On the other hand, the nonlinear behaviour of chemical processes has motivated researchers and practitioners to consider predictive control strategies based on nonlinear process models referred to as nonlinear model predictive controllers (NMPC) (Findeisen and Allgöwer, 2002).

Some of the challenging requirements related to the industrial implementation of NMPC are: (1) a reliable nonlinear model of the process is needed that can be effectively used for real time control and (2) ensuring robustness to model error.

Both first principles as well as empirical input-output models have been used in the past for nonlinear predictive control strategies. Although first principles models have the advantage of formally satisfying basic energy and mass balances of the process, they are often too complex for real time control and their structure is generally not amenable for formal robust analysis and design. NMPC strategies based on empirical models such as Hammerstein and Volterra series (Hérnandez and Arkun, 1993; Maner et al., 1996; Parker and Doyle III, 2001) have been reported but their robustness with respect to model error have not been thoroughly studied.

The need to address robustness arises from the fact that the models used for predictive control are never exact. Although a good amount of research work has been conducted on robustness of linear predictive controllers, the robustness of nonlinear predictive controllers has not been extensively studied. The lack of robustness guarantees is currently perceived as one of the key obstacles for wide industrial acceptance of NMPC strategies (Nagy and Braatz, 2003a).

The current work investigates the design of a robust NMPC algorithm based on an empirical Volterra series model. Volterra series models have been shown to efficiently describe general nonlinear systems (Schetzen, 1980). A key idea for this study is that based on the Volterra model it is possible to formulate the robust predictive control problem as a μ-Structured Singular Value test that can be used on-line to calculate optimal control actions. The μ (Structured Singular Value, Doyle, 1982) norm is used, at each sampling instant, to calculate a bound on the norm of a vector containing both output and input predictions along a predefined prediction time horizon in the presence of disturbances and uncertainty in the Volterra model coefficients. Then, the calculated bound is minimized with respect to the optimal control actions to be sent to the process.

The paper is organized as follows. In section 2 the formulation of the μ test and the optimization problem based on the calculated bound is presented. In the same section terminal conditions to enforce stability as well as conditions to enforce manipulated variable constraints are also presented. Section 3 presents two case studies and conclusions are presented in Section 4. Mathematical details are presented in Appendix A.

2. METHODOLOGY

2.1 Model Predictive Control

MPC minimizes a cost function that considers the future errors whit respect to the manipulated variables. For simplicity of notation a single input single output (SISO) case is shown but the formulation can be easily extended to the multivariable case. Considering that y^{pr} is the predicted value of the controlled variable and y^{sp} is the controlled variable set point, a vector of predictions can be written as follows:

$$\mathbf{Y} = \begin{bmatrix} y^{pr}(k_0) + d(k_0) - y^{sp}(k_0) \\ \vdots \\ y^{pr}(k_0 + p) + d(k_0 + p) - y^{sp}(k_0 + p) \end{bmatrix} \quad (1)$$

where k_0 is the initial sampling instant, p is the prediction horizon and d is a measured disturbance. The objective function of the controller proposed in this work minimizes the maximum absolute value of each element of the **Y** vector with respect to the manipulated variables u as follows:

$$\min_{\text{wrt } u(k_0),\ldots,u(k_0+m)} \left\| \mathbf{Y} \right\|_\infty \tag{2}$$

where m is the control horizon. In principle norms other than the infinity norm of the output may be considered in the formulation but are beyond the scope of the current study. It will be also shown in subsections 2.4 to 2.6 that the vector **Y** in (2) may be augmented by additional variables, other than predicted outputs, to enforce a terminal condition and manipulated variables constraints. The following subsection discusses the Volterra models used to calculate the prediction **Y**.

2.2 Volterra series

The general structure of a Volterra series model is given as follows:

$$y^{pr}(k_0) = \sum_{\sigma_1=1}^{\infty} \sum_{\sigma_i=1}^{\infty} h_i(\sigma_1,\ldots,\sigma_i) u(k_0-\sigma_1) u(k_0-\sigma_i) \tag{3}$$

where u is the manipulated variable, y is the controlled variable and h_i are the coefficients of the Volterra series. For practical purposes the series is truncated and the resulting expression is

$$\begin{aligned} y^{pr}(k_0) = \; & h_0 + \\ & \sum_{n=1}^{N} \sum_{i_1=1}^{M} \cdots \sum_{i_n=1}^{M} \left(h_n(i_1,\ldots,i_n) \times \right. \\ & \left. u(k_0-i_i) u(k_0-i_n) \right) \end{aligned} \tag{4}$$

without loss of generality it can be consider that h_0=0. For example, for N=2, the value of the controlled variable is

$$\begin{aligned} y^{pr}(k_0) = \; & \sum_{n=1}^{M} h_n^L u(k_0-n) + \\ & \sum_{i=1}^{M} \sum_{j=i}^{M} h_{i,j}^{NL} u(k_0-i) u(k_0-j) \end{aligned} \tag{5}$$

where M is the memory of the system. The linear $h_n{}^L$, and nonlinear $h_{i,j}{}^{NL}$, Volterra series coefficients can be obtained by least squares regression using process input-output data by imposing an appropriate input sequence. For a system with polynomial degree N, it has been shown that is necessary to use a N+1 level pseudorandom multilevel sequence (Nowak and Van Veen, 1994). Confidence intervals for the coefficients, to be used in the calculations as uncertainty bounds associated to these coefficients, can be obtained using least squares regression.

2.3 Calculation of the worst predicted output

The worst predicted output calculation can be performed by a Structured Singular Value (SSV) test (Nagy and Braatz, 2003b). The main motivation to use the SSV test is that it allows finding the worst $\left\| \mathbf{Y} \right\|_\infty$ when uncertainty in the Volterra coefficients is considered. Accordingly, (5) can be modified to include parameter uncertainty as follows where $h_n{}^L$ and $h_{i,j}{}^{NL}$ are the nominal value of the coefficients and $\delta h_n{}^L$ and $\delta h_{i,j}{}^{NL}$ are the uncertainty associated to the coefficients:

$$\begin{aligned} y^{pr}(k_0) = \; & \sum_{n=1}^{M} \left[h_n^L + \delta h_n^L \right] u(k_0-n) + \\ & \sum_{i=1}^{M} \sum_{j=i}^{M} \left[h_{i,j}^{NL} + \delta h_{i,j}^{NL} \right] u(k_0-i) u(k_0-j) + \\ & w(k_0) \end{aligned} \tag{6}$$

w is a feedback term that considers the current difference between the actual process output and the predicted output:

$$w(k_0) = y^{real}(k_0-1) - y^{pr}(k_0-1) \tag{7}$$

By selecting an appropriate interconnection matrix **M** and uncertainty block structure **Δ**, the worst value of a variable in the presence of model error can be calculated by the following SSV test (Braatz et al., 1994; Nagy and Braatz, 2003b)

$$\max_{\text{wrt } \delta h_i^L, \delta h_{i,j}^{NL}, w} \left\| \mathbf{Y} \right\|_\infty \geq k_{ssv} \Leftrightarrow \mu_\Delta(\mathbf{M}) \geq k_{ssv} \tag{8}$$

Thus, a bound on the worst deviation of $\left\| \mathbf{Y} \right\|_\infty$, i.e. the norm of the prediction vector can be obtained by the following skew μ problem:

$$\max_{\text{wrt } \delta h_i^L, \delta h_{i,j}^L, w} \left\| \mathbf{Y} \right\|_\infty = \max_{\substack{\text{wrt } k_{ssv} \\ \text{st } \mu_\Delta(\mathbf{M}) \geq k_{ssv}}} (k_{ssv}) \tag{9}$$

A key idea in (9) is that the feedback term in (7) is also treated as an uncertainty and the maximization in (9) is carried out with respect to both this feedback error and the uncertainties in coefficients. Accordingly, the uncertainty block **Δ** is as follows:

$$\mathbf{\Delta} = \text{diag}(\mathbf{\Delta}_1, \mathbf{\Delta}_2, \mathbf{\Delta}_3) \tag{10}$$

where $\mathbf{\Delta}_3$ is a complex scalar square matrix of dimensions $p \times p$ related to performance and $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$ are real scalar square matrices related to the uncertainty in feedback and Volterra series coefficients respectively with the following dimensions:

$$\mathbf{\Delta}_1 = \left(p + 2\sum_{i=1}^{p} i + \sum_{i=1}^{p-1} \sum_{j=1}^{i} j \right) \times \left(p + 2\sum_{i=1}^{p} i + \sum_{i=1}^{p-1} \sum_{j=1}^{i} j \right) \tag{11}$$

$$\mathbf{\Delta}_2 = \text{diag}\left[\mathbf{\Delta}_{2_1} \quad \ldots \quad \mathbf{\Delta}_{2_p} \right] \tag{12}$$

$$\mathbf{\Delta}_{2_i} = \left[p \times p \quad \ldots \quad 1 \times 1 \right]^{\mathrm{T}} \; i \leq 2 \tag{13}$$

$$\mathbf{\Delta}_{2_i} = \left[(p+2-i) \times (p+2-i) \quad \ldots \quad 1 \times 1 \right]^{\mathrm{T}} \; i > 2 \tag{14}$$

The problem stated in (8) and (9) can be used within the predictive control problem defined in (2) as follows:

$$\min_{u(k_0),\ldots,u(k_0+m)} \left[\max_{\text{wrt } k_{ssv}}\left(k_{ssv}\right)\right] \quad \text{st } \mu_{\mathbf{\Delta}}(\mathbf{M}) \geq k_{ssv} \tag{15}$$

The vector **Y** can be modified, as mentioned in section 2.1 to include additional terms as follows: (1) a penalty term to prevent an excessive movement of the manipulated variables, (2) manipulated variables to enforce constraints and (3) a terminal condition to ensure convergence. These terms are explained in the following subsections.

2.4 Manipulated variables movements penalization

Define:

$$\mathbf{y}^{\Delta u} = \begin{bmatrix} W_1^{\Delta u}\left[u(k_0)-u(k_0-1)\right] \\ \vdots \\ W_m^{\Delta u}\left[u(k_0+m)-u(k_0+m-1)\right] \end{bmatrix} \tag{16}$$

Redefining: $\mathbf{Y}=[\,\mathbf{y}^{pr}\ \mathbf{y}^{\Delta u}\,]^{\mathrm{T}}$ it is ensured by (9) that the elements of $\mathbf{y}^{\Delta u}$ satisfy $\max(y_i^{\Delta u}) \leq k_{ssv}$ for $i=1,\ldots,m$. Thus, the maximum weighted manipulated variable movement is bounded at each sampling instant by k_{ssv}.

2.5 Manipulated variables constraints

Define:

$$\mathbf{y}^{uc} = \left[k_{ssv}\frac{u(k_0)}{u_{\max}(k_0)} \quad \ldots \quad k_{ssv}\frac{u(k_0+m)}{u_{\max}(k_0+m)} \right]^{\mathrm{T}} \tag{17}$$

Redefining: $\mathbf{Y}=[\,\mathbf{y}^{pr}\ \mathbf{y}^{\Delta u}\ \mathbf{y}^{uc}\,]^{\mathrm{T}}$ it is ensured by (9): that the elements of $\mathbf{y}^{uc}$ satisfy $\max(k_{ssv}u(i)/u_{\max}(i)) \leq k_{ssv}$ for $i=k_0,\ldots,k_0+m$ which can be simplified to: $\max(u(i)) \leq u_{\max}(i)$ for $i=k_0,\ldots,k_0+m$. Thus, the manipulated variables are bounded at each sampling instant by $u_{\max}(i)$ for $i=k_0,\ldots,k_0+m$.

2.6 MPC terminal condition

A terminal condition is used to ensure that at steady state the predicted output stays within a neighborhood ε near the origin (Chen and Allgöwer, 1998). Although not shown here for brevity, it can be shown that the use of the terminal condition together with the manipulated variables constraints ensures stability providing that the terminal condition is feasible with respect to constraints. Define:

$$\begin{aligned} \mathbf{y}^{tc} = \ & \frac{k_{ssv}}{\varepsilon}\left[u(k_0+m)\sum_{i=1}^{p}\left(h_i^L+\delta h_i^L\right)\right] + \\ & \frac{k_{ssv}}{\varepsilon}\left[u^2(k_0+m)\sum_{i=1}^{p}\sum_{j=i}^{p}\left(h_{i,j}^{NL}+\delta h_{i,j}^{NL}\right)\right] \end{aligned} \tag{18}$$

ε can be selected by the user but a smaller value results in more conservative control. Redefining: $\mathbf{Y}=[\mathbf{y}^{pr}\ \mathbf{y}^{\Delta u}\ \mathbf{y}^{uc}\ \mathbf{y}^{tc}]^{\mathrm{T}}$ it is ensured by (9) that

$$\max\left[\begin{array}{l} \frac{k_{ssv}}{\varepsilon}\left(u(k_0+m)\sum_{i=1}^{p}\left(h_i^L+\delta h_i^L\right)\right) + \\ \frac{k_{ssv}}{\varepsilon}\left(u^2(k_0+m)\sum_{i=1}^{p}\sum_{j=i}^{p}\left(h_{i,j}^{NL}+\delta h_{i,j}^{NL}\right)\right) \end{array}\right] \leq k_{ssv} \tag{19}$$

which can be simplified to:

$$\max\left[\begin{array}{l} u(k_0+m)\sum_{i=1}^{p}\left(h_i^L+\delta h_i^L\right)+ \\ u^2(k_0+m)\sum_{i=1}^{p}\sum_{j=i}^{p}\left(h_{i,j}^{NL}+\delta h_{i,j}^{NL}\right) \end{array}\right] \leq \varepsilon \tag{20}$$

Details on the construction of **M** for a Volterra series model NMPC strategy considering the additional terms of subsections 2.4, 2.5 and 2.6 can be found on Appendix A at the end of this paper.

3. CASE STUDIES

Different case studies are presented to show the more important features of the proposed algorithm. For simplicity a SISO case is presented where an approximated Volterra model describing the effect of coolant temperature on reactor concentration for a CSTR is as follows (Gao, 2004):

$$\begin{aligned} y(k_0) = \ & h_1^L u(k_0)+h_2^L u(k_0-1)+h_3^L u(k_0-2)+ \\ & h_{1,1}^{NL}u^2(k_0)+h_{1,2}^{NL}u(k_0)u(k_0-1)+ \\ & h_{1,3}^{NL}u(k_0)u(k_0-2)+h_{2,2}^{NL}u^2(k_0-1)+ \\ & h_{2,3}^{NL}u(k_0-1)u(k_0-2)+h_{3,3}^{NL}u^2(k_0-2) \end{aligned} \tag{21}$$

$h_1^{\mathrm{L}} = 0.2835$, $h_2^{\mathrm{L}} = 0.1445$, $h_3^{\mathrm{L}} = 0.0594$, $h_{1,1}^{\mathrm{NL}} = -0.0072$, $h_{1,2}^{\mathrm{NL}} = -0.049$, $h_{1,3}^{\mathrm{NL}} = -0.0281$, $h_{2,2}^{\mathrm{NL}} = -0.0379$, $h_{2,3}^{\mathrm{NL}} = -0.017$, $h_{3,3}^{\mathrm{NL}} = -0.0081$. The MPC prediction and control horizons are $p=3$ and $m=2$.

The first study is intended to illustrate the possibility of tuning the proposed algorithm through the value of $W_1^{\Delta u}$, i.e. the weight used to penalize manipulated variables from sampling instant (k_0–1) to sampling instant (k_0). The response of the process to a pulse disturbance is studied and the set point is equal to zero. The disturbance is as follows: from $0<k_0\leq 20$ $d=5$, then from $20<k_0\leq 40$ $d=0$. For this case it is considered that there is no uncertainty in the Volterra model coefficients.

Figures 1 and 2 show that the weight imposed on the movement on the manipulated variables can be effectively used to tune the closed loop response. To illustrate the significance of the nonlinear terms, a simulation is carried on with a controller based solely on the linear part of the Volterra model. The results (dotted line in Figures 1 and 2) illustrate that the nonlinear model based controller provides as expected, a better and more consistent performance than the linear model based one.

To illustrate the constraint handling capabilities of the algorithm, the response to a pulse disturbance is studied. The disturbance is: for $0<k_0\leq 5$ $d=5$, for $5<k_0\leq 10$ $d=50$ and for

$10<k_0\leq 15$ d=5. The value of the manipulated variable is restricted to $|u(k_0)| \leq 5.5$ and $W_1^{\Delta u}$=2. For this case it was considered that the Volterra series coefficients are known accurately, i.e. there is no model uncertainty. It can be seen from Figure 3 that the controller keeps the value of the manipulated variable within the allowed limits.

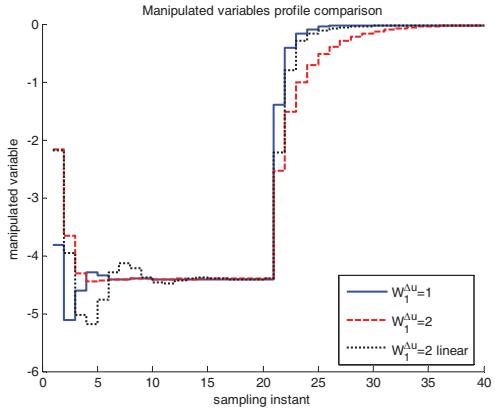


Fig. 1. Manipulated variable profile for case study 1.

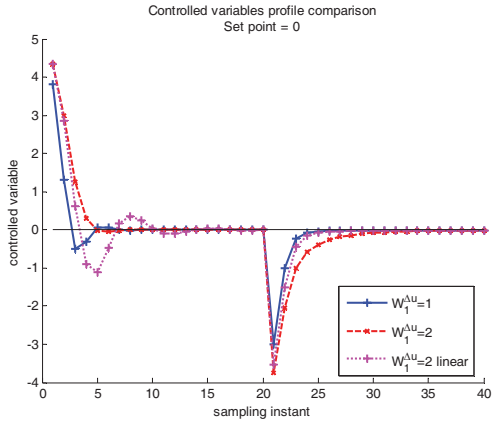


Fig. 2. Controlled variable profile for case study 1.

A key feature of the proposed MPC is that it allows considering that the Volterra series coefficients are not exactly known. To illustrate this feature of the algorithm it is assumed that certain coefficients are uncertain as follows: $h_1^{L} = 0.2551\pm0.0383$, $h_{2,2}^{NL} = -0.0360\pm0.0072$ and $h_{3,3}^{NL} = -0.0089\pm0.0018$. In this case the Volterra series coefficients of the plant are the same as those used for case study 1 and 2. Furthermore, the disturbance affecting the process is the same as that of case study 1. Figures 4 and 5 show the manipulated and controlled variable profile when $W_1^{\Delta u}$=2. The response obtained with the uncertain model MPC is more oscillatory but still acceptable. The figures show that the control variable converges to a value very close to zero and the manipulated variables are kept within limits. The small offset observed in the manipulated variable with respect to u=0 arises from the requirement of the terminal condition in the presence of model uncertainty.

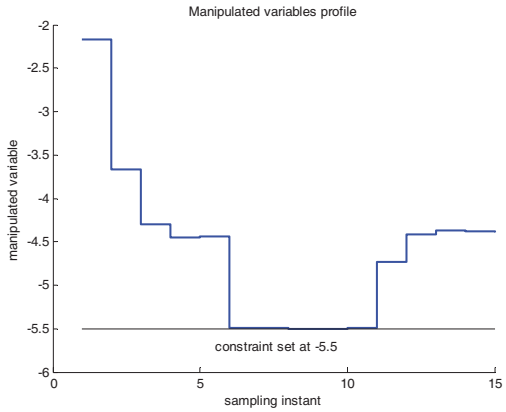


Fig. 3. Manipulated variable profile for case study 2.

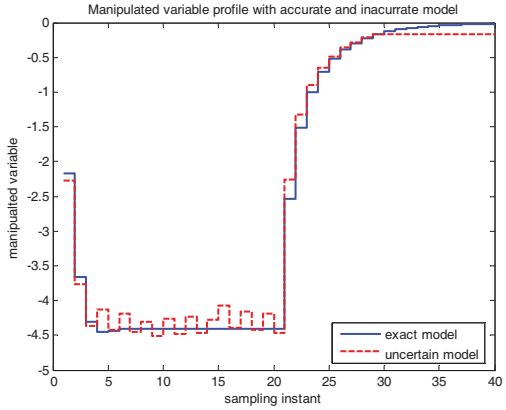


Fig. 4. Manipulated variable profile for case study 3.

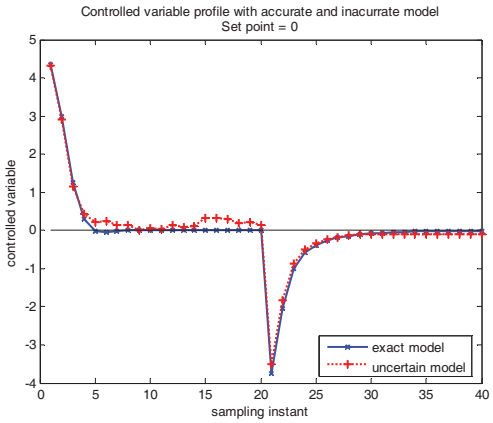


Fig. 5. Controlled variable profile for case study 3.

4. CONCLUSIONS

A novel robust NMPC controller based on a Volterra model was presented. The methodology uses μ analysis to calculate, for an uncertain plant model, the worst possible norm of a vector of inputs and outputs. The interconnection matrix can include terms to account for manipulated variables movement weighting, manipulated variables constraints and robust stability properties enforced through a terminal condition. The application of this technique to MIMO problems is currently being investigated.

ACKNOWLEDGMENTS

The financial support provided by Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

REFERENCES

Braatz, R. D., Young, P. M., Doyle, J. C., and Morari, M. (1994). Computational complexity of μ calculation, *IEEE Transactions on Automatic Control*, 39 (5), 1000–1002,

Chen, H., and Allgöwer, F. (1998). A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability, *Automatica,* 34 (10), 1205–1217.

Doyle, J., (1982). Analysis of feedback systems with structured uncertainties, *IEE Proceedings D Control Theory & Applications,* 129 (6), 242–250.

Findeisen, R., and Allgöwer, F. (2002). An introduction to nonlinear model predictive control. In Bram de Jager and Hans Zwart (ed.), *21st Benelux Meeting on Systems and Control,* 119–141. Technische Universiteit Eindhoven, Veldhoven.

Gao, J. (2004). *Robust control design of gain-scheduled controllers for nonlinear processes*. PhD Thesis University of Waterloo, Waterloo.

Hérnandez, E. and Arkun, Y. (1993). Control of nonlinear systems using polynomial ARMA models, *AIChE Journal,* 39 (3), 446–460.

Maner, B. R., Doyle III, F. J., Ogunnaike, B. A., and Pearson, R. K. (1996). Nonlinear model predictive control of a simulated multivariable polymerization reactor using second-order Volterra models, *Automatica,* 32 (9), 1285–1301.

Nagy, Z. K., and Braatz, R. D. (2003a). Robust nonlinear model predictive control for batch processes, *AIChE Journal,* 49 (7), 1776–1786.

Nagy, Z. K., and Braatz, R. D. (2003b). Worst-case and distribution analysis of finite-time control trajectories for nonlinear distributed parameter systems, *IEEE Transactions on Control Systems Technology,* 11 (5), 694–704.

Nowak, R. D., and Van Veen, B. D. (1994). Random and pseudorandom inputs for Volterra filter identification, *IEEE Transactions on Signal Processing,* 42 (8), 2124–2135.

Parker, R. S. and Doyle III, F. J. (2001). Optimal control of a continuous bioreactor using an empirical nonlinear model, *Industrial Engineering & Chemistry Research*, 40 (8), 1939–1951.

Qin, S. J., and Badgwell, T. A. (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, 11 (7), 733–764.

Schetzen, M. (1980). *The Volterra & Wiener theories of nonlinear systems.* Robert E. Krieger, Florida.

Appendix A. CONSTRUCTION OF THE MATRIX **M**

The use of an appropriate interconnection matrix **M** allows quantifying the effect that an input has on the system's output in the presence of uncertainty through a linear fractional transformation (LFT). If **M** is built according to the following structure

$$\mathbf{M} = \left[\begin{array}{c|c} \mathbf{M}_{11}^{\mathrm{LFT}} & \mathbf{M}_{12}^{\mathrm{LFT}} \\ \hline \mathbf{M}_{21}^{\mathrm{LFT}} & \mathbf{M}_{22}^{\mathrm{LFT}} \end{array}\right] \tag{22}$$

the effect that the input has on the output in the presence of uncertainty is:

$$\mathbf{Y}(k) = \left[\mathbf{M}_{21}^{\mathrm{LFT}}\boldsymbol{\Delta}\left[\mathbf{I} - \mathbf{M}_{11}^{\mathrm{LFT}}\boldsymbol{\Delta}\right]^{-1}\mathbf{M}_{12}^{\mathrm{LFT}} + \mathbf{M}_{22}^{\mathrm{LFT}}\right]\begin{bmatrix}\boldsymbol{w}(k)\\ 0\end{bmatrix} \tag{23}$$

where $\boldsymbol{w}(k) = [w(k), \ldots, w(k)]^{\mathrm{T}}_{p}$. The interconnection matrix **M** that considers manipulated variables movement penalization, manipulated variables constraints and terminal condition has the following structure:

$$\mathbf{M} = \left[\begin{array}{ccc|c} \mathbf{M}_{11} & \mathbf{M}_{12} & \mathbf{M}_{13} & \mathbf{M}_{14} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \mathbf{M}_{24} \\ \mathbf{M}_{31} & \mathbf{M}_{32} & \mathbf{M}_{33} & \mathbf{M}_{34} \\ \hline \mathbf{M}_{41} & \mathbf{M}_{42} & \mathbf{M}_{43} & \mathbf{M}_{44} \end{array}\right] \tag{24}$$

In (24) $\mathbf{M}_{11}$, $\mathbf{M}_{12}$, $\mathbf{M}_{13}$, $\mathbf{M}_{21}$, $\mathbf{M}_{22}$, $\mathbf{M}_{23}$, $\mathbf{M}_{31}$, $\mathbf{M}_{33}$, $\mathbf{M}_{34}$ and $\mathbf{M}_{44}$ are matrices of appropriate dimensions that have all its elements equal to zero. $\mathbf{M}_{14}$, $\mathbf{M}_{24}$, $\mathbf{M}_{32}$, $\mathbf{M}_{41}$, $\mathbf{M}_{42}$ and $\mathbf{M}_{43}$ are defined as:

$$\mathbf{M}_{14} = \mathrm{diag}\left[\mathbf{D} \quad \mathbf{A}\right] \tag{25}$$

$$\mathbf{M}_{24} = \mathrm{diag}\left[\begin{bmatrix}\mathbf{U}^{L}\\ \mathbf{U}^{CP}\end{bmatrix} \quad \mathbf{B}\right] \tag{26}$$

$$\mathbf{M}_{32} = \mathbf{E} \tag{27}$$

$$\mathbf{M}_{41} = \mathrm{diag}\left[\mathbf{I}_{p} \quad \mathbf{C}\right] \tag{28}$$

$$\mathbf{M}_{42} = \begin{bmatrix}\mathbf{H}^{L} & 0 & \mathbf{H}^{CP}\\ 0 & \mathbf{F} & 0\end{bmatrix} \tag{29}$$

$$\mathbf{M}_{43} = \left[\mathbf{V}^{Lac} \quad \mathbf{V}^{NLac} \quad \mathbf{V}^{CPac}\right] \tag{30}$$

The rest of the matrices are defined as follows:

$$\mathbf{A} = \mathrm{diag}\left[\mathbf{A}_{11}, \mathbf{A}_{12}, \mathbf{A}_{13}\right] \tag{31}$$

$$\mathbf{A}_{11} = k_{ssv}\, d\left(k_0 + p\right) \tag{32}$$

$$\mathbf{A}_{12} = k_{ssv}\mathrm{diag}\left[\begin{bmatrix} u(k_0-1) \\ u(k_0) \end{bmatrix} \quad \dots \quad \begin{bmatrix} u(k_0+m-1) \\ u(k_0+m) \end{bmatrix}\right] \tag{33}$$

$$\mathbf{A}_{13} = k_{ssv}\mathrm{diag}\left[u(k_0),\dots,u(k_0+m)\right] \tag{34}$$

$$\mathbf{D} = k_{ssv}\mathrm{diag}\begin{bmatrix} d(k_0)+y^{sp}(k_0) \\ \vdots \\ d(k_0+p)+y^{sp}(k_0+p) \end{bmatrix} \tag{35}$$

$$\mathbf{U}^{L} = k_{ssv}\begin{bmatrix}\mathbf{U}_1^{L} & \dots & \mathbf{U}_l^{L} & \dots & \mathbf{U}_p^{L}\end{bmatrix}^{\mathrm{T}} \tag{36}$$

$$\mathbf{U}_l^{L} = \begin{bmatrix} 0_{((p+1-l)\times(l-1))} & u(x)\mathbf{I}_{p+1-l} \\ 0_{((p+1-l)\times(l-1))} & u(x)^2\mathbf{I}_{p+1-l} \end{bmatrix} \tag{37}$$

$$\mathbf{U}^{CP} = k_{ssv}\begin{bmatrix}\mathbf{U}_{1,1}^{CP} & \mathbf{U}_{1,2}^{CP} & \dots & \mathbf{U}_{i,j-1}^{CP} & \mathbf{U}_{i,j}^{CP}\end{bmatrix} \tag{38}$$

$$i = 1,\dots,p-1,\ j = 1,\dots,p-i$$

$$\mathbf{U}_{i,j}^{CP} = \begin{bmatrix}\mathbf{U}_{i,j\,1,1}^{CP} & \mathbf{U}_{i,j\,1,2}^{CP}\end{bmatrix} \tag{39}$$

$$\mathbf{U}_{i,j\,1,1}^{CP} = 0_{((p+1-i-j)\times(i+j-1))} \tag{40}$$

$$\mathbf{U}_{i,j\,1,2}^{CP} = u(i+j)u(j)\mathbf{I}_{p+1-i-j} \tag{41}$$

$$\mathbf{B} = \begin{bmatrix} k_{ssv}u(k_0+m) & k_{ssv}u^2(k_0+m)\end{bmatrix}^{\mathrm{T}} \tag{42}$$

$$\mathbf{C} = \begin{bmatrix} \frac{k_{ssv}}{\varepsilon}, \mathrm{diag}\left[\begin{bmatrix} W_i^{\Delta u} & -W_i^{\Delta u}\end{bmatrix}, \frac{k_{ssv}}{u_{\max}(i)}\right] \end{bmatrix} \quad i = 1,\dots,m \tag{43}$$

$$\mathbf{H}^{L} = \begin{bmatrix}\mathbf{H}_1^{L},\dots,\mathbf{H}_l^{L},\dots,\mathbf{H}_p^{L}\end{bmatrix} \tag{44}$$

$$\mathbf{H}_l^{L} = \begin{bmatrix} \mathbf{H}_{l\,1,1}^{L} & \mathbf{H}_{l\,1,2}^{L} \\ \mathbf{H}_{l\,2,1}^{L} & \mathbf{H}_{l\,2,2}^{L} \end{bmatrix} \tag{45}$$

$$\mathbf{H}_{l\,1,1}^{L} = \mathbf{H}_{l\,1,2}^{L} = 0_{((l-1)\times(p+1-l))} \tag{46}$$

$$\mathbf{H}_{l\,2,1}^{L} = \mathrm{diag}\left(h_1^{L},\dots,h_{p+1-l}^{L}\right) \tag{47}$$

$$\mathbf{H}_{l\,2,2}^{L} = \mathrm{diag}\left(h_{1,1}^{NL},\dots,h_{p+1-l,p+1-l}^{NL}\right) \tag{48}$$

$$\mathbf{H}^{CP} = \begin{bmatrix}\mathbf{H}_{1,1}^{CP},\mathbf{H}_{1,2}^{CP},\dots,\mathbf{H}_{i,j-1}^{CP},\mathbf{H}_{i,j}^{CP}\end{bmatrix} \begin{matrix} i = 1,\dots,p-1 \\ j = 1,\dots,p-i \end{matrix} \tag{49}$$

$$\mathbf{H}_{i,j}^{CP} = \begin{bmatrix} 0_{((i+j-1)\times(p+1-i-j))} \\ \mathrm{diag}\left(h_{1,i+1}^{NL},\dots,h_{p+1-i-j,p+1-j}^{NL}\right) \end{bmatrix} \tag{50}$$

$$\mathbf{F} = \begin{bmatrix} \frac{k_{ssv}}{\varepsilon}\sum_{i=1}^{p} h_i^{L} & \frac{k_{ssv}}{\varepsilon}\sum_{i=1}^{p}\sum_{j=i}^{p} h_{i,j}^{NL} \end{bmatrix} \tag{51}$$

$$\mathbf{V}^{Lac} = \begin{bmatrix}\mathbf{V}_1^{Lac},\dots,\mathbf{V}_l^{Lac},\dots,\mathbf{V}_p^{Lac}\end{bmatrix} \tag{52}$$

$$\mathbf{V}_l^{Lac} = \begin{bmatrix} 0_{((l-1)\times(p+2-l))} & \delta h_x^{L}\mathbf{I}_{p+2-l}\end{bmatrix}^{\mathrm{T}} \tag{53}$$

$$\mathbf{V}^{NLac} = \begin{bmatrix}\mathbf{V}_1^{NLac},\dots,\mathbf{V}_l^{NLac},\dots,\mathbf{V}_p^{NLac}\end{bmatrix} \tag{54}$$

$$\mathbf{V}_l^{NLac} = \begin{bmatrix} 0_{((l-1)\times(p+2-l))} & \delta h_{x,x}^{NL}\mathbf{I}_{p+2-l}\end{bmatrix}^{\mathrm{T}} \tag{55}$$

$$\mathbf{V}^{CPac} = \begin{bmatrix}\mathbf{V}_{1,1}^{CPac},\mathbf{V}_{1,2}^{CPac},\dots,\mathbf{V}_{i,j}^{CPac}\end{bmatrix} \begin{matrix} i = 1,\dots,p-1 \\ j = 1,\dots,p-i \end{matrix} \tag{56}$$

$$\mathbf{V}_{i,j}^{CPac} = \begin{bmatrix} 0_{((i+j-1)\times(p+3-i-j))} & \delta h_{i,j}^{NL}\mathbf{I}_{p+3-i-j}\end{bmatrix}^{\mathrm{T}} \tag{57}$$

In order to obtain the matrix **E** it is necessary to build a column vector that contains the Volterra series coefficients according to the following structure:

$$\mathbf{VE} = \begin{bmatrix}\mathbf{VE}_1 & \dots & \mathbf{VE}_{p+1}\end{bmatrix}^{\mathrm{T}} \tag{58}$$

The matrix **VE** has the following dimensions:

$$\mathbf{VE} = \left(p + p + \sum_{i=1}^{p-1} p - i\right)\times 1 \tag{59}$$

The rule to construct **VE** is as follows:

$$\mathbf{VE}_1 = \begin{bmatrix} h_1^{L} & \dots & h_p^{L}\end{bmatrix}^{\mathrm{T}} \tag{60}$$

$$\mathbf{VE}_2 = \begin{bmatrix} h_{1,1}^{NL} & \dots & h_{p,p}^{NL}\end{bmatrix}^{\mathrm{T}} \tag{61}$$

$$\mathbf{VE}_i = \begin{bmatrix} h_{1,i-1}^{NL} & \dots & h_{p+2-i,p}^{NL}\end{bmatrix}^{\mathrm{T}} \quad i \geq 3 \tag{62}$$

Finally the matrix **E** is constructed according to the following program:

```
for i = 1,...,p + p + Σ_{q=1}^{p-1} p - q
    for ir = 1,...,p
        for ic = 1,...,p + p + Σ_{iq=1}^{p-1} p - iq
            if [H^L H^CP]_{ir,ic} = VE_{i,1}
                index_{i,1} = ic                                   (63)
            end
        end
    end
end
```

The elements of the matrix **E** are zero except the following:

$$\text{for } ir = 1,\dots,2\sum_{i=1}^{p} i + \sum_{i=1}^{p-1}\sum_{j=1}^{i} j$$

$$\mathbf{E}_{ir,index(ir,1)} = k_{ssv} \tag{64}$$

end

OUTPUT-FEEDBACK DISSIPATIVE CONTROL OF EXOTHERMIC CONTINUOUS REACTORS

A. Schaum * **J.A. Moreno** * **J. Alvarez** ** **J. Diaz-Salgado** *

* *Instituto de Ingeniería, Universidad Nacional Autónoma de México (e-mail: {ASchaum,JMorenoP,JDiazS}@ii.unam.mx).*
** *Departamento de Procesos e Hidraulica, Universidad Autónoma Metropolitana de México (e-mail: jac@xanum.uam.mx)*

Abstract: The problem of controlling a (possibly open-loop unstable) continuous exothermic reactor with temperature measurements and manipulation of reactant feed and heat exchange rates is addressed within a passivity-dissipativity framework. The combination of a nonlinear passive state-feedback (SF) controller with a dissipative observer yields the dissipative output-feedback (OF) controller closed-loop stability conditions with: (i) the identification of the underlying gain-behavior interplay, and (ii) simple tuning guidelines. The approach is tested through numerical simulations, with a representative worst-case example: an exothermic reactor with Langmuir-Hinshelwood nonmonotonic kinetics, which must be regulated about an open-loop unstable steady-state which is not observable.

Keywords: Chemical Reactor Models, Output-Feedback Control, Dissipativity, Observability.

1. INTRODUCTION

Continuous exothermic chemical reactors are complex nonlinear dynamical systems with nonlinear behavior, asymmetric MIMO coupling, parametric sensitivity, multiplicity, hysteresis, bifurcation, and limit cycling. Most of the industrial reactors are controlled by combining conventional (ratio, and cascade) feedforward and (P, PI and PID) feedback linear control component with supervisory or advisory material-energy balance and optimizing controllers (Shinskey [1988], Gonzalez and Alvarez [2005]). The process design or redesign to meet tighter safety, productivity, quality and environmental requirements motivates the development of more capable and systematic reactor control designs. Advanced nonlinear control studies have been performed in the chemical process systems engineering field, the related state of the art can be seen elsewhere, and here it suffices to mention that: (i) with a few exceptions (Alvarez et al. [1991], Viel and Jadot [1997], Antonelli and Astolfi [2003]) most of the studies lack rigorous stability and performance assessments, and (ii) only the optimality-based MPC (which stems from industrial control developments) has reached the stage of acceptance for plant scale testing or implementation (Eaton and Rawlings [1990]). Recently, in the context of polymer reactor (Gonzalez and Alvarez [2005], Diaz-Salgado et al. [2007]) and distillation column output-feedback control studies (Castellanos-Sahagun and Alvarez [2006]) with constructive nonlinear control, connections between PI, inventory and MP control designs have been identified, and the closed-loop stability assessment and tuning aspects have been handled either with conceptual arguments or with the small gain theorem. The dissipativity notion offers a unifying framework to handle design-oriented tools in constructive control (Sepulchre et al. [1997]) according to fundamental connections between optimality, passivity, robustness and dissipativity, with emphasis on interlaced observer-control designs and rigorous stability assessments. The dissipativity ideas (i) were originally developed in the context of state-feedback (SF) control problems (Willems [1972]), (ii) have been extended to design of nonlinear observers (Moreno [2005]), and observer-control separation (Moreno [2006]), and (iii) enable the tackling of the difficult problem of estimating and controlling reactors with non-monotonic kinetics, and lack of observability at maximum reaction rate (Schaum et al. [2008]).

The preceding considerations motivate the present reactor output-feedback (OF) control study, where the problem of controlling a continuous exothermic (possibly open-loop unstable) reactor with either monotonic or non-monotonic kinetics, temperature measurements, and manipulation of reactant and heat exchange rates is addresed within a combined passivity-dissipativity approach, including (i) the derivation of rigorous closed-loop stability conditions coupled with easy-to-apply tuning guidelines, and (ii) the identification of the underlying interplay between regulation speed, robustness, and observer-control gains. The proposed approach is tested, through numerical simulations, with an exothermic reactor with nonmonotonic kinetics, open-loop instability, and lack of observability.

In our previous study (Schaum et al. [2008]) the reactor problem was adressed by *ad hoc* combining a passive controller with a dissipative observer, and drawing closed-loop stability conditions with the small gain theorem. However, the passivity (controller) and dissipativity (observer) approaches were methodologically disconnected, and the

stability characterization was not reflected in a practical tuning. In the present work: (i) the controller-observer design and the closed-loop stability assessment are performed with a united framework, and (ii) a simple tuning scheme that is clearly related with closed-loop functioning features is obtained .

2. CONTROL PROBLEM

Consider a continuous chemical reactor where a reactant is converted into product via an exothermic reaction, heat being removed through a diathermal wall with a cooling jacket. Assuming the volume (V) and the jacket temperature (T_c) are controlled with fast (conventional, linear decentralized) feedback loops which manipulate the exit and coolant flowrates (Shinskey [1988]) the *reactor model* is given by the dynamic mass-energy balance:

$$\begin{array}{l} \dot{c} = -r(c,T,\pi) + \theta(c_e - c), \qquad c(0) = c_0 \\ \dot{T} = \beta r(c,T,\pi) + \theta(T_e - T) - \eta(T - T_c), T(0) = T_0 \end{array} \quad (1)$$

where ($\bar{(\cdot)}$ is the steady-state (SS) value of $(\cdot)$)

$$\begin{array}{l} c = C/C^0, \quad \theta = q/V, \quad \beta = (-\Delta H)C^0/(V\rho_m c_p) \\ \eta = (UA_U)/(V\rho c_p), \quad p = (p_a', \pi')', \quad p_a = (c_e, \beta, \eta)' \\ r(\bar{c},\bar{T}) + \overline{\theta}(\bar{c}_e - \bar{c}) = 0, \\ \beta r(\bar{c},\bar{T},\pi) + \overline{\theta}(\bar{T}_e - \bar{T}) - \eta(\bar{T} - \bar{T}_c) = 0 \end{array}$$

The reactant dimensionless concentration c, and the reactor temperature T are the states, the dilution rate q and the jacket temperature T_c are *control inputs*, r is the nonlinear reaction rate function, π is its parameter vector, θ is the inverse residence time, η is the heat transfer coefficient-to-capacity quotient, β is the adiabatic temperature rise, the feed concentration c_e and temperature T_e are the *exogenous inputs*, C (or C^0) is the reactant (or pure reactant) concentration, q is the feed flowrate, $-\Delta H$ is the heat of reaction, ρ_m (or c_p) is the reacting mixture density (or specific heat capacity), U (or A_U) is the heat transfer coefficient (or area), and p is the model parameter. The temperatures (T and T_c) are measured, and the concentrations (c_e and c) are not. In compact vector notation the *reactor model* (1) is given by

$$\begin{array}{l} \dot{x} = f\left[x, d(t), u, p\right], \quad x(0) = x_0, \quad y = Cx, \quad z = x \\ x = [c,T]' \in X = [0,1] \times (T^-, T^+) \subset \mathbb{R}^2, \quad p = (p_a', \pi')' \\ f[\bar{x}, \bar{d}, \bar{u}, p] = 0, \quad d = [c_e, T_e]', \quad T_e = y_e - \tilde{y}_e, \ C = [0,1], \\ u = (\theta, T_c), \ T^- = \min(T_e, T_c), \ T^+ = \max(T_e, T_c) + \beta \end{array} \quad (2)$$

x is the state, u (or d) is the control (or exogenous, possibly time-varying) input, and y (or z) is the measured (or regulated) output. X is an invariant set, meaning that all state motions born in X stay in X (Alvarez et al. [1991]). Since the reactor model (1) contains constant ($\tilde{p}$) and time-varying reactor (or feed) temperature measurement $\tilde{y}$ (or $\tilde{y}_e$), and dilution rate ($\tilde{\theta}$ (or coolant temperature ($\tilde{T}_c$)) actuator bounded errors, the *actual reactor system dynamics* are given by

$$\begin{array}{l} \dot{x} = f[x, d + \tilde{d}(t), u + \tilde{u}(t), p + \tilde{p}], \\ x(0) = x_0, \quad y = Cx + \tilde{y}(t), \quad z = x \\ \tilde{p} = (\tilde{p}_a, \tilde{\pi}), \ \tilde{d}(t) = [\tilde{c}_e(t), \tilde{y}_e(t)]', \ \tilde{u}(t) = [\tilde{\theta}(t), \tilde{T}_c(t)]', \\ \tilde{y}(t) = y - T|\tilde{p}| \le \delta_p, \quad ||\tilde{d}(t)|| \le \delta_d, \quad ||\tilde{u}(t)|| \le \delta_u, \\ ||\tilde{y}(t)|| \le \delta_y, \quad ||(.)(t)|| = \sup_{t \in [0,\infty)} |(.)(t)|, \end{array} \quad (3)$$

where δ_p, δ_d, δ_u and δ_y are the error sizes, and $|(\cdot)|$ is the Euclidian norm of the vector $(\cdot)$. Our *control problem* consists in designing, on the basis of the reactor model (1) (with parameter approximation p) and flow and temperature measurement, an observer-based dynamical *output-feedback (OF) controller* to regulate the concentration-temperature pair z, about a (possibly opn-loop unstable and unobservable) SS by manipulating the dilution rate-cooling temperature pair u.

3. OUTPUT-FEEDBACK (OF) CONTROLLER

The reactor dynamics represent mass and energy accumulation due to advective, reaction and heat exchange input/output mechanisms. From the abstract energy perspective associated with the dissipativity control (Willems [1972], Sepulchre et al. [1997]) and estimation (Moreno [2005]) framework, our OF control problem amounts to designing the observer-control pair in such a way that the dissipation rate is negative, and robust, and implies nonwasteful control action.

In deviation form referred to the SS regime, the reactor system (1) is written as follows

$$\begin{array}{l} \dot{e} = f_e[e, \tilde{u}_e(t)], \ e(0) = e_0, \ e = x - \bar{x}, \\ \tilde{u}_e = (\tilde{p}', \tilde{d}', \tilde{u}')', \ f_e(0,0) = 0. \end{array} \quad (4)$$

According to the definition of *nonlocal input-to-state stability (ISS)* (Freeman and Kokotovic [1996]), the SS $e = 0$ is said to be *practically uniformly* (P) *stable* if an admissible disturbance size (δ_u) produces an admissible state deviation size (ε_x): given $(\delta_u, \varepsilon_x)$ there is a KL-class (increasing-decreasing) function β and a K-class (increasing) γ so that the state responses of system (4) are bounded as follows

$$\begin{array}{l} |e_0| \le \delta_0, |\tilde{u}_e(t)| \le \delta_u, \\ \Rightarrow |e(t)| \le \tau(|e_0|, t) + \alpha(||\tilde{u}_e(t)||) \le \tau(\delta_0, 0) + \alpha(\delta_u) = \varepsilon_x \end{array} \quad (5)$$

where τ (or α) bounds the transient (or asymptotic) response. The (necessary and sufficient) Lyapunov characterization of the ISS property is given by

$$\alpha_1(|e|) \le V(e) \le \alpha_2(|e|), \quad \dot{V} = -\alpha_3(|e|) + \alpha_4(|\tilde{u}_e|) \quad (6)$$

where V is a positive definite radially unbounded function and $\alpha_1, \cdots, \alpha_4$ are K-class functions.

3.1 Passive state-feedback (SF) controller

The notion of *passivity* plays a key role in the design of robust nonlinear SF controllers (Khalil [2002]), with: (i) fundamental connections between optimality, robustness and passivity, and (ii) means to analytically construct optimal controllers via inverse optimality. An optimal SF controller is passive and underlien by a minimum phase (MP) system (with relative degree less or equal than one). A nonlinear system is passive if it is *dissipative* (Willems [1972]) with storage function-supply rate pair and MP.

The reactor (1) is feedback-passive (after input coordinate change) with respect to the input-output pair (u, z) and the storage function $V = e^T e$ if and only if the reactor relative degree equal to one condition is met, i.e. (Schaum et al. [2008]):

$$\mathrm{rd}(u,z) = (1,1), \ z = x \ \Leftrightarrow c_e \neq c \ , \ \eta \neq 0 \quad (7)$$

Thus, the state-input coordinate change $e = x - \bar{x}, v = f(x,d,u)$ takes the reactor into the passive normal form (8)

$$\dot{e} = v, e(0) = e_0, \psi = e; V = e^T e, \dot{V} = 2\psi^T v, \quad (8)$$

with storage function V and input-output pair (v, ψ). The SF controller (9) yields the closed-loop (decoupled, stable, and dissipative) dynamics (10),

$$v{=}f(e,d,u){=}{-}Ke, K{=}\text{diag}(k_c,k_T) \Rightarrow u{=}\mu(x,d,u) \quad (9)$$

$$\dot{e}{=}{-}Ke, e(0){=}e_0, \psi{=}e; V{=}\ e^T e, \dot{V}{=}{-}2eTKe < 0. \quad (10)$$

In original coordinates, *the nonlinear passive SF controller* (9) is given by:

$$\begin{array}{l} \theta = [r(c,T) - k_c(c-\bar{c})]/(c_e - c), \\ T_c = T - [\beta r(c,T) + \theta(T_e - T) + k_T(T - \bar{T})]/\eta \end{array} \quad (11)$$

This controller with state, parameter, and measurement-actuator errors $(\epsilon, \tilde{d}, \tilde{p})$, yields the closed-loop dynamics (13) with dissipation (14)

$$u = \mu(x+\epsilon, d+\tilde{d}, p+\tilde{p}) := [\mu_\theta, \mu_{T_c}]^T \quad (12)$$

$$\dot{e} = -Ke + \tilde{f}\left[e; \epsilon, \tilde{d}(t), \tilde{p}\right], e(0){=}e_0, K{=}\ \text{diag}(k_c, k_T) \quad (13)$$

$$\dot{V} \leq -2\min\{k_c, k_T\}V + e'\tilde{f}\left[e; \epsilon, \tilde{d}(t), \tilde{p}\right] \quad (14)$$

$$\tilde{f}\left(e; \epsilon, \tilde{d}, \tilde{p}\right) = f\left[\bar{x}+e, \bar{d}+\tilde{d}, \mu\left(x+\epsilon, \bar{d}+\tilde{d}, p+\tilde{p}\right), p\right]$$

$$e = (e_c, e_T)' = x - \bar{x}, \quad \tilde{f}\,(e; 0, 0, 0) = 0.$$

Since the reactor has trivially stable nominal zero-dynamics $e = 0$, the errorless closed-loop is asymptotically stable. From the Lipschitz continuity of (f, μ) the system P-stability follows (Khalil [2002]), with a suitable tradeoff between the initial state (δ_0), parameter (δ_p), input (δ_d and $\delta_{\tilde{x}}$) and state excursion (ε_x) sizes, depending on the choice of the control gain pair (k_c, k_T). The P-stable closed-loop reactor dynamics (13) represents: (i) the behavior attainable with any robust controller, and (ii) the recovery target for the OF control design. The related solvability conditions (7) are generically met by the reactor class (1) because: (i) $c < c_e$, and (ii) $\eta > 0$.

3.2 Dissipative observer

The nonlinear global detectability property of any reactor motion (Schaum et al. [2008]) suggests the consideration of a dissipative observer, because (i) its functioning does not require complete observability (Moreno [2005]), and (ii) the structure-oriented approach offers a means to perform the control-estimator design (Section 4). The reactor dissipative observer is given by (Schaum et al. [2008])

$$\begin{array}{l} \dot{\hat{c}} = -r[\hat{c} - \kappa_r(\hat{T} - y), y, \pi] + \theta(c_e - \hat{c}) - \kappa_c(\hat{T} - y), \\ \dot{\hat{T}} = \beta r[\hat{c} - \kappa_r(\hat{T} - y), y, \pi] + \theta(T_e - \hat{T}) - \\ \qquad\qquad -\eta(\hat{T} - T_c) - \kappa_T(\hat{T} - y), \\ \hat{c}(0) = \hat{c}_0, \quad \hat{T}(0) = \hat{T}_0, \end{array} \quad (15)$$

where κ_c (or κ_T) is the usual concentration (or temperature) gain, and κ_r is the gain of an injection in the concentration argument of the reaction rate. The estimation error dynamics are given by the two-subsystem interconnection in Lur'e- Popov form (Khalil [2002], Willems [1972], Schaum et al. [2008])

$$\begin{bmatrix} \dot{\varepsilon}_c(T) \\ \dot{\varepsilon}_T(t) \end{bmatrix} = \begin{bmatrix} -\theta(t) & -\kappa_c \\ 0 & -\lambda_T \end{bmatrix} \begin{bmatrix} \varepsilon_c(T) \\ \varepsilon_T(t) \end{bmatrix} + \begin{bmatrix} 1 \\ -\beta \end{bmatrix} \nu \quad (16)$$

$$\psi = \zeta \triangleq \epsilon_c - \kappa_r \epsilon_T, , \quad \lambda_T \triangleq \theta(t) + \eta + \kappa_T$$

$$\nu = -\rho\,(c, y; \zeta)\,, \quad (17)$$

with (i) a linear-dynamic advective subsystem (16) with input ν and output ζ, and (ii) a nonlinear-static kinetic subsystem (17) with the reaction rate error. Since the rate r is continuously differentiable, there is a continuous secant function φ so that the estimated minus the actual rate is conically bounded (18) with the nonlinearity ρ is encompassed in the conic sector (19)

$$\rho(c, T; \zeta) \triangleq r\,(c + \zeta, T) - r(c, T) = \varphi\,(c, T; \zeta)\,\zeta, \quad (18)$$

$$\zeta \triangleq \epsilon_c - \kappa_r \epsilon_T, \quad -k_1\,(T) \leq \varphi\,(c, T; \zeta) \leq k_2\,(T)$$

$$-k_1\,(T) = \min_{0 \leq c \leq 1} r_c(c, T, \pi)\ , k_2\,(T) = \max_{0 \leq c \leq 1} r_c(c, T, \pi),$$

$$(k_2\zeta - \rho(c, T; \zeta))\,(\rho(c, T; \zeta) + k_1\zeta) \geq 0 \quad (19)$$

(Khalil [2002]). Consequently, the static system (17) is $[-1, 1/2(k_2 - k_1), -k_1 k_2]$-dissipative Moreno [2005], and its dissipation is characterized by the reaction rate slopes: the slope k_1 is positive (or negative) if the reaction rate is monotonic (or non-monotonic). The observer is designed in such a way that: (i) the open-loop estimation error dynamics consist of the feedback interconnection of two adequatly dissipative (passive) subsystems, and (ii) the estimator and control dissipativity properties are structurally compatible. The observer gains $\kappa_c, \kappa_T, \kappa_r$ are chosen so that the system interconnection (16) - (17) is dissipative with respect to the estimation storage function

$$\hat{V} = \frac{1}{2}\epsilon^T \epsilon. \quad (20)$$

The gain pair (κ_c, κ_T) shapes the dissipation of the linear dynamical subsystem, and the gain κ_r determines the interconnection form by setting the output of the linear system. Convergence conditions for the dissipative open-loop observer (15) are given in (Schaum et al. [2008]).

3.3 OF controller

The combination of the SF (9) passive nonlinear controller with the dissipative observer (15) yields the *dynamic OF controller*

$$\begin{array}{l} \dot{\hat{c}} = -r[\hat{c} - \kappa_r(\hat{T} - y), y, \pi_r] + \theta(c_e - \hat{c}) - \kappa_c(\hat{T} - y)\ , \\ \dot{\hat{T}} = \beta r[\hat{c} - \kappa_r(\hat{T} - y), y, \pi_r] + \theta(T_e - \hat{T}) - \\ \qquad\qquad -\eta(\hat{T} - T_c) - \kappa_T(\hat{T} - y), \\ \theta = [r(\hat{c}, T) - k_c(\hat{c} - \bar{c})]/(c_e - \hat{c}), \\ T_c = \hat{T} - [\beta r(\hat{c}, T) + \theta(T_e + \hat{T}) + k_T(\hat{T} - \bar{T})]/\eta \end{array} \quad (21)$$

with five adjustable gains: k_c and k_T for the passive-dissipative controller, and κ_c, κ_T and κ_r for the observer.

4. CLOSED-LOOP STABILITY AND TUNING

In this section, the closed-loop dynamics are characterized, yielding: (i) stability conditions, (ii) tuning guidelines, and (iii) a functioning assessment. The main difficulty resides in an inherent limitation: the unmeasured output concentration (c) must be regulated about a steady-state wich is open-loop unstable and not locally observable.

The application of the OF controller (21) to the actual reactor (3) yields the *closed loop dynamics*

$$\begin{array}{l} \dot{e} = -Ke + \psi(e)\epsilon + \phi\left(e, \epsilon; \tilde{d}(t), \tilde{p}\right), \\ \dot{\epsilon} = M(t)\epsilon + \tilde{g}\left(e, \epsilon; \tilde{d}(t), \tilde{p}\right), \end{array} \quad (22)$$

$$M(t) = \begin{bmatrix} -\theta(t) - \varphi(t) & -\kappa_c \\ \beta\varphi(t) & -\lambda_T \end{bmatrix}, \phi\,(e; 0) = 0$$

where e (or ϵ) is the regulation (or estimation) error, ϕ results from the replacement of ϵ_T by $\tilde{y}$ in the reaction rate term of the error function $\tilde{f}$ (13) associated with the Lyapunov closed-loop stability characterization with SF control. From the continuity of ϕ, $\tilde{f}$, $\tilde{g}$ and the compactness of their domains their Lipschitz continuity and boundedness follow.

Given that the separation principle holds for linear but not for nonlinear systems, the nominal closed-loop stability (i.e. system (22) with $(\tilde{g}, \phi) = (0, 0)$) can be established as follows: since the regulation error dynamics are individually P-stable and the estimation error dynamics are individually convergent, the reactor(1)-OF controller(21) interconnection is uniformly asymptotically stable (Angeli et al. [2004], Moreno [2006]). Motivated by the need of a more constructive stability criterion in the sense of practical applicability for gain tuning and behavior assessment purposes, in the next propositionclosed-loop stability conditions are given in terms of the five-gain set $(k_c, k_T, \kappa_c, \kappa_T, \kappa_r)$ of the proposed OF controller (21).

Proposition 4.1. (Sketch of proof in Appendix A)
The closed-loop reactor (1) with the proposed passive-dissipative OF controller (22) is P-stable if the controller five-gain set $(k_c, k_T, \kappa_c, \kappa_T, \kappa_r)$ and the regulation-estimation error set meet the conditions

$$(i)\ \theta = \mu_\theta(k_c) > -k_1, \qquad (ii)\ k_c > \iota_c(k_c)$$
$$(iii)\ k_T > \iota_T(k_c, \kappa_c, \kappa_r), \quad (iv)\ \kappa_T > \iota_\tau(k_c, k_T, \kappa_T, \kappa_c, \kappa_r),$$

with μ_θ givn in (12) and $\iota_c, \iota_T, \iota_\tau$ in Appendix A.

As it can be seen in Appendix A, the combined passivity-dissipativity approach enables the derivation of the above stability conditions in a rather straighforward way, by using the passive control (V) (13) and dissipative observer $(\hat{V})$ (20) storage functions and applying Lyapunovs direct method. In the absence of modeling error the closed-loop stability becomes asymptotic. Condition (i) is a closed-loop detectability requirement, Condition (ii) ensures the stability of the regulation-estimation concentration dynamics and imposes lower and upper limits ($k_c^- \approx 1, k_c^+ \approx 3$) on the composition control gain k_c (Gonzalez and Alvarez [2005]), and Conditions (iii) and (iv) ensure the stability of the regulation-estimation temperature dynamics and of the entire interconnection. Thus, for $\kappa_r \approx 1/\beta, k_c \approx 3\bar{\theta}$, the preceding inequality conditions can be met by choosing: (i) k_T sufficiently large to dominate $\iota_T(k_c, \kappa_c, \kappa_r)$, and (ii) κ_T sufficiently large to dominate $\iota_\tau(k_c, k_T, \kappa_c, \kappa_r)$.

From the preceding P-stability conditions the conventional-like tuning guidelines follow: (i) set the gains conservatively at $(k_c, k_T) \approx (1, 3), \kappa_r \approx 1\beta, \kappa_c \approx k_c, \kappa_T \approx 10\kappa_c$, (ii) increase the T-estimation gain κ_T until oscillatory response is obtained at κ_T^+, back off and set $\kappa_T = \kappa_T^+/2$-to-3, (iii) in the same way set $k_T = k_T^+/2$-to-3, (iv) carefully increase k_c (sufficiently below $k_c^+ \approx 4\bar{\theta}$) until there is no improvement, and adjust κ_r. If necessary, repeat steps (ii) to (iv).

The solvability of the robust OF reactor control problem is a consequence of: (i) the solvabilities of the OF control (7) and dissipative closed-loop observer (condition (i) in Proposition 4.1) problems, and (ii) the adequate choice of gains according to Proposition 4.1.

5. APPLICATION EXAMPLE

To subject the proposed OF controller to a severe test, let us consider an extreme case of an industrial situation: the operation of the continuous reactor (3) with the Langmuir-Hinshelwood (LH) kinetics model

$$r(c, T, \pi) = \frac{cke^{-(\frac{\gamma}{T})}}{(1 + \sigma c)^2}, \quad r_c = (c^*, T, \pi) = 0, c^* = 1/\sigma$$

adapted from a previous (partial open-loop or asymptotic and full measurement injection) estimation study with EKF and experimental data for the catalyzed carbon monoxide oxidation reaction (Baratti et al. [1993]). With the nominal parameters and inputs

$\bar{d}' = (\bar{c}_e, \bar{T}_e) = (1, 1), \bar{u}' = (\bar{\theta}, \bar{T}_c) = (1, 370), p = (p_a', \pi')'$,
$p_a = (\bar{c}_e, \bar{T}_e, \eta)' = (1, 370, 1), \pi' = (k, \gamma, \sigma) = (e^{25}, 10000, 3)$

the reactor has three steady-states (Diaz-Salgado et al. [2007]: two stable (extinction and ignition), and one unstable at maximum concentration rate $r^* = 0.6614$ with $c^* = 1/3$. The application of the tuning guidelines associated with Proposition 4.1 yielded: $\kappa_c = 0.62$, $\kappa_T = 30$, $\kappa_r = \frac{1}{50}$, $k_c = 2$, $k_T = 3.$, and the initial reactor and estimator conditions were set at $x_0 = [430, 0.28]'$, $\hat{x}_0 = [425, 0.35]'$, about the unstable steady-state with maximum rate. The relative degree (7) and global detectability (Schaum et al. [2008]) conditions are well met, because: $c_e - \bar{c} = 2/3 > 0, \eta = 1 > 0, \bar{\theta} = 1, 1/3 = \theta^- \leq \theta \leq \theta^+ = 3/2$. In the spirit of the nonlocal P-stability framework, the closed-loop reactor with nominal SF, nominal and perturbed OF will be subjected to initial state, and persistent parameter and exogenous input disturbances.

5.1 Nominal behavior with SF control

The closed-loop reactor behavior with exact model-based nonlinear passive SF controller (9) is shown in Figure 1. As expected, the concentration (or temperature) response is about one half (or quarter) settling residence time $(4/\theta = 4)$, with smooth-coordinated dilution rate-coolant temperature control action, safely away from saturation. This agrees with the optimality-based non-wasteful feature of passive SF controllers (Sepulchre et al. [1997]).

5.2 Nominal behavior with OF control

Initially, the reactor was in the above stated deviated initial state, and subjected to known constant feed concentration $c_e = 1$ and temperature $T_e = 370K$. The behavior with exact model-based OF control (21) is shown in Figure 2: (i) the state responses are quite similar to the ones of the nonlinear SF controller (Figure 1), in spite of a sluggish concentration estimate response (about 3/4th of the natural settling time), and (ii) as expected from the FF component of the OF controller, the control inputs practically annihilate the effect of the known oscillatory input, and (iii) the control actions are smooth and efficient, reasonably away from saturation. Thus, the nominal OF controller recovers rather well the behavior of its exact model-based nonlinear SF counterpart. This test verifies

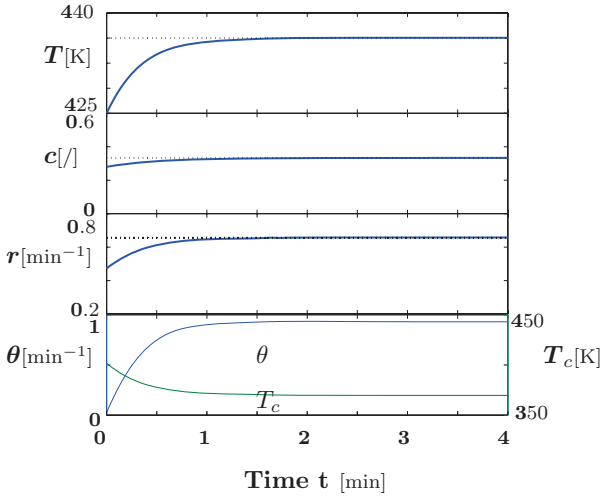


Fig. 1. Closed-loop nominal behavior with nonlinear SF controller: input and response (—) , estimate (− −), and set point (· · ·).

the closed-loop P-stability property with OF dynamic control, with asymptotic convergence to the prescribed SS.

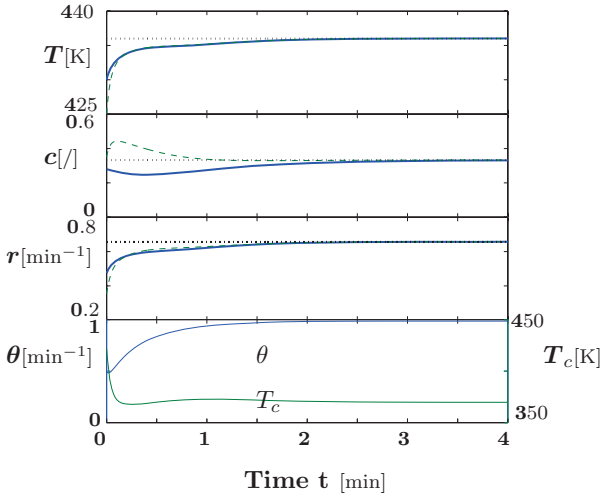


Fig. 2. Closed-loop nominal behavior with nonlinear OF controller: input and response (—) , estimate (− −), and set point (· · ·).

5.3 Robust behavior with OF control

To test the robustness of the OF controller, the reactor and the estimator initial states were deviated from the nominal open-loop unstable and maximum reaction rate steady-state, and subjected to the oscillatory feed concentration and temperature inputs

$$c_e = 0.99 + 0.01\cos(4\pi t), \quad T_e = 370 + 2\sin(4\pi t)$$

The constant errors in the estimation model correspond to: (i) $\hat{c}_e(t) = 0.991$, (ii) measured feed and reactor temperatures with considerable periodic error $\hat{T}_e(t) - T_e(t) = y(t) - T(t) = 2\cos(40\pi t)$ (four degrees amplitude band and frequency close to natural resonance mechanism), and (iii) -1.5, -10, and $+3$ % errors in the activation energy (γ), heat transfer coefficient (η), and adiabatic temperature rise (β), respectively. These errors represents a worst-case situation to subject the OF controller to a severe robustness test. The resulting closed-loop behavior is presented in Figure 3: (i) the reactor is adequately P-stable with a transient response trend that basically coincides with the one of the errorless model case (see Figure 2), (ii) as expected from the severe modelling errors, the unmeasured concentration exhibits a significant ($\approx -30\%$) asymptotic offset, some reaction rate offset ($\approx -20\%$) and the temperature estimate generated by the linear-dynamical advective (that is mass-energy balance based) estimation component yields an offset-less trend response, and (iii) given the flatness feature of the reaction kinetics in the isotonic branch of the reaction rate function, in spite of the -30% concentration trend offset, the reaction rate trend is only a -20% of its maximum set point value. Should it be necessary, the optimal rate offset can be reduced by online kinetic parameter model calibration on the basis of the occasional or periodic concentration measurements that are usually taken for quality monitoring purposes.

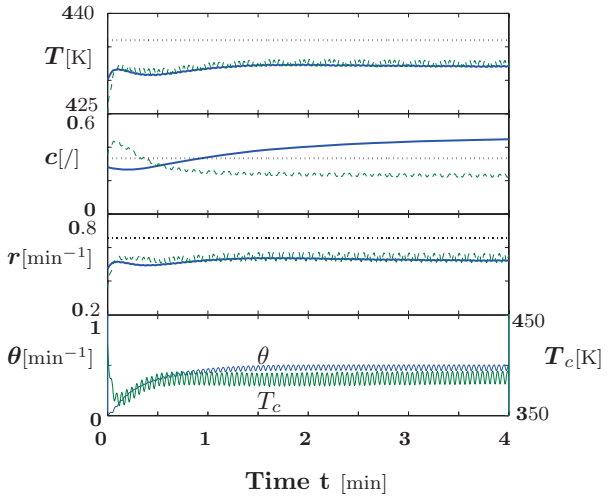


Fig. 3. Closed-loop robust behavior with nonlinear OF controller: input and response (—) , estimate (− −), and set point (· · ·).

5.4 Concluding Remarks

In agreement with the theoretically drawn methodology, the proposed passive-dissipative OF controller: (i) recovers rather well the behavior of its exact model-based nonlinear SF counterpart, with optimality-based robustness and control non-wastefulness, and (ii) exhibits P-(robust and non local) stability with respect to model, and measurement

errors. The closed-loop behavior assessment through simulations made quantitative the P-stability features (like transient response speed, overshoot, high frequencies oscillatory components, and asymptotic response offsets), and verified the effectiveness of the gain tuning scheme obtained from the P-stability characterization.

6. CONCLUSIONS

A robust OF control design methodology for continuous reactors with temperature measurements has been presented. Structural (relative degree and global detectability) solvability conditions were identified and exploited to design a nonlinear dynamic dissipative-passive OF controller. The interlaced estimator-control design led to a robust OF control scheme with: (i) a systematic construction procedure, and (ii) a rigorous closed-loop (nonlinear-nonlocal) P-stability criterion, (iii) simple tuning guidelines, and (iv) behavior recovery, up to estimator convergence, of the exact model-based FF-SF nonlinear control. A Langmuir Hinshelwood kinetics (carbon monoxide oxidation) in an open-loop unstable reactor at maximum reaction rate was considered as a representative case example with numerical simulations.

ACKNOWLEDGEMENTS

This work has been supported by DGAPA-UNAM, Project IN112207-3, and CONACyT research grant (Design and Control of Batch Processes) Project 51244. A. Schaum and J. Díaz-Salgado thank CEP-UNAM and CONACyT for financial support.

REFERENCES

Alvarez, J., Alvarez, J., and Suarez, R. (1991). Nonlinear bounded control for a class of continuous agitated tank reactors. *Chem. Eng. Sci.*, 46(12), 3235–3249.

Angeli, D., Ingalls, B., Sontag, E., and Wang, Y. (2004). An output feedback separation principle for input-output and integral-input-to-state stability. *SIAM J. Contr. Optim.*, 43 (1), 256–276.

Antonelli, R. and Astolfi, A. (2003). Continuous stirred tank reactors: easy to stabilise? *Automatica*, 39(10), 1817–1827.

Baratti, R., Alvarez, J., and Morbidelli, M. (1993). Design and experimental identification of nonlinear chemical reactor estimator. *Chem. Eng. Sci.*, 48 (14), 2573–2585.

Castellanos-Sahagun, E. and Alvarez, J. (2006). Synthesis of two-point linear controllers for binary distillation columns. *Chem. Eng. Comm.*, 193, 206–232.

Diaz-Salgado, J., Schaum, A., Moreno, J., and Alvarez, J. (2007). Interlaced estimator-control design for continuous exothermic reactors with non-monotonic kinetics. In *Eighth International IFAC DYCOPS, Cancun, Mexico.*

Eaton, J. and Rawlings, J. (1990). Feedback control of chemical processes using on-line optimization techniques. *Comp. & chemical eng.*, 14(4-5), 469–479.

Freeman, R. and Kokotovic, P. (1996). *Robust Nonlinear Control Design: State-Space and Lyapunov Techniques.* Birkhäuser, Boston.

Gonzalez, P. and Alvarez, J. (2005). Combined pi-inventory control of solution homopolymerization reactors. *Ind. Eng. Chem. Res. J.*, 7147–7163.

Khalil, H. (2002). *Nonlinear Systems (Third Edition).* Prentice Hall, Upper Saddle River, New Jersey.

Moreno, J. (2005). Approximate observer error linearization by dissipativity methods. *Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems, Springer LNCIS*, 35–51.

Moreno, J. (2006). A separation property of dissipative observers for nonlinear systems. In *45th IEEE Conference on Decision and Control, San Diego, CA, USA*, 1647 – 1652.

Schaum, A., Moreno, J., Alvarez, J., and Diaz-Salgado, J. (2008). Dissipativity-based observer and feedback control design for a class of chemical reactors. *Journal of Process Control*, 18(9), 896 905.

Sepulchre, R., Jankovic, M., and Kokotovic, P. (1997). *Constructive Nonlinear Control.* Springer-Verlag, London.

Shinskey, F. (1988). *Process Control Systems 3rd. Edt.* McGraw-Hill, New York.

Viel, F. and Jadot, G. (1997). Robust feedback stabilization of chemical reactors. *IEEE Trans. Autom. Cont.*, 42(4), 473 – 481.

Willems, J. (1972). Dissipative dynamical systems: Part i - general theory, part ii - linear systems with quadratic supply rates. *Archive for Rational Mechanics and Analysis*, 45, 321–393.

Appendix A. PROOF OF PROPOSITION 4.1

Recall the control (V) (13) and observer ($\hat{V}$) (20) storages, set the composed storage W, and write the corresponding dissipation ($\dot{W}$) along the closed-loop reactor motion:

$$W = V + \hat{V}, \quad \dot{W} = -z^T Q z, z = [e_c, \epsilon_c, e_T, \epsilon_T],$$

$$Q = \begin{bmatrix} k_c & -\iota_c(\theta_r+\varphi) & 0 & \kappa_r\varphi(c_e-c)/[2(c_e-\hat{c})] \\ \star & \theta_r+\varphi & (\beta\varphi)/2 & [\kappa_c-(\kappa_r+\beta)\varphi]/2 \\ \star & \star & k_T & 2\iota_1+\iota_r \\ \star & \star & \star & \kappa_T+\iota_r \end{bmatrix}$$

$$\iota_c(k_c) = \frac{(\theta^* - [k_c-\varphi](c_e-c))^2}{4(c_e-\hat{c})(\theta_r+\varphi)}, \iota_r(\kappa_r) = \mu+\eta+\kappa_r\beta\varphi$$

$$\iota_T(k_c) = \frac{k_c\beta^2\varphi^2}{4(\theta_r+\varphi)(k_c-\iota_c(k_c))}, \iota_1 = \frac{(k_T-\kappa_r\beta\varphi)^2}{4k_T} - \iota_r$$

$$\iota_2 = \frac{\varpi(k_c,\kappa_c,\kappa_r}{(\theta_r+\varphi)} - \iota_r, \quad \iota_\tau = \max\{\iota_1,\iota_2\},$$

and ϖ is a class-$\mathcal{K}$ function of its arguments. The enforcement of the positive defintness property in each of the four leading principal minors $(M_1,\ldots,M_4)$, yields the conditions stated in Proposition 4.1, or equivalently the positive definitness of Q implying the closed-loop P-stability property. QED

Modeling and Simulation

Oral Session

Non-linear model order reduction using input to state Hammerstein structures

O. Naeem, * **A.E.M. Huesman,** ** **O.H. Bosgra.** ***

* *Delft Center for Systems & Control (DCSC), Technische Universiteit Delft, 2628 CD, Delft, the Netherlands (e-mail: O.Naeem@tudelft.nl)*
** *(e-mail: a.e.m.huesman@tudelft.nl)*
*** *(e-mail: o.h.bosgra@tudelft.nl)*

Abstract: In this paper, the focus will be on approximating original model of process systems using block-structured models. The context of model reduction is to improve the computational efficiency (simulation time). The reduced order models are important for online applications. Hammerstein structures have been used to approximate a mathematical non-linear model of a process. Input-Output Hammerstein structure can be defined as classical Hammerstein model but the technique is extended here to Input-State Hammerstein structure. It is shown that Input-State Hammerstein structure can be derived from Taylor series. Approximation accuracy has been improved by approximation for second term. The approximated Input-state Hammerstein block structure model gives good approximation of the original non-linear system. Over an operating domain of a process, the Input-State Hammerstein structure provides opportunities for reducing the computational load by order reduction of states and Jacobians. The methodology has been applied to a high purity distillation benchmark and satisfactory results are obtained as far as approximation is concerned. Reduction in states and Jacobian size by 70% is attained.

Keywords: Nonlinear model order reduction, Hammerstein, Taylor series, high purity distillation column

1. INTRODUCTION

First order principle models (rigorous models) are stiff and large, thus are computationally inefficient. Since the (rigorous) NL models are not always exact match of real processes and there is mismatch at some point between the two, reduced models can be very useful if they match the rigorous NL model over a certain operation window. Advantage of reduced mathematical models for NL processes include low computational effort, better approximation of process within the operating window and beneficial for the real-time applications (e.g; control and optimization purposes).

The rigorous models available for large industrial processes can be characterized as a set of differential and algebraic equations (DAE). DAE class of models is capable to express the majority of processes. Thus the methodology to achieve a reduced model should be capable of handling DAE models. The transformation from DAE to ordinary differential equation (ODE) format is regarded as a major model reduction step; but it is not possible generally. A methodology which involves this step is advantageous for the process models of DAE class.

There is not much literature available on model reductions when it comes to model reduction in context of computational load. Balasubramhanya and Doyle (2000) developed a reduced order model of batch distillation column using travelling waves. The closed loop simulation of this reduced model was six times faster than original model in closed loop. Aling et al. (1997) used POD to get reduced model for rapid thermal processing system. Reduction of computational load by a factor ten was reported. Hahn and Edgar (2000) elaborated on model reduction by balancing empirical Gramians and showed model order reduction but reduction in computational effort and time was limited. Perregaard (1993) simplified and reduced chemical processes models for simulation and optimization purposes. He achieved the reduction by simplifying the calculation of algebraic equations, which resulted in computational effort reduction. Gani et al. (1990) replaced the true (symbolic) Jacobian by approximated Jacobian (from local models). They reported the reduction of computational times by factor of $20 \sim 60$. Empirical modelling has been one of the major approaches for achieving low computational complexity (which allows fast simulations). Ling and Rivera (1998) used a Hammerstein structure for model reduction, but did not report reduction in computational load (on polymerization benchmark). Berg (2005) reported that if computational load has to be reduced, not only model order reduction is to be targeted but the complexity (and stiffness) of reduced model has to be lower; as Gani (Gani et al. (1990)) achieved the computational load reduction by reducing the complexity (discussed above). Block structure models have been used for the identification purposes (see Eskinat et al. (1991); Billings and Fakhouri (1977), Norquay et al. (1999), Harnischmacher and Marquardt (2007) etc.). Though block structure models have been used for the identification purposes, but the block structure models have not been used for the model reduction purposes.

As the literature review shows, there is no reduction technique available directly related to reduction of computational load. Each model reduction technique has its specific purposes which is completely understandable. Not every model reduction methodology works for every process, but it is desired to have a model reduction methodology which is generic and applicable to wide class of processes (represented by DAE class of models). Moreover the literature review indicates that there is not much research material available on model reduction subject; whatever is available, mostly addresses the model order reduction and it does not focuses the reduction in complexity of reduced model (which is major component for computational load). Not many model reduction methodologies have addressed the problem of simplification of reduced model. The field is open for research to achieve reduced models, which are simple and reduced order to achieve computational load reduction.

Block structure models have an advantage over other model approximation methodologies; the structure of approximation model gives insight to the complexity of the process and breaks down the complexity of the NL process. This give handles to feel for the complexity and reduce it. Use of block structure enhances the chances to get a reduced model, which is uncomplicated and is computationally efficient.

In this paper, a block structure (Hammerstein) has been used to achieve reduced model for nonlinear chemical process. In the subsequent section, Hammerstein structure is discussed. In section 3, reduction methodology is discussed. In section 4, implementation on high purity distillation column and its results are considered. The last section 5 concludes the paper with key points and future work.

2. HAMMERSTEIN STRUCTURE

There are different block structures which are known for model reduction (and empirical modeling); Wiener, Hammerstein etc. Chen (1995) has introduced and discussed a wide variety of such block structures. Wiener and Hammerstein block structure models are most widely used structures in literature for the representation of nonlinear physical processes and will be shortly discussed here. Wiener models have limitations (specifically for chemical processes) which give edge to Hammerstein structure for identification purposes (Harnischmacher and Marquardt (2007)). Wiener models not only limit the nonlinearity measure to be approximated, but they also increase the complexity involved in identifying or approximating the process. Harnischmacher (2007) investigated that Wiener models restrict the dynamic NL behavior that can be approximated and identified in comparison to Hammerstein structure.

Hammerstein structure is used for the approximation of NL processes in this study. The methodology is extended further to I/S Hammerstein structure to improve the approximation.

2.1 Classical (Input-Output I/O) Hammerstein structure

Classical Hammerstein model can be seen as nonlinear static gain, followed by linear dynamics.

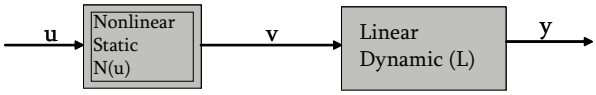


Fig. 1. Classical Hammerstein structure (input-output)

The classical I/O Hammerstein structure shown in figure 1 represents the continuous system/process. A procedure to get Hammerstein structure approximation for a process is to represent the nonlinear static block by interpolation table (lookup table), neural network or spline scheduling (the steady states) and represent linear dynamic block by linear time invariant (LTI) model. The intermediate variable v is a low dimensional vector. Mathematically, classical Hammerstein structure is given as;

$$\begin{aligned} \dot{x} &= A\ x + g(\mathbf{u}) \\ y &= C\ x \end{aligned} \tag{1}$$

where, A is the state matrix, C is the output matrix which can be identified from process data or can be obtained by linearizing the nonlinear system at 'nominal operating point'. Nominal operating point is an operating point within the operating domain, chosen by the input design (discussed in later section 3.1.

The input-output (I/O) Hammerstein model shown can be modified to Input-state (I/S) Hammerstein model under few assumptions (Naeem et al. (2008)). I/S Hammerstein model can be derived from expansion of Taylor series (shown in section 2.3)

2.2 Taylor Series

The Taylor expansion of a function $f(x)$ that is differentiable in the neighborhood of real or complex number 'a' is mathematically given as:

$$f(x) = f(a) + \frac{1}{1!}\left.\frac{\partial f}{\partial x}\right|_a (x-a) + \frac{1}{2!}\left.\frac{\partial f^2}{\partial^2 x}\right|_a (x-a)^2 + \dots \tag{2}$$

Typically process models are of DAE format and transformation from DAE to ODE is a model reduction step (for a large scale process). The ODE can be approximated by I/S Hammerstein structure.

Given an ODE $\dot{x} = f(x, y)$, which is modeled in an environment (gPROMS, MATLAB, SIMULINK), the first order Taylor expansion around point (x^*, u^*) is given mathematically as:

$$\dot{x} = f(x,u) = f(x^*,u^*) + J_x|_{x^*,u^*}\,(x-x^*) + J_u|_{x^*,u^*}\,(u-u^*) + \dots \tag{3}$$

Equation 3 evaluates the function $f(x, u)$ given linearization at $f(x^*, u^*)$. Figure 2 shows the equation 3 in block diagram.

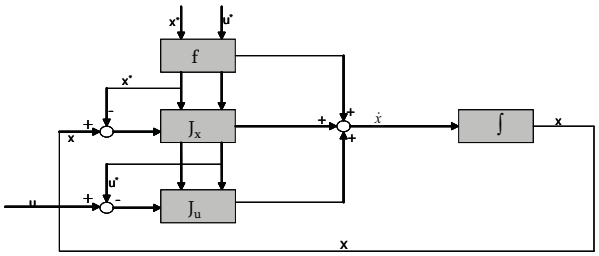


Fig. 2. Block structure for Taylor series of $\dot{x} = f(x, u)$

2.3 Taylor Series expansion

Taylor series illustrated in preceding section can be extended to I/S Hammerstein structure (at steady-state point (x^*, u^*)) under following assumptions:

a. It is assumed, within the operating domain that every u^* leads the system finally to steady-state 'x_{ss}' which means a stable process is considered. Moreover, it is assumed that steady-state is calculated by u. Mathematically, $x_{ss} = g(\mathbf{u})$. Setting $x^* = x_{ss}$ results in output of the (constant) block (f) zero (the system is being evaluated at steady-state x^*).
b. Input u is chosen freely, but is chosen such that it is equal to the input at steady-state, mathematically; $u = u^*$; this implies that gradient input to block ($J(u)$) becomes zero; (since $u - u^* = 0$).

Under above assumptions, adding $g(\mathbf{u})$, removing blocks f and J_u and rearranging the block structure in figure 2, we get the block structure shown in figure 3. Observing this structure it can be considered an *I/S Hammerstein structure*; it has two blocks, a NL steady-state mapping block, followed by linear dynamic block.

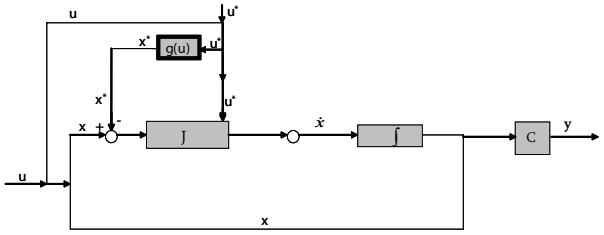


Fig. 3. Taylor series extension to I/S Hammerstein structure

For a linear system, state-space model can be mathematically given as below:

$$\begin{aligned} \dot{x} &= A\ x + B\ u \\ y &= C\ x \end{aligned} \tag{4}$$

For the state-space linear model (in equation 4), I/S Hammerstein structure can be shown as figure 4.

The block structure shown in figure 4 can be extended for NL processes, shown in figure 5. I/S Hammerstein structure shown in figure 5 is similar to the structure derived by Taylor series expansion (shown in figure 3).

The block structure shown in figure 3 and figure 5 is I/S Hammerstein block structure, with separated NL statics(where NL static mapping takes place), followed by linear dynamic block. The dynamic linear block is driven by difference between the steady state 'x_{ss}' and current state 'x'.

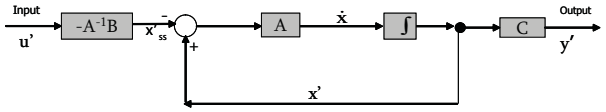


Fig. 4. Input-state Hammerstein structure for linear system

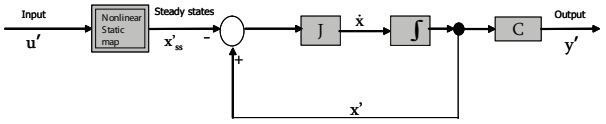


Fig. 5. Input-state Hammerstein structure for non-linear system

Mathematically, for NL case, I/S Hammerstein structure can be given as:

$$\begin{aligned} \dot{x} &= J\ (x - x_{ss}) + g(\mathbf{u}) \\ y &= C\ x \end{aligned} \tag{5}$$

where, J = *Jacobian*; C = *output state matrix*; y = *output*; x_{ss}= $g(\mathbf{u})$ *is steady-state, scheduling (implemented by lookup table).*

The I/S Hammerstein block structure shown in figures 3 and 5 is used for the approximation of NL processes. The accuracy of approximation of I/S Hammerstein structure is improved by estimating Jacobian online. Jacobian is estimated (and updated) based on information of Jacobian basis J_b, reduced state z and input u. Jacobian basis J_b are calculated by SVD analysis of Jacobian data. Jacobian data is collected by taking snapshots of Jacobians over the operating domain ('*input design*') by exciting the NL system with inputs to acquire most information in operating envelope. Similarly the reduced order states z is calculated by transformation matrix (U_1), obtained by SVD analysis of steady-state and dynamic state (snapshot) data, taken over the operating domain.

First order I/S Hammerstein approximation structure (with updated Jacobian), is shown in figure 6. As figure shows,the Jacobian estimation is based on (reduced) current state information (z).

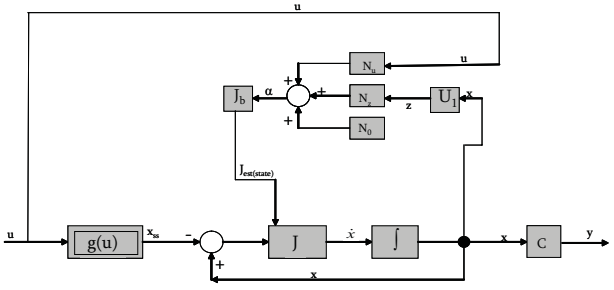


Fig. 6. First order I/S Hammerstein approximation structure

In the figure 6, J = *Jacobian* ; U_1 = *transformation matrix to transform full state x to reduced state z* ; J_b =

Jacobian basis obtained by SVD analysis of snapshots of Jacobians (within operating domain) ; N_0, N_1, N_2 = *the parameters relating Jacobian with reduced-state z, input u and constant.*

It is to be noted, I/S Hammerstein can be derived by extending Taylor expansion as proved above. This is not possible for Wiener structure.

2.4 Accuracy improvement by higher order approximation

First order approximation of NL system by expansion of Taylor series to I/S Hammerstein structure is shown in above section. The approximation accuracy can be improved with higher order terms.

The Taylor series is extended to second order. Taylor series around (steady-state) point (x^*, u^*) is given as:

$$f(x,u) = f(x^*,u^*) + \left.\frac{\partial f}{\partial x}\right|_{x^*,u^*} (x - x^*) + \frac{1}{2!} \left.\frac{\partial f^2}{\partial^2 x}\right|_{x^*,u^*} (x - x^*)^2 \tag{6}$$

Similarly, Taylor series expansion around any point (x, u) is given as below:

$$f(x^*,u^*) = f(x,u) + \left.\frac{\partial f}{\partial x}\right|_{x,u} (x^* - x) + \frac{1}{2!} \left.\frac{\partial f^2}{\partial^2 x}\right|_{x,u} (x^* - x)^2 \tag{7}$$

Equation 6 and equation 7 are the Taylor series expansions at two points (x^* and x), given by Taylor series extension to second order. Rearranging equation 7, we get;

$$f(x,u) = f(x^*,u^*) + \left.\frac{\partial f}{\partial x}\right|_{x,u} (x^* - x) + \frac{1}{2!} \left.\frac{\partial f^2}{\partial^2 x}\right|_{x,u} (x^* - x)^2 \tag{8}$$

Adding equation 6 and equation 8 (while higher order terms are canceled, assuming $\frac{1}{2!} \left.\frac{\partial f^2}{\partial^2 x}\right|_{x,u} = \frac{1}{2!} \left.\frac{\partial f^2}{\partial^2 x}\right|_{x^*,u^*}$), we get;

$$f(x,u) = f(x^*,u^*) + \frac{1}{2} \left[\left.\frac{\partial f}{\partial x}\right|_{x^*,u^*} + \left.\frac{\partial f}{\partial x}\right|_{x,u} \right] (x - x^*) \tag{9}$$

Equation 9 is the approximation of $f(x, u)$ using higher order terms. There are two Jacobian evaluations involved in this approach, an approximation using knowledge at steady-state (x^*) and approximation using current state (x) knowledge. The block structure representation of this approximation is shown in figure 7.

3. REDUCED ORDER HAMMERSTEIN STRUCTURE

Approximation block structure shown in figure 7 is full state model. Since it is a prerequisite for the approximation block structure, to be valid within certain operating domain.

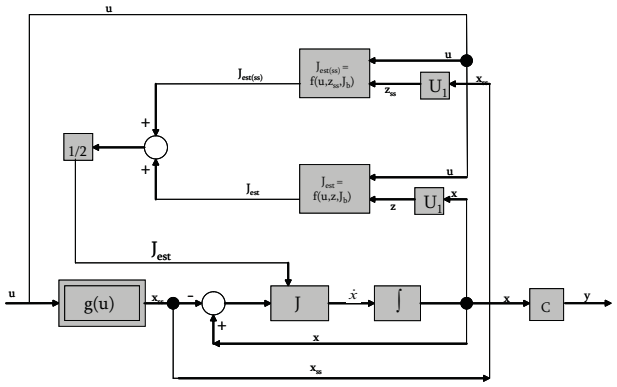


Fig. 7. Approximation model (I/S Hammerstein); Higher order approximation)

3.1 Operating domain/Input Design

Within defined domain, inputs are designed with the purpose to define an input trajectory that travels through complete domain. The process is known as '*input design*'. As the name indicates, it is the process of designing inputs based on constraints on input or output (depending upon a process). In summary, the operating domain is a region, where the approximated model is supposed to be valid, once identified (based on data from the physical process) and input design is the procedure, which defines the boundaries of this operating domain.

The steady-state and dynamic state data is obtained by taking snapshots over the operating domain. Similarly, Jacobian data is collected by taking the snapshots of Jacobians over the operating domain. Jacobian basis J_b are computed by SVD analysis of Jacobian data.

The I/S Hammerstein block structure gives scope to get the reduced order structure by;

i. Reduction in state size.
ii. Reduction in Jacobian size.

i. Reduction in state size.
The singular value analysis on data of states indicates that there is a low dimensional space, such that low order state (z) can represent the whole operating domain. The state reduction is performed by transformation matrix. The transformation matrix U_1 (to obtain reduced state 'z') is obtained by SVD analysis of data over the operating domain. Reduced states are back-transformed to full state by back transformation matrix $\hat{U}_1$. The block structure of the reduced approximation model is shown in figure 8. The Jacobian reduction takes place online (in the block $U_1 * J * U_1^T$).

ii. Reduction in Jacobian size.
The scheme in figure 7 shows that Jacobian estimation takes place using Jacobian basis J_b, state z and input u information. The estimated Jacobian J_{est} is full order Jacobian. There is possibility to obtain reduced size Jacobian, by using reduced order Jacobian basis ($J_{b_{red}}$) and reduced state (z). With reduced basis $J_{b_{red}}$ and reduced states z, estimated Jacobians are also reduced sized and the block structure is a reduced order I/S

Hammersteins approximation model. The block structure is shown in figure 9.

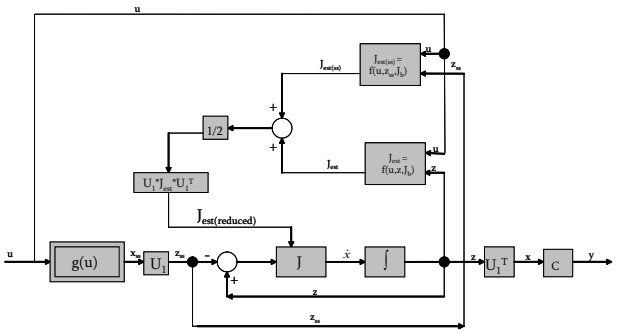


Fig. 8. Reduced order approximation model (I/S Hammerstein); Higher order approximation

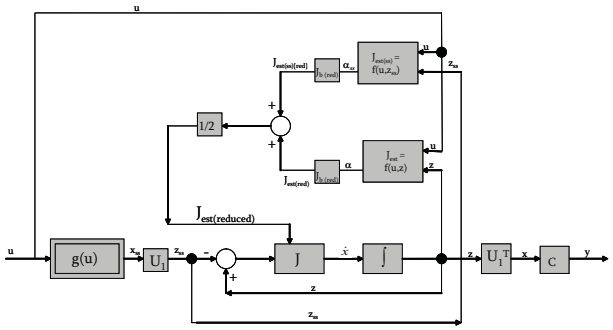


Fig. 9. Approximation block structure with reduced Jacobian and state size

4. APPLICATION TO HIGH PURITY DISTILLATION COLUMN

The approximation block structure and reduced block structure model has been applied to the benchmark. A benchmark is high purity distillation column, and its properties will be discussed in subsequent sections. But before the methodology is implemented on benchmark, a prerequisite for the methodology is to define the operating domain, within which the approximation/reduced model is valid. Input design, discussed in section 3.1 is designed for distillation column; the operating domain has been finalized by constraint in output purity. A set of input variables (reflux (L) and vapour boilup (V)) is chosen, for which output variables are observed. The product purity for output variables sets the boundary for operating domain.

High purity distillation column

A high purity distillation column is used as test case, on which the approximated and reduced model estimation is applied. The distillation column has following properties; The column has 72 trays, a total condenser and partial reboiler. It is a nonlinear system. The thermodynamics of the column are governed by constant relative volatility. The relative volatility for this specific system is 1.33. Pressure is assumed to be constant. Vapour holdups are considered negligible and liquid holdups are considered

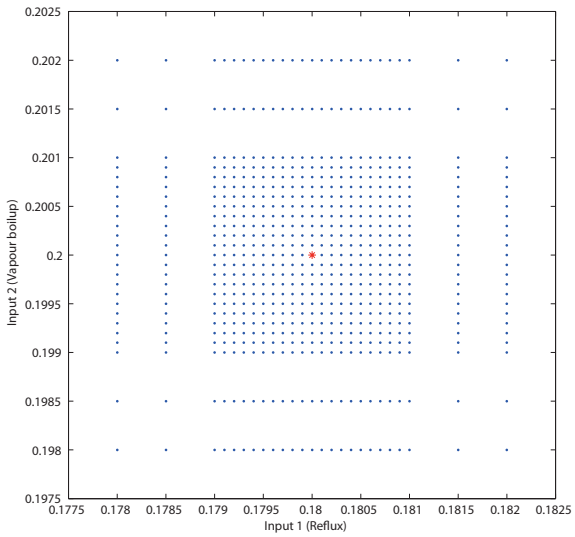


Fig. 10. Figure of operating domain '*input design*'

to be constant. Moreover, the column is assumed to be working with equimolal flow (which results in eliminating energy balances). The distillation benchmark model has been explained in detail by Lévine and Rouchon (1991).

The benchmark is modelled in gPROMS while the approximation model is modelled in MATLAB and SIMULINK. The approximation technique is implemented on the benchmark. The operating domain is finalized (input design) shown in figure 10. The static part consists of lookup table which interpolates the steady-states. The steady-states are fed into the dynamic part. The difference between current state and steady-state is fed to Jacobian block, which is estimated based on state and steady-state data. This represents the linear block as bilinear system.

Two types of changes

There are two types of input signals tested for the validation of approximated (and reduced) model.

a) 'Separation index' (SI) is change in distillation, when both the input variables (reflux rate (L) and vapour boilup (V)) are given *same* steps at the same time, or the rate of flow of distillate D and bottom B does not change.
b) 'Effective Cut Point' (ECP) is change in distillation, when one input variable (reflux rate (L) or vapour boilup (V)) is kept constant and step change is given to the other input variable. This change is known to be highly non-linear for high purity distillation column (in process industry).

Figure 11 show the step in both inputs at the same time (separation index). The approximation model structure (in full state) and reduced order model structure approximate the behavior very well. The mismatch between the approximations and original is cause by offset form lookup table (NL block of approximated model).

Figure 12 show the result comparison of original, full order approximated model and reduced order model, when step in vapour boilup (V) input (effective cut point) is

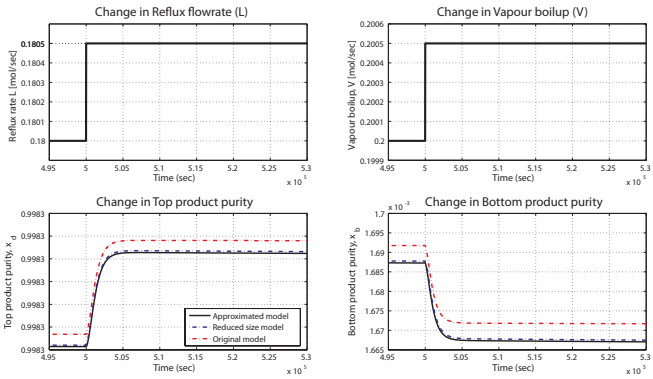


Fig. 11. Results for comparison of SI change between original, approximation '*full*' and '*reduced*' model

applied. The approximation model structure (in full state) and reduced order structure approximates the behavior satisfactorily. There is a very small mismatch in dynamics between the approximation model and original. The offset is acceptable (and sufficiently accurate) for this application.

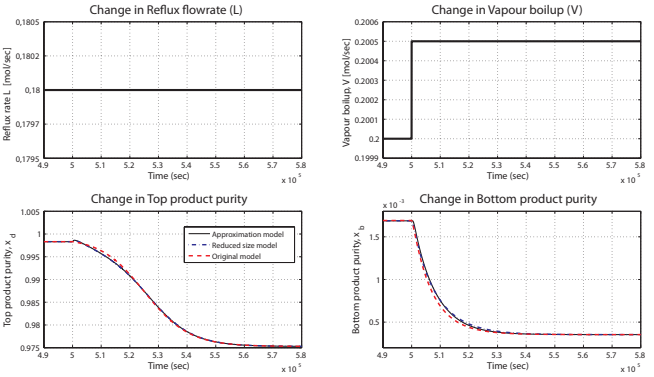


Fig. 12. Results for comparison of ECP change between original approximation '*full*' and '*reduced*' model

5. CONCLUSIONS & FUTURE WORK

In this paper, it is shown that Input-State Hammerstein structure can be derived from a Taylor expansion. The approximation model's accuracy can be improved by including higher order terms. The approximation results were shown for a high purity distillation benchmark are acceptable for the kind of application. Order reduction of 70% is possible using the methodology with satisfactory results (high accuracy).

Work on the following tasks is done presently or is to be considered in future:

i) The computational load reduction for the benchmark example (high purity distillation column) is to be investigated. The computational load and simulation time reduction has to be compared with original NL model.
ii) It is planned to extend the methodology to industrial case. The industrial models make use of dynamic link library (dll) files (as foreign process) to compute different task (such as thermodynamic properties). Such foreign processes buildup overhead costs,resulting in increased computational load. It is anticipated, that transformation of large DAE model to ODE structure will reduce the computational effort (and simulation time), since the algebraic computations are vanished in ODE structure, replaced by NL mapping.

ACKNOWLEDGEMENTS

This work has been supported by the European union within the Marie-Curie Training Network PROMATCH under the grant number MRTN-CT-2004-512441.

REFERENCES

Aling, H., Banerjee, S., Bangia, A., Cole, V., Ebert, J., Emami-Naeini, A., and Jensen, K. (1997). Nonlinear model reduction for simulation and control of rapid thermal processing. *In Proceedings of the American Control Conference*, 2233–2238.

Balasubramhanya, L. and Doyle, F. (2000). Nonlinear model-based control of a batch reactive distillation column. *Journal of Process Control*, 10, 209–218.

Berg, J.V.D. (2005). *Model reduction for dynamic real-time optimization of chemical processes*. OPTIMA, P.O. Box 84115, 3009 CC Rotterdam, The Netherlands.

Billings, S. and Fakhouri, S. (1977). Identification of nonlinear systems- a survey. *IEEE Proc. Pt.D*, 127–272.

Chen, H. (1995). Modelling and identification of parallel nonlinear systems: Structural classification and parameter estimation methods. *P.IEEE*, 83(1), 39–66.

Eskinat, E., Johnson, S., and Luyben, W.L. (1991). Use of hammerstein models in identification of nonlinear systems. *AlChE Journal*, 37, 255–278.

Gani, R., Perregaard, J., and Johansen, H. (1990). Simulation strategies for design and analysis of complex chemical processes. *Trans. IChemE.*, 68 (Part A), 407–417.

Hahn, J. and Edgar, T. (2000). Reduction of nonlinear models using balancing of emperical gramians and galerkin projections. *In Proceedings of the American Control Conference*, 2864–2868.

Harnischmacher, G. and Marquardt, W. (2007). A multivariate hammerstein model for processes with input directionality. *Journal of Process Control*, 17, 539–550.

Harnischmacher, G. (2007). *Block structured modeling for control*. VDI Verlag GmbH, Dusseldorf, Germany.

Lévine, J. and Rouchon, P. (1991). Quality control of binary distillation columns via nonlinear aggregated models. *Chemical Engineering Science*, 27(3), 463–480.

Ling, W. and Rivera, D. (1998). Control relevant model reduction of voltera series models. *Journal of Process Control*, 8(2), 78–88.

Naeem, O., Huesman, A., and Bosgra, O. (2008). Nonlinear model order reduction using input to state hammerstein structures. *In the proceedings of REDUCIT Symposium*.

Norquay, S., Palazoglu, A., and Romangnoli, J. (1999). Application of wiener model predictive control (wmpc) to an industrial c2-splitter. *J. Proc. Control*, 9(6), 461–473.

Perregaard, J. (1993). Model simplification and reduction for simulation and optimization of chemical processes. *Computers and Chemical Engineering*, 17(5/6), 465–483.

A Clean-Coal Control Technology Application Study: Modelling and Control Issues for a Coal Gasifier

S. Bittanti* L. Calloni S. Canevese** A. De Marco* V. Prandoni****

** Dipartimento di Elettronica e Informazione (DEI), Politecnico di Milano, Piazza Leonardo Da Vinci 32, 20133, Milan, Italy (e-mail: sergio.bittanti@polimi.it)*
*** CESI RICERCA, Via Rubattino 54, 20134, Milan, Italy (e-mail: calloni@cesiricerca.it, canevese@cesiricerca.it, prandoni@cesiricerca.it)*

Abstract: The dynamic behaviour of a coal slurry gasifier in an Integrated Gasification Combined Cycle is modelled by means of mass, energy and momentum conservation equations as well as reaction kinetics descriptions. The main phenomena taken into consideration are (i) slurry drying and devolatilisation, (ii) char and volatile gas combustion, char gasification and water-gas shift reaction, and (iii) syngas cooling. The proposed 0-dimensional description is sufficient to capture process dynamics and it is a useful starting point for control design and verification. In particular, basic control strategies are discussed. Both model and control implementation is carried out in the Matlab-Simulink environment. Simulation results are shown to support model reliability and control effectiveness.

Keywords: Process modelling, Process automation, Process control, Process simulators, Power generation, Coal gasification.

1. INTRODUCTION

As it is well known, nowadays the scenario in electric energy production is characterised by a constant increase in demand, a decrease in fossil fuel reserves, more and more demanding restrictions on pollutant levels. Feasible solutions can be increasing efficiency and reducing pollutant emissions in thermoelectric power plants, and contributing to the development of the so called "green energy". Coal can play a major role, especially because of the important amount of its proven reserves worldwide; a main challenge for research is then to develop high-efficiency coal-based energy production systems with near zero emissions. In this paper, reference is made to a 70 MWe coal-fed Integrated Gasification Combined Cycle (IGCC) pilot plant (Fantini et al., 2007), allowing flexible production of electric energy and hydrogen. For the design, work is in progress to build up a simulator of the whole process, in order to obtain reliable predictions of its dynamic behaviour in different operating conditions and to study the operating manoeuvres. Dynamical models of the shift reactor and of the Pressure Swing Adsorption (PSA) unit have already been studied ((Bittanti et al., 2008), (Canevese et al., 2007)). Here, we focus on the gasifier, working out a first-principle model useful for control design.

In Section 2 of this paper, the gasification process is analysed, and its main phases are represented by a dynamical model, based on a thermo-fluid-dynamical and a kinetic-chemical description; such model has been developed in full detail starting from the basic conservation equations and the constitutive equations (including the kinetic equation of char gasification); here, of course, we will present only a partial outline of this model. Section 3 deals with control problems. Section 4 reports some simulation results highlighting control effectiveness. Finally, Section 5 reports some conclusions and hints to future work.

2. PROCESS ANALYSIS

The gasifier under study is an entrained-flow gasifier working at about 65 bar and 1650-1700 K. It is formed essentially by two coaxial cylinders: in the inner one, the gasification process occurs, while the outer one is employed for a first syngas cooling. As illustrated in Fig. 1, the reactor is fed with slurry (pulverized coal mixed with water which can be handled like a liquid fuel) and highly pure oxygen and it produces a gaseous mixture whose main components are CO, CO_2, H_2, N_2, H_2O, and pollutants (COS, H_2S, and dusts). We sketch the overall process of gasification and cooling as composed of the following phases (Smoot and Smith, 1985):

- *drying and devolatilisation*:

slurry can be described as coal powder where each particle is coated with a water film; when it is pumped into the inner cylinder, the high temperature that it meets makes water evaporate, thus yielding dry char, and then makes volatile gases (such as N_2, H_2S, H_2O and several kinds of hydrocarbons, among which CH_4) leave char;

- *oxidation and gasification*:

oxygen burns both the volatile gases and the char, according to the reactions

$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O \quad (volatile\ combustion), \tag{1}$$

$$C + \left(1/\phi\right)O_2 \rightarrow 2\left(1 - 1/\phi\right)CO + \left(2/\phi - 1\right)CO_2, \quad 1 \leq \phi \leq 2 \quad (char\ combustion) \tag{2}$$

(the mechanism factor ϕ indicating whether CO or CO_2 is transported from the particle surface is calculated according to (Wen and Dutta, 1979), (van der Looij, 1988)). The related temperature increase sustains the drying and volatile emission process and the endothermic reactions

$$C + H_2O \rightarrow CO + H_2 \quad (gasification)\,, \tag{3}$$

$$C + CO_2 \rightarrow 2CO \quad (char - CO_2\ reduction)\,, \tag{4}$$

$$CO + H_2O \rightleftarrows CO_2 + H_2 \quad (water - gas\ shift)\,; \tag{5}$$

- *cooling*:

the thus obtained hot syngas (its temperature is around 1760 K) is sent to the bottom of the reactor, where contact with relatively cold liquid water causes a thermal shock which decreases the gas temperature abruptly and stops all reactions still going on; besides, unburned char residuals and char ashes solidify, fall down and are extracted as slag. After bubbling into water, the syngas is pushed by pressure difference (65 bar Vs 62 bar in the considered case) to the outer cylinder, where it is further cooled by a counter-current water spray. Spray temperature and flow rate can be used to control the outlet fluid mixture temperature and humidity. ■

Actually, the first two phases occur in the same region (gasification region). A detailed 3-D description of the phenomena taking place here is suggested, e.g., in (Chen, et al., 2000). Simpler models can be worked out by assuming that both temperature and pressure are uniform in the whole region (0-D assumption). This assumption, motivated by the intense recirculation of gases and adopted also in (Schoen, 1993) for a different type of gasifier, is adopted herein. The proposed model is able to capture the process fundamental dynamics with a low complexity degree, thus ensuring both clear physical insight and short simulation times, in view of the study of control strategies.

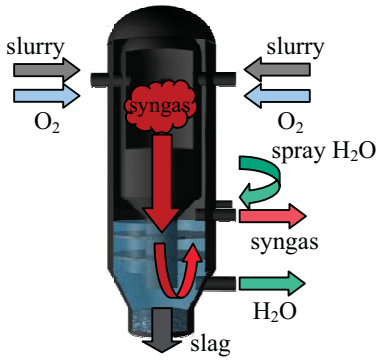


Fig. 1. The GE-Texaco gasifier: schematic view.

The following subsections report the conservation equations employed to model each phase and a reaction kinetic description for the second phase. Table 1 collects the main symbols employed.

Table 1. Nomenclature

Symbol	Description	Unit
c	Specific heat	kJ/(kg·K)
e	Relative energy	kJ/kg
f	Mass fraction	-
h	Relative enthalpy	kJ/kg
$j_k^{(r)}$	Component k stoichiometric coefficient in reaction r=1,..,5	-
p	Pressure	bar
w	Mass flow rate	kg/s
$\tilde{w}$	Molar flow rate	kmol/s
x	Molar fraction	-
A	Area	m^2
L	Length	m
M	Mass	kg
PM_k	Component k molecular weight	kg/kmol
Q	Thermal power	W
S_{ch}	Char active surface	m^2
T	Temperature	K
V	Volume	m^3
β	Mass exchange coefficient	kg/(N·s)
γ	Convective energy exchange coefficient	W/(K·m^2)
ρ	Density	kg/m^3
Ω	Equivalent perimeter	m
Subscripts		
c	Cooling region	
ch	Char	
d/u	Volume under/above the cooling liquid surface	
ev/vol	Resulting from evaporation/devolatilisation	
g/l	Gas/liquid phase	
$g\text{-}l$	Exchange between gas and liquid phases	
in/out	At the inlet/outlet of the region under study	
int/ext	Internal/external chamber	
sat	Saturation	
sh	Shift reaction	
sl	Slurry	
$surf$	Char surface	

2.1 Drying and Devolatilisation

The coal slurry injected into the gasifier forms a jet of length L_{ev} where the water evaporates heated by the hot gases. Then, after water evaporation, the volatile release takes place, due to further heating of the mass of dried coal.

The mass conservation equations for the liquid water and the dry char are respectively

$$\rho_{H_2O}\bar{A}_{H_2O}\dot{L}_{ev} = w_{sl}f_{H_2O,in} - w_{ev}\,, \tag{6}$$

$$\dot{M}_{dry} \equiv \frac{d(\rho_{dry}\bar{A}_{dry}L_{ev})}{dt} = w_{sl}(1 - f_{H_2O,in}) + - M_{dry}u_j/L_{ev} + \rho_{dry}\bar{A}_{dry}\dot{L}_{ev}, \tag{7}$$

where $\bar{A}_{H_2O}$ is the evaporation average equivalent area for water and $\bar{A}_{dry}$ is the equivalent area for char,

$$w_{ev} = \beta_{g-l} L_{ev} \Omega_{ev} (p_{sat}(T_{ev}) - p_{H_2O,g}), \quad (8)$$

$L_{ev}\Omega_{ev}$ is the average evaporation surface and u_j is the particle average velocity in the drying region.

As to energy conservation, a unique average temperature T_{ev} can be adopted for the overall particle (water and char), so

$$\begin{aligned}(\rho_{H_2O}\overline{A}_{H_2O}c_{H_2O} + \rho_{dry}\overline{A}_{dry}c_{dry})[(T_{ev} - T_{sl})\dot{L}_{ev} + \\ + L_{ev}\dot{T}_{ev}] = Q_{ev} - w_{ev} \cdot (h_{g,sat}(T_{ev}) - h_{H_2O}(T_{sl})),\end{aligned} \quad (9)$$

$$Q_{ev} = \gamma_{g-l} L_{ev} \Omega_{ev} \cdot (T_g - T_{ev}). \quad (10)$$

Volatile emission is assumed to occur almost instantaneously after drying, since its dynamics are of the order of few tens of ms ((Thambimuthu and Whaley, 1987), (Kobayashi et al., 1977)); therefore, such dynamics, together with the related heating process, are neglected here. Evaporation is very fast as well, with dynamics of the order of a second at most (Thambimuthu and Whaley, 1987), so it is neglected in the simulations reported in Section 4.

2.2 Oxidation and Gasification

The mass conservation equation for the *k*-th gaseous component in the gasification volume can be written as

$$\begin{aligned}\dot{M}_k = w_{k,in} - w_{g,out} f_{k,out} + \sum_{r=2,3,4} \tilde{w}_{ch}^{(r)} PM_k j_k^{(r)} + \\ + \tilde{w}_{sh} PM_k j_k^{(5)} + \tilde{w}_{CH_4,vol} PM_k j_k^{(1)},\end{aligned} \quad (11)$$

where the component inlet flow rate is

$$w_{k,in} = \begin{cases} w_{ev}, & \text{for } k = H_2O \\ w_{k,vol}, & \text{for } k = CO_2, H_2S \\ 0, & \text{for } k = CO, H_2 \\ w_{N_2,vol} + w_{air} f_{N_2,air}, & \text{for } k = N_2 \end{cases} \quad (12)$$

and $w_{g,out}\, f_{k,out}$ its outlet flow rate. O_2 can be assumed to be completely consumed, and with extremely fast dynamics, by the combustion reactions (1) and (2) (Cotone, 2003); therefore, it is not necessary to write a mass conservation equation for it. For char, which is in the solid phase, we write

$$\dot{M}_{ch} = w_{dry} - \eta_g \frac{w_{g,out} M_{ch}}{\rho_g V_g} + \sum_{r=2,3,4} \tilde{w}_{ch}^{(r)} PM_{ch} j_{ch}^{(r)}, \quad (13)$$

where η_g <<1 is a shape factor (De Marco, et al., 1991) which accounts for the reaction spatial development and which can be identified from experimental results,

$$\tilde{w}_{ch}^{(r)} = \frac{K_r M_{ch} M_r}{V_{\text{int}}}, r = 3,4, M_3 = M_{H_2O}, M_4 = M_{CO_2}, \quad (14)$$

$$\tilde{w}_{ch}^{(2)} = \phi \tilde{w}_{O_2}. \quad (15)$$

$\tilde{w}_{O_2}$ is the difference between the O_2 inlet molar flow rate and the O_2 molar flow rate consumed by reaction (1). We remark that (14) is taken from the unreacted-core-shrinking model described in (Wen and Chaung, 1979); K_r accounts both for reaction kinetics, by means of an Arrhenius-type term, and for component diffusion between the gaseous bulk and the char reacting surface. The reversible shift reaction (5) kinetics is assumed to be at equilibrium on the char surface, so that the gaseous components surface diffusion is the limiting phenomenon, except for H_2, whose diffusion coefficient can be assumed as ideally infinite.

The energy conservation equations, for the gaseous mixture in the gasification volume and for solid char, read as

$$M_g c_g \dot{T}_g = Q_{g,in} - \tilde{w}_{sh} \Delta H_{sh} + w_{vol} \Delta H_{vol} + Q_{ch-g} - Q_{ev}, \quad (16)$$

$$M_{ch} c_{ch} \dot{T}_{ch} = -Q_{ch,in} + Q^{(2)} - Q^{(3)} - Q^{(4)} - Q_{ch-g}, \quad (17)$$

where

$$\begin{aligned}Q_{g,in} = w_{ev} \cdot (h_{H_2O}(T_{ev}) - h_g(T_g)) + w_{O_2} \cdot (h_{O_2}(T_{ev}) + \\ - h_g(T_g)) + w_{ev} \cdot (h_{ev}(T_{ev}) - h_g(T_g)) + w_{dry,out,vol} \cdot \\ \cdot (e_{ch}(T_{ev}) - h_g(T_g)),\end{aligned} \quad (18)$$

$$Q_{ch-g} = \gamma_{ch-g} S_{ch} \cdot (T_{ch} - T_g), \quad (19)$$

$$\begin{aligned}Q_{ch,in} = w_{dry,out,vol}(e_{ch}(T_{ch}) - e_{ch}(T_{ev})) + w_{O_2} \cdot \\ \cdot (e_{ch}(T_{ch}) - h_{O_2}(T_{ev})) + w_{ev}(e_{ch}(T_{ch}) - h_{H_2O}(T_{ev})) + \\ + w_{vol} \cdot (e_{ch}(T_{ch}) - h_{vol}(T_{ev})),\end{aligned} \quad (20)$$

$$Q^{(r)} = \tilde{w}_{ch}^{(r)} PM_{ch} \Delta H^{(r)}, r = 2,3,4 \quad (21)$$

and $\Delta H^{(2)}$ is a function of ϕ, since CO and CO_2 have different heating values.

Finally, as to pressure, the perfect gas law yields

$$p_g = \frac{R T_g}{V_g} \sum_k \frac{M_k}{PM_k}, k = H_2O, H_2, CO, CO_2, N_2, H_2S. \quad (22)$$

2.3 Cooling

The cooling volume is composed of two regions, separated by the liquid water surface: a lower "pool" region, where gas bubbles into water, and an upper spray region, where water droplets further cool humid gas leaving the pool region. Mass and energy conservation equations will be written separately for the gas mixture and for liquid water in each region.

As to the overall gas and to its water part under the surface, first of all, one can write mass conservation as

$$\dot{M}_{gd} = w_b - w_{bout} - w_{g-l}, \quad (23)$$

$$\dot{M}_{H_2O,gd} = w_b f_{H_2O,in} - w_{bout} f_{H_2O,out} - w_{g-l}, \quad (24)$$

where w_b is the gasification outlet flow rate, w_{bout} the flow rate leaving the liquid surface and w_{g-l} the exchanged (usually condensating, anyway) water flow rate inside the gas phase.

Adopting a steady-state model for the gas phase under the free liquid surface, from (23) and (24) we can write

$$w_{bout} = w_{g,out} \left(1 - f_{H_2O,in}\right) \Big/ \left(1 - f_{H_2O,out}\right) \quad (25)$$

$$w_{g-l} = w_b \left(f_{H_2O,out} - f_{H_2O,in}\right) \Big/ \left(1 - f_{H_2O,out}\right), \quad (26)$$

and, adopting an equilibrium model,

$$f_{H_2O,out} = \frac{PM_{H_2O}p_{sat}(T_{ld})}{\overline{PM_{gd}}[p_c - p_{sat}(T_{ld})] + PM_{H_2O}p_{sat}(T_{ld})}, \quad (27)$$

where $\overline{PM_{gd}}$ is the average molecular weight of the "dry" gas, i.e. without considering its water contents.

As to energy conservation, we have

$$M_g c_g \dot{T}_{gd} = Q_{g,ex} - Q_{gl} - w_b(h(p_g, T_{in}, f_{gd}) + -h(p_c, T_{gd,out}, f_{gd,out})), \quad (28)$$

$$Q_{g,ex} = w_{g-l}(h_{g,sat}(T_{ld}) - h_g(T_{gd})), \quad (29)$$

$$Q_{gl} = \gamma_{g-l} S_{g-l}(T_{gd} - T_{ld}). \quad (30)$$

In (28), the last product term in the right-hand member describes the heat lost to decrease the inlet gas temperature; p_c is the upper cooling region pressure.

As to the liquid water mass M_{ld} under the surface, one has

$$\dot{M}_{ld} = w_{H_2O\_down} + w_{g-l} - w_{ld,out}, \quad (31)$$

where the three terms on the right-hand member are due to water falling down from the cooling upper region into the pool because of gravity, to the condensating water and to the outlet water respectively. In particular,

$$w_{H_2O\_down} = M_{w\_spr} / \tau_{down}, \quad (32)$$

where the average delay factor τ_{down} (Lydersen, 1983) models the residence time (typically of the order of a few seconds) of spray water (whose mass is M_{w_spr}) in the spray region. $w_{ld,out}$ is a control variable which can be employed for level regulation, by means of a suitable valve.

For the liquid water temperature T_{ld} under the surface, one can write

$$M_l c_l \dot{T}_{ld} = -Q_{H_2O\_down} + Q_{l,ex} + Q_{gl}, \quad (33)$$

$$Q_{H_2O\_down} = w_{H_2O\_down}(h_{H_2O}(T_{lu}) - h_{H_2O}(T_{ld})), \quad (34)$$

$$Q_{l,ex} = w_{g-l}(h_{g,sat}(T_{ld}) - h_l(T_{ld})). \quad (35)$$

For gas mass conservation above the free surface, one has

$$\dot{M}_{gu} = w_{bout} - w_{gas,mix} - w_{cnd}, \quad (36)$$

where $w_{gas,mix}$ is the overall gasifier outlet syngas flow rate, which is assumed, for simplicity, to be regulated by a critical valve, and

$$w_{cnd} = \frac{6M_{lu}\beta_{g-l,u}}{\rho_{lu} d_{drop}} \cdot \left(\frac{M_{steam} R T_{gu}}{PM_{H_2O} V_{spr}} - p_{sat}\left(\frac{T_{lu} + T_{w\_spr}}{2} \right) \right) \quad (37)$$

is due to humid gas condensation. T_{w_spr} is the spray inlet temperature, d_{drop} the average spray drop diameter (water drops are assumed as spherical), V_{spr} the spray region volume. M_{steam} can be derived from the conservation equation

$$\dot{M}_{steam} = w_{bout} f_{H_2O,out} - w_{cnd} - w_{gas,mix} M_{steam} / M_{gu}. \quad (38)$$

For gas energy conservation above the free surface, one has

$$M_g c_g \dot{T}_{gu} = -Q_{g,sc} - Q_{exch} + w_{bout}(h(T_{gd}, p_c) - h(T_{gu}, p_c)), \quad (39)$$

$$Q_{g,sc} = w_{cnd}(H_{g,sat}(T_{lu}) - h(T_{gu})), \quad (40)$$

$$Q_{exch} = \gamma_{g-l,u} S_{g-l,u}(T_{gu} - T_{lu}). \quad (41)$$

For the liquid water mass M_{lu} and temperature T_{lu} above the surface, respectively, one can write

$$\dot{M}_{lu} = -w_{H_2O,down} + w_{cnd} + w_{spr}, \quad (42)$$

$$M_l c_l \dot{T}_{lu} = Q_{l,sc} + Q_{exch} + w_{spr}(h(T_{w\_spr}) - h(T_{lu})), \quad (43)$$

where w_{spr} is the inlet spray water flow rate and

$$Q_{l,sc} = w_{cnd}(H_{g,sat}(T_{lu}) - h_l(T_{lu})). \quad (44)$$

Again, pressure in the cooling chamber can be derived from the ideal gas law:

$$p_c = R T_{gu} M_{gu} / (V_{spr} \overline{PM_{gu}}). \quad (45)$$

2.4 Model Verification

As for the coal composition and steady-state nominal conditions, we have made reference to the data published in (Cotone, 2003). The main model parameters, especially for the correlations, have been drawn from the literature as well. In particular, the K_r's in (14) have been taken from (Wen and Chaung, 1979). This way, our model has been fully specified. For its verification, the molar fractions supplied by the model have been compared with the molar fractions in the literature. In Table 2 such comparison is carried out by referring to the situation occurring after the thermal shock at T=1077 K.

Table 2. Comparison between literature data (Cotone, 2003) and simulation results at nominal steady state

	Reference data	Simulation results	Relative Error
x_{CO}	34%	38%	12%
x_{H_2O}	14.9%	13%	13%
x_{CO_2}	16%	15.9%	1%
x_{H_2}	33%	30.6%	7%
x_{N_2}	1.8%	2%	11%

3. CONTROL ISSUES FOR THE GASIFICATION PLANT

In this work, attention is focused on problems related to fulfilling the electrical network's needs, such as supplying the requested power variations, in normal operation (load following) or in emergency conditions, taking part in primary frequency control, or even contributing to secondary frequency control. In particular, the problem of coordinate control is dealt with here: the gasifier load and the global plant operating conditions are mastered so as to satisfy power requests from the electrical network, while preserving plant integrity and correct operation, of course. A simplified plant scheme is shown in Fig. 2. It is composed of three main parts: (i) the gasification island; (ii) a lower pressure system for the syngas treatment, together with thermal energy recovery; (iii) a conventional gas turbine, with its own fuel feed system

controlled by valve v_3. A valve, v_2 (or an expansor for power recovery), connects part (i) and part (ii). The symbol θ is adopted for actuator command signals.

We now focus on the problem of supplying fast and relatively large power variations, in order to fulfil the network's requests. For this purpose, the control scheme of Fig. 3 can be considered. Fast power variations are obtained by acting on valve v_3, regulating the turbine inlet flow rate (as in conventional power plants). By means of two feed-forward actions (FFW in Fig. 3), such variation results in corresponding changes in command signals θ_1 and θ_2. Signal θ_1 controls the slurry flow rate as well as the oxygen flow rate. θ_2 determines the flow rate of the outlet syngas. The two feed-forward actions have to be designed so as to keep constant pressure p_1, at the gasifier outlet, pressure p_2, at the turbine inlet, and temperature T_g inside the gasifier. To this purpose, in Fig. 3 a decentralized control scheme is proposed, where control signals θ_1 and θ_2 are adopted for the regulation of p_1 and p_2 (dashed rectangle of Fig. 3). Alternatively, one can resort to a centralized controller, here omitted for reasons of conciseness. As for T_g, the control action is manually operated and indicated in Fig. 3 by a dash-dotted rectangle.

Note that, for the overall control system, variations of θ_3 can be seen as main (measurable) disturbances.

We now conclude with some observations about the regulation problems.

Variable θ_1 acts simultaneously on the slurry and O_2 flow rates. Here the main objective is to keep constant the ratio between the two flow rates. However, these cannot be varied simultaneously, in order to avoid excessive over- or under-elongations in the gasification temperature. More precisely, when there is a load variation, the corresponding variation of the O_2 flow rate must take place with some delay after the variation of the slurry flow rate. The reason is that the oxygen reacts extremely fast, so that the temperature variation occurs abruptly. Such delay is represented in Fig. 3 as well (lag).

Let us finally consider the 2x2 MIMO system where θ_1 and θ_2 are the input signals, and p_1 and p_2 the output signals.

Conventional power plant operating experience would suggest regulating independently p_1 by the gasifier load and p_2 by valve v_2, as shown at the top of Fig. 3. However, the variables under study are rather interacting with each other: for instance, increasing the gasifier inlet load implies an increase in pressure p_1, which makes valve v_2 flow rate increase and therefore pressure p_2 increase as well. The degree of coupling is *quantitatively* captured by the relative gain matrix RGA, whose elements are not far from 0.5. Therefore, a centralized control solution, carried out by a forward decoupling technique, has been also analysed. The open-loop SISO transfer functions employed for controllers tuning have been identified from the system step responses around the chosen steady-state nominal point (see Section 2.4). Summing up, both centralized and decentralized controllers have been designed and simulated. Simulation results are reported in the subsequent section.

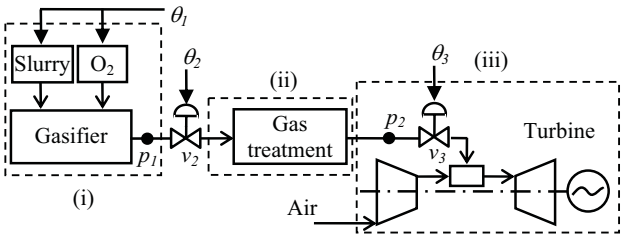


Fig. 2. The controlled simplified plant scheme.

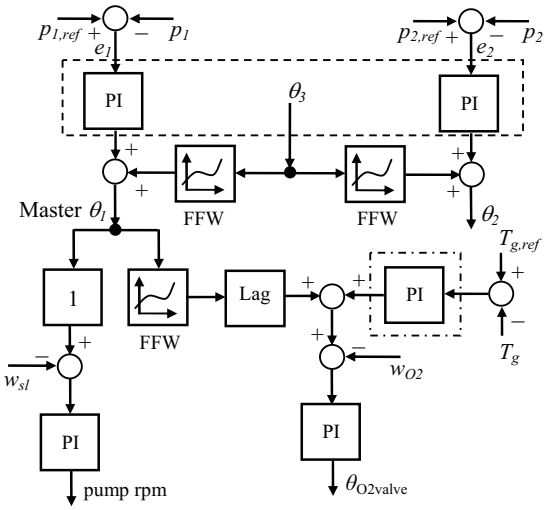


Fig. 3. A controller structure for the plant.

4. SIMULATION RESULTS

The simplified plant depicted in Fig. 2 has been simulated in the Matlab-Simulink environment, with the gasifier model, in particular, implemented by means of an S-function written in the C++ language. Simulations have been carried out both in open loop and in closed loop by considering both the centralized and the decentralized schemes. Integration has been executed in the continuous-time domain, by resorting to the standard Matlab algorithms.

Some of the results of a dynamical simulation with the centralized controller are now reported. Starting from the steady-state nominal conditions, a positive 10% step on the turbine valve position is given, at time t = 1500 s; this simulates a variation of power request from the network, so that the turbogas control system requires more inlet fuel flow rate. The top of Fig. 4 shows the responses of pressure p_1 and temperature T_g: as expected, their steady-state values are unaffected by the disturbance, and elongations around such values are very small. In Fig. 4 - bottom the control variables slurry and oxygen flow rates are depicted: they both increase in a rather slow manner, so as to preserve the integrity of the gasifier itself and of the other devices (with their dynamic operating constraints). Finally, it turns out that also the opening of valve v_2 exhibits a smooth behaviour.

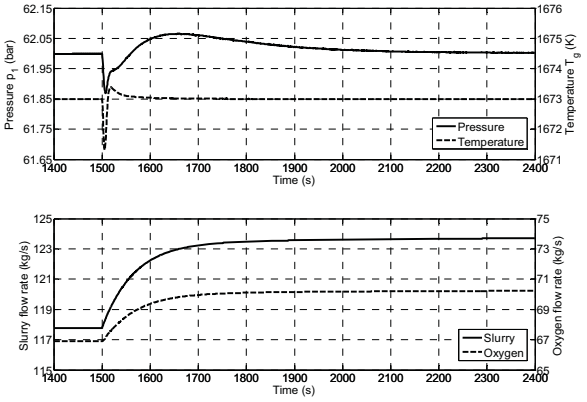


Fig. 4. Simulation results with centralized control.

5. CONCLUSIONS

A control-oriented first-principle dynamical model for a coal slurry gasifier has been proposed. Model parameters have been identified from literature data. Classical control schemes have been proposed for the 2x2 problem of controlling gasifier pressure and turbine inlet pressure by the gasifier inlet slurry and O_2 flow rates and by the gasifier outlet syngas flow rate. Simulations have shown satisfactory performance for disturbance rejection. Also the set-point tracking (not presented in this paper, for brevity) leads to good results. Future activities include model validation in transient conditions and integration of this model and models of other plant devices (see Section 1) into an overall IGCC plant simulator. Then, control strategies will be studied concerning the interaction between the gasifier and the other devices, in normal operating conditions and during startups/shutdowns; such strategies will be implemented both by standard techniques, based on SISO PID controllers, and by more involved MIMO techniques.

ACKNOWLEDGEMENTS

We are grateful to P. D'Adamo and P. De Francesco, formerly students at Politecnico di Milano, for their collaboration in the phase of parameter tuning.

This work has been financed by the Research Fund for the Italian Electrical system under the Contract Agreement between CESI RICERCA and the Ministry of Economic Development - General Directorate for Energy and Mining Resources stipulated on June 21, 2007 in compliance with the Decree n. 73 of June 18, 2007. Research has also been supported by the Italian National Research Project "New Techniques of Identification and Adaptive Control for Industrial Systems" and partially by CNR-IEIIT.

REFERENCES

Bittanti, S., Canevese, S., De Marco, A., Prandoni, V., and Serrau, D. (2008). Towards clean-coal control technologies: modelling conversion of carbon oxide into hydrogen by a shift reactor, *17th IFAC World Congress*, Seoul, South Korea.

Canevese, S., De Marco, A., Murrai, D., and Prandoni, V. (2007). Modelling and control of a PSA reactor for hydrogen purification, *ICPS'07 (1st IFAC Workshop on Convergence of Information Technologies and Control Methods with Power Plants and Power Systems)*, Cluj-Napoca, Romania.

Chen, C., Horio, M., and Kojima, T. (2000). Numerical simulation of entrained flow coal gasifiers. Part I: modeling of coal gasification in an entrained flow gasifier. *Chemical Engineering Science*, 55 (18), 3861-3874.

Cotone, P. (2003). Section D - Basic information for each alternative. In Domenichini, R., *IEA GHG Gasification power generation study, Final Report, Rev. 1*.

De Marco, A., De Michele, G., Miccio, M., Prandoni, W., and Traniello Gradassi, A. (1991). A model for the dynamic simulation of a pilot fluidized bed combustor: Predictions and first validation, *Proceedings of the 11th International Conference on Fluidized Bed Combustion*, vol. 1, 433–438. ASME, New York, NY.

Fantini, V., Mazzocchi, L., Moia, F., Prandoni, V., and Savoldelli, P. (2007). Pre-feasibility study of a flexible hydrogen-electricity co-production IGCC coal-fed plant with CO_2 capture and sequestration, *Third International Conference on Clean Coal Technologies for our Future (CCT 2007)*, Cagliari, Italy.

Kobayashi, H., Howard, J.B., and Sarofim, A.F. (1977). Coal devolatilization at high temperatures, *15th Symposium (International) on Combustion*, 16 (1), 411-425.

Lydersen, A.L. (1983). *Mass transfer in engineering practice*. John Wiley & Sons.

Schoen, P. (1993). *Dynamic modeling and control of integrated coal gasification combined cycle units*. PhD thesis, Delft University of Technology, Delft, The Netherlands.

Smoot, L.D. and Smith, P.J. (1985). *Coal combustion and gasification.* Plenum Press, New York, NY.

Thambimuthu, K.V. and Whaley, H. (1987). The combustion of coal-liquid mixtures. In C.J. Lawn (ed.), *Principles of combustion engineering for boilers*, Chapter 4, 402-408. Academic Press, London.

van der Looij, J.M.P. (1988). *Dynamic modeling and control of coal fired fluidized bed boilers*. PhD thesis, Delft University of Technology, Delft, The Netherlands.

Wen, C.Y. and Chaung, T.Z. (1979). Entrainment coal gasification modeling. *Industrial & Engineering Chemistry Process Design and Development*, 18 (4), 684-695.

Wen, C.Y. and Dutta, S. (1979). Rate of coal pyrolysis and gasification reactions. In C. Wen, Lee (ed's), *Coal conversion technology*, 57-170. Addison-Wesley Publishing Co., Reading, MA.

Identification of Reaction Mechanisms with a Dynamic PFR Model ⋆

Jan C. Schöneberger * **Harvey Arellano-Garcia** *
Holger Thielert ** **Günter Wozny** *

* *Berlin Institute of Technology,Chair of Process Dynamics and Operation, Berlin, Germany (e-mail: jan.schoeneberger@tu-berlin.de)*
** *Uhde GmbH, Dortmund, Germany*

Abstract: In this work, a dynamic model of a catalytic fixed bed reactor (FBR) based on partial differential equations (PDE) is introduced and used for the identification of reaction mechanisms which take place during the oxidation of sulfur dioxide over a vanadium pentoxid catalyst. The measured data is collected from a pilot plant, which uses commercial sized catalyst particles. In order to reduce the experimental effort, a developed framework based on the methods of nonlinear optimal experimental design is applied using a steady state FBR model. The systematic procedure is improved using a dynamic reactor model. This makes the time dependent measurement data valuable for the identification procedure.

Keywords: Catalytic fixed bed reactor (FBR), reaction mechanisms, partial differential equations (PDE), parameter identification, optimal experimental design.

1. INTRODUCTION

Catalytic gas phase reactions have a high relevance in chemical engineering. The majority of chemical processes will not be profitable and in some cases not even viable without the usage of catalysts. An important application represents the utilization of FBR for waste gas treatment processes. In this work, the oxidation of sulfur dioxide to sulfur trioxide is considered, which is converted with water to sulphuric acid. It should be noted that huge effort is made in the development of new catalysts with a higher activity. Commonly, new catalysts are designed and tested at micro scale, i.e. a pulverized catalyst. However, the catalyst used in industrial plants are much larger and the previously identified mechanisms and kinetic parameters can not be transferred without further investigations in a scale up procedure w.r.t. the reactor layout. Thus, measurements with commercial catalyst particles are inevitable implying a high experimental effort. The sized particles require a larger reactor diameter, and thus, high gas flow rates are necessary in order to hold the operation conditions close to the industrial scale reactor. Moreover, corresponding requirements for process automation and safety engineering have to be met. Consequently, in order to reduce the experimental effort while reaching a desired model quality, methods of nonlinear optimal experimental design can be applied. In addition, due to the fact that a good deal of data enhances the mechanism and parameter identification, the proposed framework can still be improved when using a model, which describes the pilot plant dynamic behavior. In this work, a homogeneous and a two-phase FBR model are presented. The latter can be used for dynamic simulations. Both are compared to each other based on a set of measured data taken from an identification campaign. The experiments were performed in a pilot plant (Fig. 1) using a commercially available vanadium pentoxid catalyst.

⋆ This work is supported by the Max-Buchner-Forschungsstiftung.

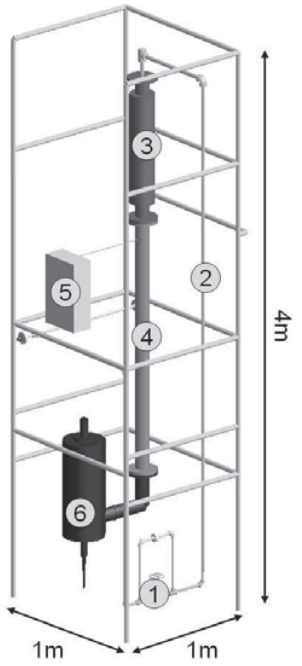


Fig. 1. Pilot plant set-up: (1) Quality measurement, (2) Heating section, (3) Reactor, (4) Tube bundle cooler, (5) Quality measurement, (6) Bubble column.

2. PILOT PLANT DESCRIPTION

The core of the pilot plant is the tubular reactor which has a diameter of $D_R = 0.1$ m and a length of $L_R = 1$ m.

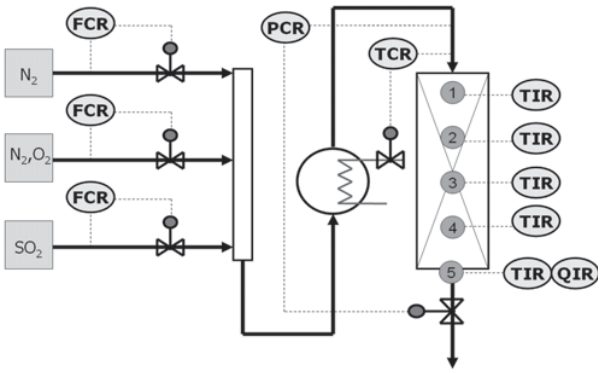


Fig. 2. Control structure: (FCR) Flow control and measurement, (TCR) Temperature control and measurement, (PCR) Pressure control and measurement, (TIR) Temperature measurement, (QIR) Quality measurement.

The diameter results from the size of the catalyst particle, which is about $D_P = 0.005$ m for the examined catalyst. It represents the minimum diameter required in order to get a uniform particle distribution. The reactor consists of five beds of catalyst packing, where each is $H_B = 0.1$ m high. The additional length is used to reduce in and outlet effects. The reactor is operated nearly adiabatic using an electric heated isolation. The majority of the available space is occupied by the secondary units such as the heating, cooling and gas scrubbing. The last two steps are combined in a bubble column. Fig. 1 gives an idea of the pilot plant set-up. The pilot plant is automated using ABB Freelance. Beside the implemented safety engineering procedures, the reactor inlet variables are controlled using the structure given in Fig. 2. The gas flow rate $\dot{V}$ can be varied between 200 and 400 norm liter per minute at inlet temperatures up to $T_{in} = 500$°C.

3. REACTOR MODELS

Two reactor models are presented, which describe the FBR behavior, namely the *homogeneous model* and the *two-phase model*. The chemical reaction is modeled with six different reaction mechanisms for which several steps of parameter identification and model discrimination are performed in order to find the best suitable mechanism. Besides the kinetic parameters in the reaction rate equations, the reactor models contain several parameters. Some of them can be obtained from the particles geometry e.g. the packing porosity ϵ, the relative particle diameter D_P, the specific surface area a, and the catalyst density ρ_C, other parameters are derived from property functions such as the component heat capacities $c_{P,i}$ and $c_{V,i}$, the heat of reaction Δh_R, and the thermodynamic equilibrium constant, K_P. The heat loss can be determined from experimental runs without reactions and is then calculated with the estimated heat transfer coefficient, k_W, and the measured wall temperature, T_W. Four components are considered in the model with the indices as given in Tab. 1.

Table 1. Component indices

Index	1	2	3	4
Component	$\mathbf{SO_2}$	$\mathbf{O_2}$	$\mathbf{SO_3}$	$\mathbf{N_2}$

3.1 Homogeneous Model

The model equations composed of mass balances, (1), energy balance, (2), and momentum balance, (3), imply the assumption of a plug flow profile in the reactor and an instant heat and mass transfer between the gas and the solid phase. This means that the temperature of the catalyst particle is the same as the gas phase bulk temperature. Due to this issue the model can not be used for the description of the dynamic behavior, since the heat capacities of gas and solid differ a lot. A detailed description and derivation of the homogeneous model can be found in Arellano-Garcia et al. (2007).

$$\frac{dF_i}{dz} = \nu_i \dot{r} \cdot \rho_C (1-\epsilon) \frac{\pi}{4} D_R^2 \tag{1}$$

$$\frac{dT}{dz} = \frac{-\dot{r}\Delta h_R(T)\rho_C(1-\epsilon)\frac{\pi}{4}D_R^2 - k_W \pi \; D_R (T - T_W)}{\sum F_i c_{P,i}} \tag{2}$$

$$\frac{dP}{dz} = -\frac{\sum F_i M_i}{\rho \frac{\pi}{4} D_R^2 D_P} \left(\frac{1-\epsilon}{\epsilon^3}\right) \cdot \left[\frac{150(1-\epsilon)\eta}{D_P} + 1.75 \frac{\sum F_i M_i}{\frac{\pi}{4} D_R^2}\right] \tag{3}$$

In the model equations, F_i denotes the component flow rates, ν_i the stoichiometric coefficient, and M_i, the molar mass. T stands for the gas and the catalyst temperature, P for the gas phase pressure. The viscosity η is calculated assuming an ideal mixing.

Following the simulation results based on this model, the pressure drop can be neglected for the given reactor set-up and is not considered anymore in the two-phase model.

3.2 Two-Phase Model

The key idea of developing a two-phase model is to perform dynamic simulations of the reactor behavior in order to include time variant measurement data in the parameter estimation procedure. Due to the varieties in the resulting time constants in the energy balances, which are mainly influenced by the heat capacity and density in its corresponding phase, a split modeling of the gas and solid phase becomes inevitable. The resulting equation system comprises mass and energy balances for the two phases,(4)- (7), which are coupled via mass and heat transfer correlations,(8)-(10). Instead of the component flow rates (see (1)), the concentrations in the gas phase $c_{G,i}$ and the solid phase $c_{C,i}$ are used here as state variables.

$$\frac{dc_{G,i}}{dt} = D_{ax_i} \frac{\partial^2 c_{G,i}}{\partial z^2} - \frac{\dot{V}}{\epsilon \frac{\pi}{4} D_R^2} \frac{\partial c_{G,i}}{\partial z} - a \frac{1-\epsilon}{\epsilon} \beta (c_{G,i} - c_{C,i}) \tag{4}$$

$$\frac{dc_{C,i}}{dt} = \nu_i \dot{r} \rho_k + a \, \beta (c_{G,i} - c_{C,i}) \tag{5}$$

$$\sum_{i=1}^{4} c_{G,i} c_{V_i} \frac{dT_G}{dt} = a \frac{1-\epsilon}{\epsilon} \alpha (T_C - T_G) - \frac{\dot{V}}{\epsilon \frac{\pi}{4} D_R^2} \cdot \sum_{i=1}^{4} c_{G,i} c_{P_i} \frac{\partial T_G}{\partial z} - k_W \frac{4}{\epsilon \; D_R} (T_G - T_W) \tag{6}$$

$$\frac{dT_C}{dt} = \frac{a\alpha}{\rho_C c_{P,C}} (T_G - T_C) - \frac{\Delta h_R \dot{r}}{c_{P,C}} \tag{7}$$

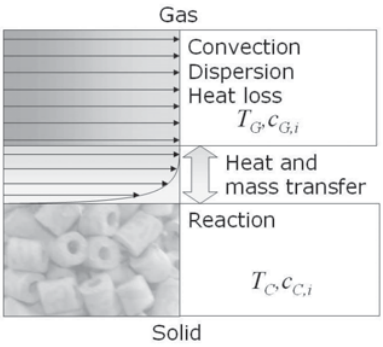


Fig. 3. Control volume of the two-phase model.

The control volume and the considered state variables of the two phases are shown in Fig. 3. The transfer coefficients (α, β, D_{ax}) are calculated with dimensionless numbers (Nusselt number Nu, Sherwood number Sh, axial Peclet number Pe_{ax}) and correlations taken from literature, see e.g. Fogler (2006). Here, $D_{ab,i}$ stands for the diffusion coefficient of the component i in nitrogen and $\bar{\lambda}$ for the mixtures heat conductivity.

$$D_{ax} = \frac{\dot{V} D_P}{Pe_{ax} \epsilon \frac{\pi}{4} D_R^2} \tag{8}$$

$$\beta_i = \frac{Sh_{\epsilon,i}\, D_{ab_i}}{D_P} \tag{9}$$

$$\alpha = \frac{\bar{\lambda} Nu}{D_P} \tag{10}$$

On the one hand, a dynamic simulation becomes possible with the inclusion of the solid phase, but on the other hand, new unknown or not well-known parameters are introduced. Three parameters are used in the transport correlations, one for each equation. In addition, two parameters depend on the catalyst particle properties, namely, the catalyst heat capacity, $c_{P,C}$, and the specific surface area, a. These additional parameters are to be determined from the measurement data, but they are strong correlated with the reaction rate, and thus, with the kinetic parameters of the reaction mechanisms. This problem can be overcome by running experiments without a reaction, i.e. with pure nitrogen or air. By this means the parameter in the heat transfer correlation, and the catalyst properties can be determined independently.

3.3 Reaction Mechanisms

In the open literature, plenty of different reaction mechanisms have been published for the oxidation of sulfur dioxide over vanadium pentoxid, we refer to Mezaki and Kadlec (1972) for an overview. In this work, the five of the most promising rate laws are selected. Additionally, a generic power law mechanism is also considered. All the rate equations describe the reaction stated in equation (11). They are functions of the components partial pressures, P_i, and the temperature, T, which affects the calculation of the velocity constant k, and the equilibrium constants K_P, and K.

$$\mathbf{SO_2} \; + \; \frac{\mathbf{1}}{\mathbf{2}} \, \mathbf{O_2} \; \rightleftharpoons \; \mathbf{SO_3} \tag{11}$$

Power Law: In the rate equation for the power law mechanism (12), the component exponents a, b, and c are treated as model parameters and have to be determined from the measurements. The temperature dependent velocity constant k is calculated with equation (13), in which the kinetic parameters p_1 and p_2 have to be identified.

Rate 1:

$$\dot{r} = k \left(P_{SO_2}^a P_{O_2}^b P_{SO_3}^c - \frac{P_{SO_2}^{(a-2)} P_{O_2}^{(b-1)} P_{SO_3}^{(c+2)}}{K_P} \right) \tag{12}$$

$$k = exp\left(\frac{p_1}{T} + 0.5 \; ln(T) + p_2 \right) \tag{13}$$

Mechanistic Rate Equations: The rate equations (14) to (18) can be derived assuming a liquid metal phase on the catalyst particle formed by vanadium, in which the supplied oxygen is dissolved. This is a widely accepted assumption for the oxidation over vanadium pentoxid. The differences in the reaction rates result from different mechanisms of the catalyst activation with oxygen.

Rate 2:

$$\dot{r} = \frac{k K P_{SO_2} P_{O_2}^{1/2}}{\left[P_{SO_3}^{1/2} + \left(K P_{SO_2} \right)^{1/2} \right]^2} \left(1 - \frac{P_{SO_3}}{K_P P_{SO_2} P_{O_2}^{1/2}} \right) \tag{14}$$

Rate 3:

$$\dot{r} = \frac{k K P_{SO_2} P_{O_2}}{\left[P_{SO_3}^{1/2} + \left(K P_{SO_2} \right)^{1/2} \right]^2} \left(1 - \frac{P_{SO_3}^2}{K_P^2 P_{SO_2}^2 P_{O_2}} \right) \tag{15}$$

Rate 4:

$$\dot{r} = \frac{k K P_{SO_2} P_{O_2}^{1/2}}{\left[P_{SO_3}^{1/2} + \left(K P_{SO_2} \right)^{1/2} \right] P_{SO_3}^{1/2}} \left(1 - \frac{P_{SO_3}}{K_P P_{SO_2} P_{O_2}^{1/2}} \right) \tag{16}$$

Rate 5:

$$\dot{r} = \frac{k \left(K P_{SO_2} \right)^{1/2} P_{O_2}}{\left[P_{SO_3}^{1/2} + \left(K P_{SO_2} \right)^{1/2} \right]} \left(1 - \frac{P_{SO_3}^2}{K_P^2 P_{SO_2}^2 P_{O_2}} \right) \tag{17}$$

Rate 6:

$$\dot{r} = \frac{k K P_{SO_2} P_{O_2}}{\left[P_{SO_3}^{1/2} + \left(K P_{SO_2} \right)^{1/2} \right] P_{SO_3}^{1/2}} \left(1 - \frac{P_{SO_3}^2}{K_P^2 P_{SO_2}^2 P_{O_2}} \right) \tag{18}$$

All these rate equations utilize the same approach for the velocity constant k, and the equilibrium of the vacant sites K, which are given in the equations (19) and (20).

$$k = exp\left(p_1 - \frac{p_2}{T} \right) \tag{19}$$

$$K = exp\left(p_3 - \frac{p_4}{T} \right) \tag{20}$$

3.4 Numerical Solution

In order to keep the computational effort low, both reactor models were discretized using the orthogonal collocation (OC), see Schöneberger et al. (2009). In the case of the homogeneous model this leads to an algebraic equation system (AE) and in the case of the two-phase model to a system of ordinary differential equations (ODE). The

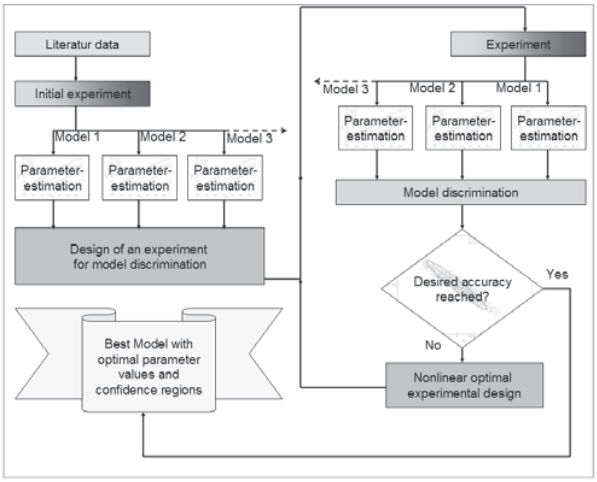


Fig. 4. Model identification framework.

AE is solved with a Newton-Raphson step and the ODE is integrated with an OC based Runge-Kutta algorithm. The use of the numerical solution in an optimization framework requires a very robust solution algorithm, in particular, when the free variables are positioned in exponential terms such as in the case of kinetic parameter estimation. Therefore, specialized initial value generation algorithms and step size control algorithms are required, see Schöneberger et al. (2007). Anyhow, the solution of the homogeneous model is more robust. Consequently, the first parameter estimation is performed with this model in order to get good initial parameter values for the two-phase model.

4. MECHANISM IDENTIFICATION FRAMEWORK

The proposed framework in Fig. 4 is based on the methods of nonlinear experimental design. It is similar to the model building framework proposed by Franceschini and Macchietto (2008), but with some improvements regarding the specific problem. There is only one experiment designed for model discrimination, and this is performed in the beginning of the identification procedure. Further experiments are exclusively designed in order to improve the parameter accuracy until the parameter spreading is in an acceptable region. Please note that all six reaction rates are considered in the steps 'Parameter estimation' and 'Model discrimination' of the loop in Fig. 4, but the 'Nonlinear optimal experimental design' is only performed for the actually best rate model.

4.1 Parameter Estimation Problem

The objective of the parameter estimation procedure is to find the parameter values which set the numerical solution of the model equations (e.g. (1), (2), and (3) for the homogeneous model) as close as possible to the measured data. For this purpose, the problem (21), here stated for the homogeneous model, has to be solved. In this work, a weighted least square functional is used as objective function and it's final value is named LSQ. High values are related to a high lack of fit.

$$\min_{p_1,p_2,a,b,c} \mathsf{LSQ} = \sum_{j=1}^{NM} \frac{\left(T_{ns,j} - T_{md,j}\right)^2}{\sigma^2_{T_j,T_j}} \tag{21}$$

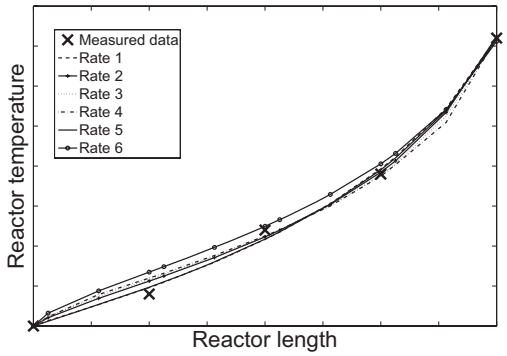


Fig. 5. Initial experiment.

In equation (21) the subscripts ns and md denote numerical solution and measured data, respectively, NM is the number of measured data points, and $\sigma^2_{T_j,T_j}$ is the standard deviation of the measured quantity (here the temperature T) at the measured point j. The calculated temperature profiles for the different rate laws after the parameter estimation are compared to the measured temperature data in Fig. 5. In Tab. 2, the LSQ values for the different rate models are given. All models are able to describe the measured data. This is not surprising, because 4 parameters (5 for rate 1) are fitted to only 5 measured points. Anyhow, the structure of the rate equations does not allow the same good fit for all rates. The best fit is reached with rate 5.
The parameter estimation problem becomes more difficult when more data points are available. The model equations have to be solved separately for each new experiment, making the parameter estimation the most expensive step regarding the computational effort.

Table 2. LSQ values for different rate models.

Rate	1	2	3	4	5	6
LSQ	0.488	0.688	0.821	0.916	0.461	1.901

4.2 Model Discrimination Problem

It is difficult to choose the best suitable rate model from Fig. 5. The discrimination step normally is performed taking the model with the lowest LSQ value. But, after only one experiment the models LSQ values are still close together, see Tab. 2. Thus, in a second step an experiment is designed that drift apart from the calculated temperature profiles for the estimated set of parameters. To do this, the optimization problem given in equation (22) is solved. By this means, the obtained inlet conditions are optimal for the discrimination step. Due to the fact that only one measured variable is considered, here a simpler formulation is chosen as proposed by Akaike (1974) and other authors.

$$\begin{aligned} &\min_{\dot{V},c_{SO_2,in},c_{O_2,in},T_{in}} \Phi_{\text{Disc.}} \text{ , with} \\ &\Phi_{\text{Disc.}} = \sum_{m=2}^{6} \sum_{i=m}^{6} \sum_{j=1}^{5} \left(T_{ns,Mod=m,j} - T_{ns,Mod=i,j}\right)^2 \end{aligned} \tag{22}$$

The temperature profiles calculated with the solution of (22) are plotted in Fig. 6. After the experiment is per-

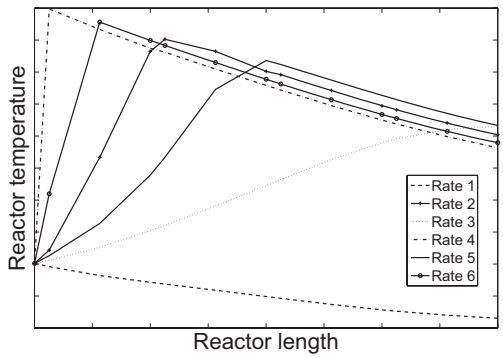


Fig. 6. Model discrimination experiment.

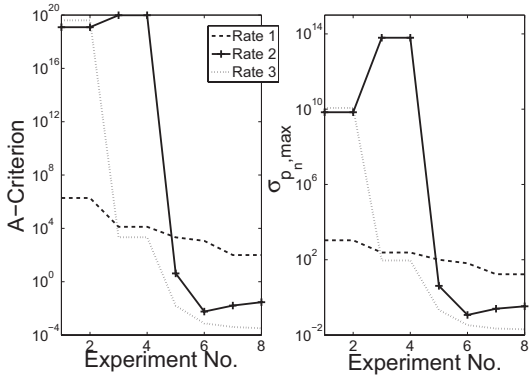


Fig. 7. Development of the parameter accuracy.

formed and the new parameters for each model are found, the profiles are not separated anymore. However, the experiment forces the rate models parameter values to move, and thus, the models flexibility and arbitrariness is reduced considerably.

Only one experiment is designed for a better model discrimination. After the discrimination experiment the rates 1, 2, and 3 are the most promising candidates. The other rates are still considered in the calculations but not plotted anymore in this paper. Further designs for discrimination are not performed because the focus on the parameter accuracy forces also an increasing difference in the models LSQ values. This means that after the convergence of the loop in Fig. 4 a good distinguishability between the rates is reached in addition.

4.3 Nonlinear Optimal Experimental Design Problem

In this work, the A-Criterion is used in order to increase the parameter accuracy, leading to the objective function given in equation (23). A detailed description of the nonlinear optimal experimental design and the different criteria so as a reason for the selection of the A-Criterion can be found in Schöneberger et al. (2008).

$$\min_{\dot{V},c_{SO_2,in},c_{O_2,in},T_{in}} \Phi_\mathrm{A}, \text{ with} \quad \Phi_\mathrm{A} = \frac{\mathrm{trace}\,(\mathbf{C})}{\dim\,(\mathbf{C})} \tag{23}$$

The covariance matrix of the model parameters $\mathbf{C}$ is approximated with the inverse of the Fisher information matrix $\mathbf{F}$ which can be calculated with equation (24), see Bard (1974). The rate model parameters are summarized in the parameter vector $\mathbf{P}$.

$$\mathbf{C} \geq \mathbf{F}^{-1} = \left[\sum_{j=1}^{NM} \left(\left(\frac{\partial T_{ns,j}}{\partial \mathbf{p}} \right) \frac{1}{\sigma^2_{T_j,T_j}} \left(\frac{\partial T_{ns,j}}{\partial \mathbf{p}} \right)^\mathbf{T} \right) \right]^{-1} \tag{24}$$

The parameter's standard deviations $\sigma_{p_n,p_n} = \sqrt{\mathbf{C}(n,n))}$ can be calculated with the diagonal elements of the covariance matrix. The framework is stopped, when a maximal standard deviation of $\sigma_{p_n,p_n,max} \leq 0.02$ is reached. This was accomplished after 8 experiments. The development of the A-Criterion and the maximal parameter standard deviation w.r.t. the experiments is shown in Fig. 7. The development of the LSQ values for the first three rate models

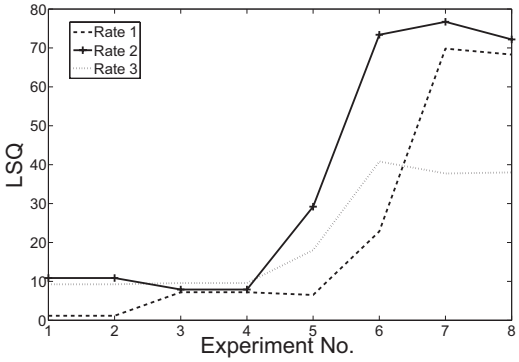


Fig. 8. Development of the lack of fit.

are depicted in Fig. 8. The experiments 1 and 2 and so 3 and 4 are repeated experiments used for the determination of the measured variables standard deviation σ_{T_j,T_j}.

5. RESULTS

In this section, the measured dynamic reactor behavior and the results obtained with the two-phase model are presented.

5.1 Catalyst Properties and Heat Transfer

The specific heat capacity of the catalyst phase $c_{P,k}$ is a model specific property, which has to be determined from experimental data. It should be noted that it is not equal to the heat capacity of the catalysts bulk material since the modeled catalyst phase contains also the particle pores. It can be estimated together with the parameter in the heat transfer correlation from experiments without reaction. This saves reactant gases and reduces the experimental effort because no off-gas treatment is necessary. The transient temperature profiles during the reactor heat up procedure can be used for this issue. They allow the independent estimation of the two parameters because they contain also the initial steady state, when the reactor inlet temperature is reached. These steady state profiles are independent of the catalyst phase heat capacity, but a function of the heat transfer coefficient. In Fig. 9 the numerical solution (surface) is fitted to the measured temperature data (black lines) based on the procedure described in section 4.1.

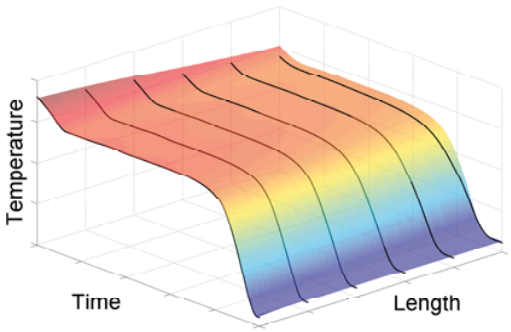


Fig. 9. Time-space surface without reaction.

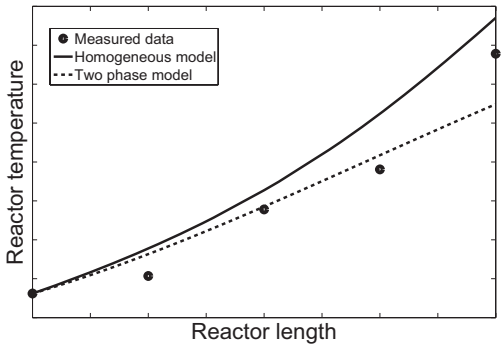


Fig. 10. Steady state comparison of the reactor models.

5.2 Parameter Transfer

Under the assumption of an instantaneous heat and mass transfer and a neglecting of axial dispersion, the model equations of the two-phase model can be rearranged to the form of the homogeneous model in the case of steady state. To show this, the profiles of a solution with the homogeneous model are compared to the steady state solution of the two-phase model in Fig. 10. The differences arise from the stated assumptions, which are not completely fulfilled. However, a steady state examination of a catalyst would not justify the use of the more complex two-phase model.

5.3 Dynamic Reactor Behavior

In Fig. 11, the transient temperature data is plotted. The temperatures after the first and the second bed show a strong overshooting. This behavior can be explained with the two-phase model. The coupled balances for gas and solid phase lead to a PT2 behavior when linearized. This second order element has the potential to produce a swinging solution. The effect is reduced in relation to the distance from the reactor inlet. The last two profiles show inflexion points instead. This can be explained with the increasing reactor wall temperature, which is also disturbing the system as well.

6. CONCLUSIONS

First calculation results show that the two-phase model is able to describe the dynamic reactor behavior. It has been demonstrated that for steady state experiments with a FBR the more complex two-phase model is not required. However, the information content of the transient profiles is much higher than the one from the steady state profiles. The effect on the parameter accuracy still has to be examined, in particular, because of the inclusion of the additional parameters. The knowledge of the dynamic reactor behavior enables the design of optimal transient experiments and their implementation in the proposed identification framework.

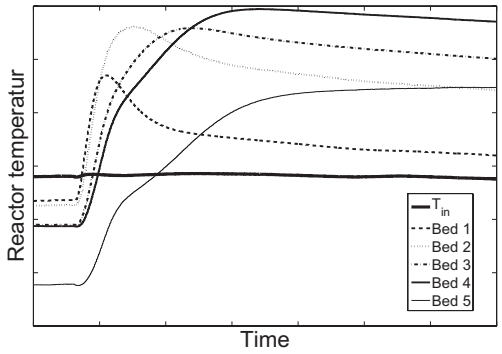


Fig. 11. Transient temperature data with reaction.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the Max-Buchner-Forschungsstiftung, and the provision of the catalyst from the BASF SE.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.

Arellano-Garcia, H., Schöneberger, J.C., and Körkel, S. (2007). Optimale Versuchsplanung in der chemischen Verfahrenstechnik. *Chem. Ing. Tech.*, 79(10), 1625.

Bard, Y. (1974). *Nonlinear Parameter Estimation*. Academic Press, London.

Fogler, H.S. (2006). *Elements of Chemical Reaction Engineering*. Perason International Edition, 4th edition.

Franceschini, G. and Macchietto, S. (2008). Model-based design of experiments for parameter precision: State of the art. *Chem. Eng. Sci.*, 63, 4846.

Mezaki, R. and Kadlec, B. (1972). Remarks on the reduction-oxidation mechanism of sulfur dioxode oxidation on vanadium catalyst. *J. Catal.*, 25, 454.

Schöneberger, J.C., Arellano-Garcia, H., Thielert, H., Körkel, S., and Wozny, G. (2007). An efficient approach to robust simulation of claus processes in coking plants. *Computer Aided Chem. Eng.*, 24, 521.

Schöneberger, J.C., Arellano-Garcia, H., Thielert, H., Körkel, S., and Wozny, G. (2008). Experimental analysis of a fixed bed reactor for catalytic SO2 oxidation. *Ind. Eng. Chem. Res.*, in press.

Schöneberger, J.C., Arellano-Garcia, H., and Wozny, G. (2009). An efficient discretization approach for partial differential equations describing chemical reaction systems. *Computer Aided Chem. Eng.*, in press.

Modeling and Simulation of the Polymeric Nanocapsule Formation Process

Luciane S. Ferreira and Jorge O. Trierweiler

GIMSCOP – Chemical Engineering Department – UFRGS. Rua Luis Englert, s/n. Porto Alegre. RS. Brasil.
(e-mail: {luciane,jorge}@enq.ufrgs.br)

Abstract: In this work the modelling and simulation of nanoparticle formation according to the technique of nanoprecipitation was done. In this method, the particle is formed due to the further diffusion of solvent into the water, resulting in the aggregation of the associated polymer chains. In order to predict the characteristics of the nanoparticle and also to improve the process, it was developed a mathematical model that considers: (a) the type of polymer; (b) interaction between solvent and polymer; and, (c) dynamics of solvent diffusion. The diffusivity between polymer-solvent was modelled by means of the Vrentas & Duda Free Volume Theory, including the Sanchez-Lacombe equation-of-state. The model was written in terms of Partial Differential Equation, and solved with MAPLE for a given initial size distribution. Additionally, it is a moving boundary problem because the diffusion of the solvent out of the droplet leads to its size reduction. Based on a given initial droplet size distribution, the transient behaviour and the final droplet size distribution can be evaluated. The dynamic simulation shows both the evolution of the solvent inside the droplet and the variation of size in time. Additionally, the comparison between experimental and simulated results showed a very good agreement.

Keywords: nanoprecipitation, modeling, simulation, diffusion, moving boundary, droplet size distribution.

1. INTRODUCTION

Polymeric nanoparticles are of especial interest from the pharmaceutical point of view. First, they are more stable in the gastrointestinal tract than other colloidal carriers and can protect encapsulated drugs from gastrointestinal environment. Second, the use of various polymeric materials enable the modulation of physicochemical characteristics (e.g. hydrophobicity, zeta potential), drug release properties, and biological behavior (e.g. targeting, bioadhesion, improved cellular uptake) of nanoparticles. Finally, their submicron size and large specific surface area favor their absorption compared to larger carriers (Des Rieux *et al.*, 2006). For instance, nanoparticles encapsulating proteins and vaccines (Des Rieux *et al.*, 2006) and chemotherapeutic agents (Jabr-Milane *et al.*, 2008) have been investigated in the last years.

One of the methods applied to produce nanoparticles is the so-called Nanoprecipitation. This method, first presented in 1989 (Fessi *et al.*, 1989), was largely applied by other authors in the subsequent years (Guterres *et al.*, 1995, Thioune *et al.*, 1997, Govender *et al.*, 1999, Chorny *et al.*, 2002, Galindo-Rodriguez *et al.*, 2004, Bilati *et al.*, 2005, Galindo-Rodriguez *et al.*, 2005). It consists of a simple procedure for the preparation of nanocapsules (NC) by interfacial deposition of a preformed, well-defined, and biodegradable polymer following displacement of a semi-polar solvent miscible with water from a lipophilic solution. The method of preparation yielded spherical vesicular nanocapsules, which consisted of an oily cavity – where the drug is dissolved - surrounded by a thin wall formed by interfacial deposition of the polymer. When organic and aqueous phases are in contact, it is assumed that solvent diffuses from the organic phase into the water and carries with it some polymer chains, which are still in solution. Then, as the solvent diffuses further into the water, the associated polymer chains aggregate forming NC. Therefore, this method involves the equilibrium among a polymer, its solvent and a non-solvent.

In order to predict the characteristics of the nanoparticle and also to improve the process, it was developed a mathematical model that takes in account: (a) the type of polymer; (b) interaction between solvent and polymer; and, (c) solvent diffusion process. After the description of the nanocapsules preparation, the mathematical model is explained. Then, the numerical simulation and its results are discussed. Finally, in the appendix the Free Volume model, applied to calculate the diffusivity polymer/solvent, is explained.

2. NANOCAPSULES PREPARATION

As mentioned in the previous section, the nanocapsules are prepared according to the method of nanoprecipitation. This method is based on the spontaneous emulsification of the organic internal phase, in which the polymer is dissolved, into the external aqueous phase. In this work nanocapsules of poly(ε-caprolactone) (PCL) containing 3-benzophenon (solar protection factor) were prepared according to the following procedure (Fessi *et al.*, 1989): 100 mg of PCL, 76.6 mg of sorbitan monostearate, 333 mg of Mygliol 810 (caprylic/capric triglyceride) and 30mg of Benzophenon-3 are first dissolved in acetone (27 ml). The resulting organic solution is poured in 53 ml of water containing 76.6 mg of polysorbate 80. The aqueous phase immediately turns milky with bluish opalescence as a result of the formation of nanocapsules, the wall of which is mainly constituted by PCL, and the oily core by the benzophenon-mygliol solution. The size of the nanoparticles is then analyzed by *Dynamic Light Scattering* (Zetasizer Nano, Malvern).

3. MATHEMATICAL MODELLING

It was considered that the nanoprecipitation produces perfect spherical particles and also that each nanoparticle is originated from one droplet formed immediately after the mixing of organic phase and aqueous phases. The major model assumptions are: (a) there is a negligible relative velocity between the droplet and the water; therefore, the external mass transfer is approximated by diffusion. This assumption can be done based on the order of the Stokes number, which is related to the particle velocity and is defined as (Crowe, 2005, Rielly and Marquis, 2001):

$$St \equiv \frac{\tau_v}{\tau_F} = \frac{d_p^{\,2}\rho_p V}{18\mu L} \quad [1] \tag{1}$$

For small Stokes numbers, the particles follow the fluid motion; but for large *St*, the particles follow different trajectories from the fluid elements. The Stokes number (Figure 1) was calculated for fluid velocities between $1x10^{-3}$ and 1 $m.s^{-1}$ (based on CFD simulations for a stirred tank that are not showed here), and particle diameters from 100 to 4000 nm. As the Stoke number is in all cases less than 10^{-5}, the assumption of negligible relative velocity between the droplets and the external phase can be considered valid.

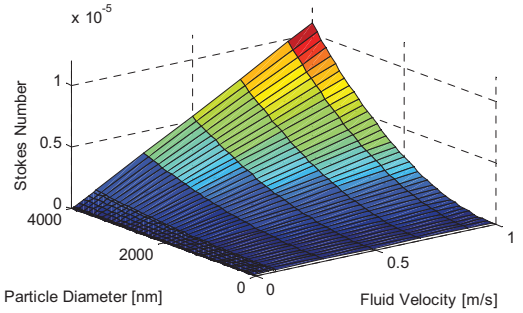


Fig. 1. Stokes Number.

(b) the diffusion is one-dimensional along the radial direction; and, (c) the diffusivity varies with time and concentration only.

Based on these assumptions, the mass balance equation for the solvent, written in spherical coordinates is:

$$\frac{\partial c_1(r,t)}{\partial t} = \frac{D}{r^2}\left(\frac{\partial}{\partial r}\left(r^2\left(\frac{\partial c_1(r,t)}{\partial r}\right)\right)\right) \quad [2] \tag{2}$$

This is a moving boundary problem, since the size of the droplet reduces because of the diffusion.

As the dimensions of dependent and independent variables are not the same, a variable normalization was be done including new variables τ, rh, and c_{1h} defined as,

$$t_0 = \frac{r_0^2}{D} \tag{3}$$

$$\left\{r = rh \cdot r_0,\ t = \tau \cdot t_0,\ c_1(r,t) = c_{1h}(rh,\tau)\cdot \rho_1\right\} \quad [3] \tag{4}$$

After the normalization, (2) is then rewritten as:

$$\frac{1}{t_0}\cdot\frac{\partial c_{1h}(rh,\tau)}{\partial \tau} = \frac{D}{r_0^{\,2}\cdot rh^2}\left(\frac{\partial}{\partial rh}\left(rh^2\left(\frac{\partial c_{1h}(rh,\tau)}{\partial rh}\right)\right)\right). \tag{5}$$

3.1 Initial and Boundary Conditions

It is assumed that the solution is well mixed and therefore, the concentration inside the droplet is uniform. Thus the initial condition is,

$$c_{1h}(rh,0) = c_{1h0} \qquad 0 \le rh \le r0\,. \tag{6}$$

Where c_{1h0} is evaluated according to the experimental conditions.

The boundary condition at the center of the droplet (rh=0) arises from the symmetry,

$$\left.\frac{dc_{1h}}{drh}\right|_{rh=0} = 0 \qquad t \ge 0\,. \tag{7}$$

Additionally, the boundary condition at the interface was calculated based on the mass balance and can be written as

$$\left.\frac{dc_{1h}}{dt}\right|_{rh=R} = \frac{D_{S-W}}{4\pi R^2}\left(c_{1h}(R) - c_{1h}(\infty)\right) \quad [4] \tag{8}$$

3.2 Boundary Movement

The boundary movement is calculated based on the assumption that both the mass of polymer, oil and drug remain constant during the diffusion process. The volume of the droplet is considered to be

$$V_D = V_2 + V_1 + V_{oil} + V_{drug} \equiv \frac{4}{3}\pi R^3 \quad [5] \tag{9}$$

Per definition the volume of polymer and solvent inside the droplet are

$$V_2 = m_2/\rho_2 \qquad V_1 = (c_1/\rho_1)V_D \quad [6] \tag{10}$$

Substituting (10) in (9) and isolating for *R*, then the radius can be calculated as

$$R(t) = \sqrt[3]{\left[3\cdot\left(\frac{m_2}{\rho_2} + V_{oil} + V_{drug}\right)\cdot\left(1 - \frac{c_1(t)}{\rho_1}\right)^{-1}\right] \Big/ 4\pi}\,. \tag{11}$$

3.3 Model Parameters

The two main parameters of this model are the diffusivity solvent/polymer and the diffusivity solvent/water.

The experimental data presented by Wild (2003) was adjusted as a polynomial curve to describe the diffusivity of acetone in water.

$$D_{S\text{-}W} = -4.737 w_S + 15.92\, w_S - 14.71\, w_S + 4.738 \tag{12}$$

where w_s is the molar fraction of water in the external phase.

The diffusivity between polymer and solvent was modelled according to the Free Volume Theory (Vrentas and Duda, 1976, Vrentas and Duda, 1977a, Vrentas and Duda, 1977b, Vrentas and Duda, 1979). Those authors applied the Flory–Huggins thermodynamic model in their free volume diffusion theory to describe the polymer solvent enthalpic and entropic

[1] τ_F - characteristic time of the flow field; τ_v - particle relaxation time; d_p - particle diameter; ρ_p - particle density; V - fluid velocity, μ - fluid viscosity; L - characteristic dimension of the obstacle.

[2] c_1 - concentration of solvent; r - particle radius; D - diffusivity; t - time.

[3] r_0 - initial droplet radius; D - diffusivity; ρ_1 - solvent density.

[4] R - actual radius of the droplet; $D_{S\text{-}W}$ - diffusivity of the solvent in the external phase; $c_{1h}(R)$ - concentration of solvent at the interface; $c_{1h}(\infty)$ - bulk concentration.

[5] V_D - volume of the droplet; V_1, V_2, V_{oil}, V_{drug} - volume of solvent, polymer, oil, and drug, respectively.

[6] m_2 - mass of polymer; ρ_2 - polymer density.

interactions. For the estimation of solvent diffusion coefficient in polymer solution systems, free-volume parameters for the both polymer and solvent must be available. The free volume (FV) diffusion model developed by Vrentas & Duda describes the solvent self-diffusion coefficient (D_1) and the polymer/solvent binary mutual diffusion coefficient (D) as given by (13) and (14), respectively.

$$D_1 = D_0 \exp\left(\frac{-E}{RT}\right)\cdot\exp\left(\frac{-\left(w_1\hat{V}_1^* + w_2\xi\hat{V}_2^*\right)}{w_1\left(\frac{K_{11}}{\gamma}\right)\left(K_{21}-T_{g1}+T\right)+w_2\left(\frac{K_{12}}{\gamma}\right)\left(K_{22}-T_{g2}+T\right)}\right) \quad (13)$$

$$D = D_1\left(1-\phi_1\right)^2\left(1-2\chi\phi_1\right) \quad (14)$$

In (13) the first exponential term can be considered as the energy factor, and the second exponential term is the free-volume factor. Eq. (14) contains the following implicit assumptions (Zielinski and Duda, 1992): (a) the mutual-diffusion coefficient is related theoretically to the solvent and polymer self-diffusion coefficients through an expression developed by Bearman (1961); (b) the contribution of the polymer self-diffusion coefficient to the mutual-diffusion is negligible; and, (c) the Flory-Huggins (Flory, 1970) model accurately describes the polymer activity. In addition, the specific free volumes of the polymer and solvent are presumed to be additive (without a volume change on mixing), and thermal expansion coefficients are approximated by average values over the temperature intervals of interest (Frick *et al.*, 1990, Lodge *et al.*, 1990). There are 13 independent parameters to be evaluated in (14). Some of them can be grouped reducing this number to the following variables: K_{11}/γ, $K_{21}-T_{g1}$, K_{12}/γ, $K_{22}-T_{g2}$, $\hat{V}_1^*$, $\hat{V}_2^*$, D_0, E, , ξ, and χ, that must be determined to estimate mutual diffusivities. All of these parameters have physical significance, and therefore one must be able to evaluate every parameter from sources other than diffusion studies. The guidelines to calculate them, clarified by Zielinski and Duda (1992), were used in this work. Additionally the modification proposed by Wang *et al.* (2007), according to the Sanchez-Lacombe equation-of-state (SL EOS), was also taken in account and is explained in Appendix A. The process was considered to be isotermic and isobaric (T=298K, P=1bar and E=0). All the parameters are listed in Table 1 and more details can be found in Appendix A.

Table 1: Model Parameters.

$\hat{V}_1^*$	0.9695 cm^3/mol	(K_{11}/γ)	0.983 x 10^{-3}
$\hat{V}_2^*$	0.8181 cm^3/mol	ρ^*	1.1427 g/cm^3
δ_1	18.29 J$^{1/2}$/cm$^{3/2}$	T^*	668 K
δ_2	20.85 J$^{1/2}$/cm$^{3/2}$	P^*	4035 bar
D_0 x 10^4	14.3 cm^2/s	T_{g2}	213 K
K_{21}-T_{g1}	-12.12		

The numerical values of number- and volume-average (size and standard deviation) – measured with ZetaSizer Nano® from samples all prepared according to the same methodology – are described in Table 2.

4. NUMERICAL SIMULATION

Experimentally the mean diameter is measured by *Dynamic Light Scattering* (ZetaSizer Nano® – Malvern). Typical number and volume density distributions, measured, can be seen in Fig. 2.

According to Table 2 and Fig. 2, it is observed that: (a) The number-average size is always smaller than the volume-average, as the contribution of a spherical particle grows proportionally with D^3; (b) The standard deviation is about 32 – 41% of the averaged value; and, (c) The size distribution does not follow a normal distribution.

Table 2: Number and Volume average sizes.

	Number-average particle diameter	Volume-average particle diameter
Sample 1	222.12 ± 71.25	303.76 ± 103.68
Sample 2	199.23 ± 67.83	285.13 ± 103.97
Sample 3	200.81 ± 73.60	300.90 ± 117.73
Sample 4	200.57 ± 72.48	303.66 ± 126.88

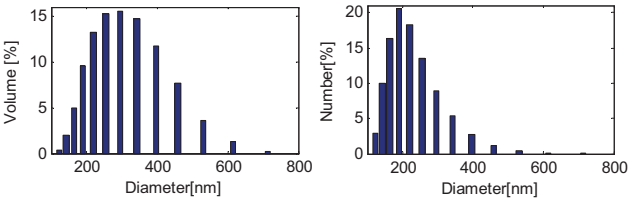


Fig. 2: Measured Number and Volume density distribution.

If it is assumed that the number of droplets/particles remains constant during all the process, the initial distribution can be calculated through the mass balance and based on the experimentally volume density distribution at final time, as explained by Fig. 3.

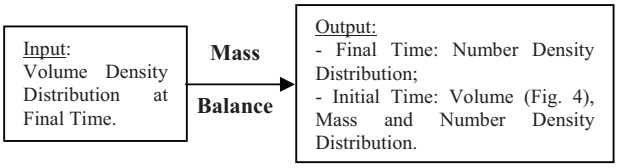


Fig. 3: Measured number and volume size distribution.

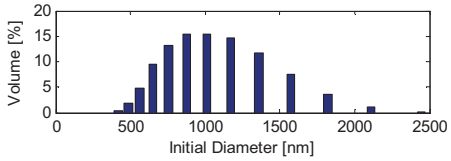


Fig. 4: Calculated Initial Size Distribution.

The calculated number density at final time was compared to the measured one, as can be seen in Fig. 5. The good agreement between both assures that the methodology used in this work is correctly applied.

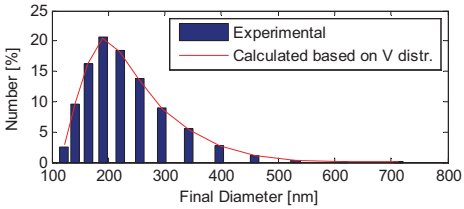


Fig. 5: Comparison between the Experimental and Calculated Number Density Size Distribution.

As the initial size distribution is now available, the diffusion

model can be simulated. It was done in Maple, which is capable of finding solutions for higher order PDE or PDE systems.

Based on the initial size distribution, (5) is solved for each class of particle size. After that, the new radius can be thus calculated. If the relative difference (*ΔDiam*) between the new and old radius is less than $1x10^{-5}$, the process ends; if not, the iteration process goes on as illustrated in Fig. 6.

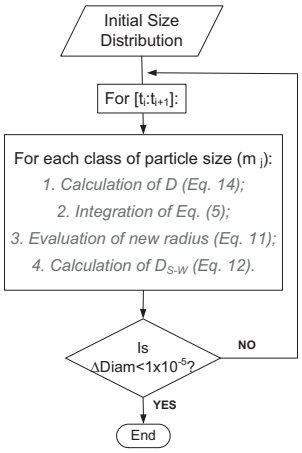


Fig. 6: Scheme of solution implemented in MAPLE.

5. RESULTS AND DISCUSSION

The model predicts that in about 12 ms all the particles reach the final diameter. This result is in qualitative agreement with what is observed experimentally, that is, as the organic phase is mixed in the aqueous phase, the suspension becomes *immediately* (at least for the human eyes) opaque, as a result of the nanoparticle formation.

Throughout the diffusion process, gradients of oil, acetone and polymer arise into the droplet leading to the reduction in size and the formation of the nanoparticle, as can be seen in Fig. 7. Each line represents one class of particle that forms the distribution.

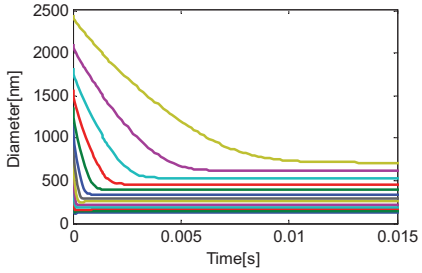


Fig. 7: Evolution of Particle Size in time.

For the smallest particles, it is observed that the diameter reduction is very fast, which is a consequence of the fast diffusion of acetone to the external medium, as can be seen in Fig. 8.

The order of magnitude of the diffusion time is in agreement with that found by Moinard-Chécot *et al.* (2008). They tried to measure the duration of the solvent diffusion step with a stopped-flow apparatus. In this experiment, only the signal corresponding to the final state could be observed, i.e., the diffusion step is less than 20 ms (the acquisition time of the apparatus).

Meanwhile the biggest particles presented a slower diffusion profile. Because of their initial big size, it takes longer for the solvent to reach the interface and consequently, to be transferred to the external medium. As more solvent diffuses out of the droplet, the concentration of polymer inside the droplet increases and consequently, the diffusivity polymer/solvent also grows up. At intermediate steps, the slow reduction in size showed by the largest particles leads almost to a bimodal volume and mass density distribution, as can be seen in Fig. 9.

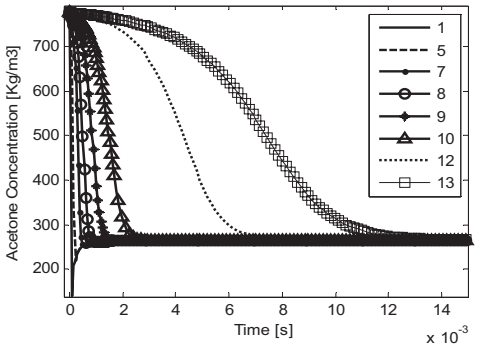


Fig. 8: Acetone concentration at the interface.

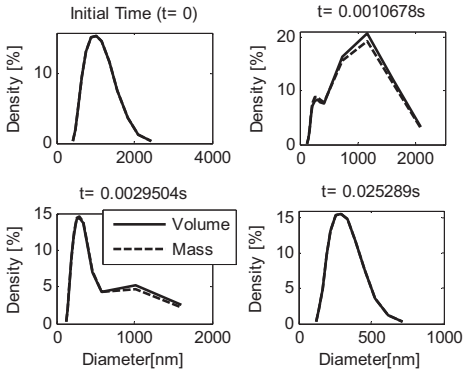


Fig. 9: Volume and Mass Distribution at the initial and final time, and 2 instants of time in-between.

The comparison between the experimental and calculated final volume density distribution (Fig. 10) shows a very good agreement, confirming that the assumptions done in sections 2 and 3 are suitable for modelling the system.

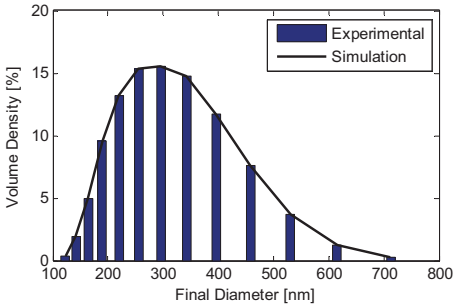


Fig. 10: Comparison between the experimental and simulated volume density distribution.

In order to generalize the model, a probability density

distribution can be employed to describe the initial size of the droplets. Applying the gamma distribution it is possible then to compare the experimental and simulation results, as show in Fig. 11. Analyzing both results and considering the standard deviation of the measurement, it is possible to confirm that the model and the methodology presented in this work are suitable to simulate the nanoprecipitation.

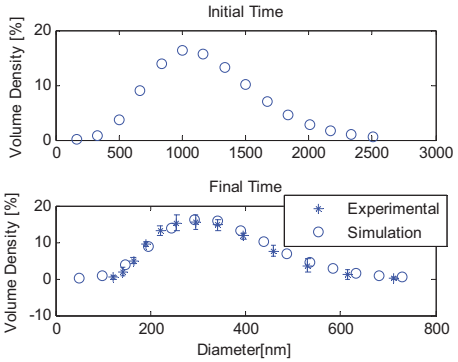


Fig. 11: Results obtained when using a gamma distribution to describe the droplets initial size. (a) Initial size distribution, (b) Final size distribution.

6. CONCLUSIONS

This work shows that it is possible to obtain satisfactory results using a simplified PDE model for the nanoprecipitation. The adopted approach considers several variables that have influence on diffusion, like type of polymer, solvent and non-solvent; affinity among them; and, polymer solubility. As a result, it is then possible to evaluate the particle size distribution during the nanoprecipitation. The comparison between the simulated and measured sizes showed a quite good agreement, suggesting that the employed methodology can describe correctly the nanoprecipitation.

The proposed model can be combined with CFD simulator in order to improve the predictions. From the simulation, one could try to evaluate more accurately the initial droplet distribution. This methodology could thus be applied to study the influence of several kinds and sizes of reactors and mixers in the final properties of the nanoparticles.

Acknowledgments: Prof. Wolfgang Marquardt (PT.AVT – RWTH Aachen) and CNPq/DAAD.

7. REFERENCES

Bearman, R. J. (1961) On the Molecular Basis of some Current Theories of Diffusion. *J. Phys. Chem.,* 65**,** 1961-1968.

Bilati, U., Allemann, E. and Doelker, E. (2005) Development of a nanoprecipitation method intended for the entrapment of hydrophilic drugs into nanoparticles. *European Journal of Pharmaceutical Sciences,* 24**,** 67-75.

Chorny, M., Fishbein, I., Danenberg, H. D. and Golomb, G. (2002) Lipophilic drug loaded nanospheres prepared by nanoprecipitation: effect of formulation variables on size, drug recovery and release kinetics. *Journal of Controlled Release,* 83**,** 389-400.

Crowe, C. T. (2005) *Multiphase Flow Handbook*, CRC Taylor & Francis.

Des Rieux, A., Fievez, V., Garinot, M., Schneider, Y.-J. and Préat, V. (2006) Nanoparticles as potential oral delivery systems of proteins and vaccines: A mechanistic approach. *Journal of Controlled Release,* 116**,** 1-27.

Fessi, H., Puisieux, F., Devissaguet, J. P., Ammoury, N. and Benita, S. (1989) Nanocapsule formation by interfacial polymer deposition following solvent displacement. *International Journal of Pharmaceutics,* 55**,** R1-R4.

Flory, P. J. (1970) Thermodynamics of Polymer Solutions. *Disc. Faraday Soc.,* 49**,** 7 - 29.

Frick, T. S., Huang, W. J., Tirrell, M. and Lodge, T. P. (1990) Probe diffusion in polystyrene/toluene solutions. *Journal of Polymer Science Part B: Polymer Physics,* 28**,** 2629-2649.

Galindo-Rodriguez, S., Allémann, E., Fessi, H. and Doelker, E. (2004) Physicochemical Parameters Associated with Nanoparticle Formation in the Salting-Out, Emulsification-Diffusion, and Nanoprecipitation Methods. *Pharmaceutical Research,* 21**,** 1428-1439.

Galindo-Rodriguez, S. A., Puel, F., Briancon, S., Allemann, E., Doelker, E. and Fessi, H. (2005) Comparative scale-up of three methods for producing ibuprofen-loaded nanoparticles. *European Journal of Pharmaceutical Sciences,* 25**,** 357-367.

Govender, T., Stolnik, S., Garnett, M. C., Illum, L. and Davis, S. S. (1999) PLGA nanoparticles prepared by nanoprecipitation: drug loading and release studies of a water soluble drug. *Journal of Controlled Release,* 57**,** 171-185.

Guterres, S. S., Fessi, H., Barratt, G., Devissaguet, J. P. and Puisieux, F. (1995) Poly (DL-lactide) nanocapsules containing diclofenac: I. Formulation and stability study. *International Journal of Pharmaceutics,* 113**,** 57-63.

Hansen, C. M. (2000) *Hansen solubility parameters: a user's handbook*, CRC Press

Jabr-Milane, L., Van Vlerken, L., Devalapally, H., Shenoy, D., Komareddy, S., Bhavsar, M. and Amiji, M. (2008) Multi-functional nanocarriers for targeted delivery of drugs and genes. *Journal of Controlled Release,* 130**,** 121-128.

Lodge, T. P., Lee, J. A. and Frick, T. S. (1990) Probe diffusion in poly (vinyl acetate) / toluene solutions. *Journal of Polymer Science Part B: Polymer Physics,* 28**,** 2607-2627.

Moinard-Chécot, D., Chevalier, Y., Briançon, S., Beney, L. and Fessi, H. (2008) Mechanism of nanocapsules formation by the emulsion-diffusion process. *Journal of Colloid and Interface Science,* 317**,** 458-468.

Rielly, C. D. and Marquis, A. J. (2001) A particle's eye view of crystallizer fluid mechanics. *Chemical Engineering Science,* 56**,** 2475-2493.

Rodgers, P. A. (1993) Pressure-volume-temperature relationships for polymeric liquids: A review of equations of state and their characteristic parameters for 56 polymers. *Journal of Applied Polymer Science,* 48**,** 1061-1080.

Sugden, S. (1927) Molecular Volumes at Absolute Zero: II. Zero Volumes and Chemical Composition. *Journal of the Chemical Society***,** 1786 - 1798.

Thioune, O., Fessi, H., Devissaguet, J. P. and Puisieux, F. (1997) Preparation of pseudolatex by nanoprecipitation: Influence of the solvent nature on intrinsic viscosity and interaction constant. *International Journal of Pharmaceutics,* 146**,** 233-238.

Van Dijk, M. A. and Wakker, A. (1997) *Polymer Thermodynamics Library - Volume 2: Concepts of Polymer Thermodynamics*.

Van Krevelen, D. W. (1990) *Properties of Polymers*, Elsevier.

Vrentas, J. S. and Duda, J. L. (1976) Diffusion of Small Molecules in Amorphous Polymers. *Macromolecules,* 9**,** 785-790.

Vrentas, J. S. and Duda, J. L. (1977a) Diffusion in Polymer-Solvent Systems. I. Reexamination of the Free-Volume Theory. *J. Polymer Sci., Polymer Phys.,* 15**,** 403 - 416.

Vrentas, J. S. and Duda, J. L. (1977b) Diffusion in Polymer-Solvent Systems. II. A Predictive Theory for the Dependence of Diffusion Coefficients on Temperature, Concentration, and Molecular Weight. *J. Polymer Sci., Polymer Phys.,* 15**,** 417 - 439.

Vrentas, J. S. and Duda, J. L. (1979) Molecular Diffusion in Polymer Solutions. *AIChE J.,* 25**,** 1 - 24.

Wang, B.-G., Lv, H.-L. and Yang, J.-C. (2007) Estimation of solvent diffusion coefficient in amorphous polymers using the Sanchez-Lacombe equation-of-state. *Chemical Engineering Science,* 62**,** 775-782.

Wild, A. (2003) Multicomponent Diffusion in Liquids. TU München.

Zielinski, J. M. and Duda, J. L. (1992) Predicting polymer/solvent diffusion coefficients using free-volume theory. *AIChE Journal,* 38**,** 405-415.

Appendix A. FREE VOLUME THEORY PARAMETERS

$\hat{V}_1^*$ *and* $\hat{V}_2^*$ – The two critical volumes were approximated as the specific volumes of solvent and polymer at absolute zero temperature. The molar volumes at 0K were estimated using a group contribution method (Sugden, 1927).

χ – The interaction parameter in terms of solubility parameters can be calculated through the following equation (Van Dijk and Wakker, 1997):

$$\chi = \frac{V_m}{RT}(\delta_1 - \delta_2)^2 \tag{15}$$

where V_m is the molar volume, R is the ideal gas constant and δ_1 and δ_2 are the solvent and polymer solubility parameters respectively.

A widely used solubility parameter approach for predicting polymer solubility is proposed by Hansen (2000). The basis of the so-called Hansen Solubility Parameters (HSP) is that the total energy of vaporization of a liquid consists of several individual parts. These arise from (atomic) dispersion forces (E_D), (molecular) permanent dipole-permanent dipole forces (E_P) and (molecular) hydrogen bonding (electron exchange) (E_H). The basic equation which governs the assignment of Hansen parameters is that the total cohesion energy, E, must be the sum of the individual energies which make it up

$$E = E_D + E_P + E_H \tag{16}$$

Dividing each one by the molar volume gives the square of the total solubility parameter.

$$\begin{aligned} E/V &= E_D/V + E_P/V + E_H/V \\ \delta^2 &= \delta_D^2 + \delta_P^2 + \delta_H^2 \end{aligned} \tag{17}$$

The solubility parameter components were predicted from group contribution method, using the following equations (Van Krevelen, 1990).

$$\delta_d = \sum F_{di} \Big/ V \qquad \delta_p = \sqrt{\sum F_{pi}^2} \Big/ V \qquad \delta_h = \sqrt{\sum E_{hi} \Big/ V} \tag{18}$$

D_0, K_{21}-T_{g1}, (K_{11}/γ) – These are solvent parameters (acetone) and the values previously collected by Zielinski and Duda (1992) were used here.

(K_{12}/γ)(K_{22}-T_{g2}+T)(Wang et al., 2007) – According to the Vrentas-Duda model (Zielinski and Duda, 1992), the hole free volume of a polymer $\hat{V}_{FH2}$ in its rubbery state can be expressed as

$$\frac{\hat{V}_{FH2}}{\gamma} = \left(K_{12}/\gamma\right)\left(K_{22} - T_{g2} + T\right) \tag{19}$$

Assuming that the hole free volume is equal to the volume defined by the WLF (Williams, Landel and Ferry) theory, at the atmospheric pressure, (19) becomes

$$\frac{\hat{V}_{FH2}}{\gamma} = V\left\{0.025 + \int_{T_{g2}}^{T}\left[\left(1+\frac{\tilde{P}}{\tilde{\rho}^2}\right)\Big/\left(T\left[\frac{\tilde{T}}{1-\tilde{\rho}} - 2\right]\right)\right]_{P0} dT\right\} \tag{20}$$

where $V(=1/\rho)$ is the volume of polymer per gram at temperature T, which can be estimated by the SL EOS according to the equation,

$$\tilde{\rho} = 1 - \exp\left[-\left(\tilde{\rho}^2 + \tilde{P}\right)/\tilde{T} - \tilde{\rho}\right] \tag{21}$$

with the definition

$$\tilde{\rho} = \rho/\rho^*, \qquad \tilde{T} = T/T^*, \qquad \tilde{P} = P/P^* \tag{22}$$

where ρ^*, T^* and P^* are characteristics parameters of mass density, temperature and pressure, respectively. These parameters are listed by Rodgers (1993) and those for PCL used in this work are listed in Table 1. This method eliminates the need to use polymer viscoelastic data for determining the polymer free volume parameters. Additionally, its scope is also extended to include not only temperature and concentration but also pressure influence on solvent diffusivities. The only new parameters introduced by this model (ρ^*, T^* and P^*) are three parameters of the SL EOS (Wang et al., 2007).

Predictive Modeling of Key Process Variables in Granulation Processes based on Dynamic Partial Least Squares

D. Ronen*, C.F.W. Sanders*, H.S. Tan, P. R. Mort***, F.J. Doyle III***

**Department of Chem. Eng., UCSB, Santa Barbara, CA 93106-5080 USA (e-mail: frank.doyle@isb.ucsb.edu)*
*** P & G Technical Centre, Ltd. Whitley road, Longbenton, Newcastle Upon Tyne, NE12 9TS UK (e-mail: tan.h.9@pg.com)*
****Procter & Gamble Co. 5299 Spring Grove Ave, Cincinnati, OH 45217 USA (e-mail:mort.pr@pg.com)*

Abstract: Granulation is a multivariable process characterized by several physical attributes that are essential for product performance, such as granule size and size distribution. An optimally operated granulation process will yield, in a reproducible manner, product with tightly controlled performance attributes. In this paper predictive models of the dynamics of these key variables are developed using a dynamic partial least squares approach. The method, demonstrated here on process simulation as well as on an industrial mixer-granulator process, result in accurate predictions. These models motivate the development of model predictive controllers for these processes.

Keywords: Granulation, Process control, Dynamic modeling, Partial least squares

1. INTRODUCTION

Granulation is a complex process in which many input variables influence many product properties. As Iveson et al. describe in a review paper (2001), the understanding of the fundamental processes that control granulation behavior and product properties have increased in recent years. This knowledge can be used during process design, in choosing the right formulation and operating conditions, and it can also be used to improve process control. Although many variables are set constant during process design, variations during production in input variables occur due to the variable nature of the powder feed. Even if all granule properties, except for size, are ignored for process control, a one dimensional granule size distribution can be constructed by multiple discrete output variables, in order to represent the shape of the distribution (these can be mean sizes (with coefficients of variation), percentile sizes, moments or size bins). Model Predictive Control (MPC) is an effective method to control such multiple input, multiple output processes (García, et al., 1989). The majority of MPC applications in the chemical process industries utilize empirical models that are constructed from plant data. In this work, we explore the use of dynamic partial least squares to construct these empirical models.

2. METHODS

2.1 Partial Least Squares

Partial Least Squares (PLS) methods have been demonstrated as a useful tool for analysis of data and modeling of the systems from which the data are collected (Kaspar and Ray, 1993). Unlike related methods, such as Principal Component Analysis (PCA), which finds factors that capture the greatest amount of variance in the predictor (X) only, the PLS method attempts to find factors which both capture variance and achieve correlation. PLS handles this by projecting the information in high dimensional spaces (X,Y) down to low dimensional spaces defined by a small number of latent vectors (t_1,t_2…t_a). These new latent vectors summarize all the important information contained in the original data sets, by representing the scaled and mean-centered values of X and Y matrices as:

$$X = \sum_{i=1}^{a} t_i p_i^T + E$$

$$Y = \sum_{i=1}^{a} u_i q_i^T + F \qquad (1)$$

where the t_i are latent (score) vectors calculated sequentially for each dimension i=1,2,…a.

In the PLS method, the covariance between the linear combinations of X and the output measurement matrix Y is maximized at each iteration, using the vectors p_i and q_i which are the loading vectors whose elements express the contribution of each variable in X and Y toward defining the new latent vectors t_i and u_i. E and F are residual matrices for X and Y blocks, respectively. The optimal number of latent vectors retained in the model is often determined by cross-validation (Dayal et al. 1994).

In an industrial environment, it is more often the case that many of the predictor variables (X) are highly correlated with one another and their covariance matrix is nearly singular,

which renders classical regression methods intractable. Reduced space methods such as PLS and PCA can overcome this problem (MacGregor and Kourti, 1995). PLS is also robust to measurement noise in the data and can be used in cases where there are random missing data and when the number of input variables is greater than the number of observations (Dayal et al. 1994). Various examples of the implementation of PLS analysis to industrial process modeling and control can be found in the literature (for example, Dayal et al., 1994, MacGregor and Kourti, 1995, and others).

Process dynamics can be incorporated into the PLS model by including columns of lagged outputs and/or inputs into the predictor block (*X*) (Dayal et al., 1994, Kaspar and Ray, 1993, Juricek et al. 2001). The resulting PLS model is actually an ARX type input-output model of the form:

$$A(q^{-1})y_i(k) = \sum_{j=1}^{nu} B_j(q^{-1})u_j(k - nk_j) \tag{2}$$

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} + \ldots + a_m q^{-m} \tag{3}$$

$$B_j(q) = b_{j,1} q^{-1} + b_{j,2} q^{-2} + \ldots + b_{j,m} q^{-m} \tag{4}$$

where *y* denotes the output variable (e.g., median particle size, d_{50}), and *u* denotes the manipulated variable (e.g., binder flow). The terms *A* and *B* contain the autoregressive and exogenous terms of the model, respectively. The autoregressive term captures dynamics through lagged terms of the output, and the exogenous term captures dynamics through lagged terms in the input.

Once the models have been calculated from the plant data, it is useful to evaluate their properties using several key statistical measures. Some of the useful statistics that are associated with reduced space models (Wise et al. 2006) are outlined below:

Q residual – is simply the sum of squares of each row of *E* (from eq. 1), i.e. for the i[th] sample in *X*, x_i:

$$Q_i = e_i e_i^T \tag{5}$$

where e_i is the i[th] row of *E*. The Q statistics is a measure of the difference between a sample and its projection into the *a* principal components retained in the model.

Hotelling T^2 is a measure of the variation in each sample within the model. Its value is the sum of normalized squared scores, defined by:

$$T_i^2 = t_i \lambda^{-1} t_i^T \tag{6}$$

where t_i are the score vectors (eq. 1) and λ is a diagonal matrix containing the eigenvalues corresponding to *a* eigenvectors (principal components) retained in the model.

Together, the T^2 and Q residual statistics are useful in evaluating the fitness of a PLS model to specific data. It is possible to calculate statistically meaningful confidence limits for both cases.

2.2 Simulation studies

In our previous work, a nonlinear one dimensional population balance model (1D-PBM) was successfully used to model a laboratory continuous drum granulation process with fine particle recycle (Glaser et al., 2008). The same model is used here as a base for a process simulation (Figure 1) for a preliminary evaluation and sensitivity test of the applicability of the dynamic PLS modeling technique for granulation.

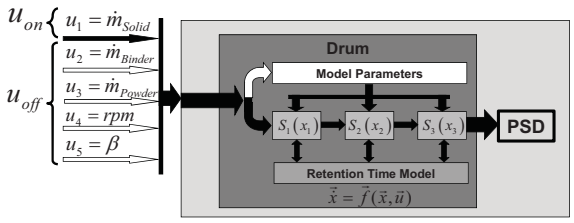


Fig. 1. Simulator structure: five inputs are included in the simulator: binder spray rate, fine powder feed-rate, drum rotation-rate and the drum inclination angle. The model is divided into three well mixed drum compartments, each described by an individual set of ODEs, a retention time model and a set of global parameters that influence the model behavior (taken from Glaser 2008).

Both particle median size (d_{50}) and, separately, particle size distribution width (d_{84}/d_{16}) were used as output variables for this study. The predictor (*X*) was constructed from 4 manipulated variables (solid feed flow rate, binder feed flow rate, drum rotation speed, recycle rate) and the computed recycle flow as an additional input variable. Process dynamics were incorporated into the *X* block by including columns of lagged output variables. The lag time was estimated using an autocorrelation function. Delay times of each of the input variables were estimated using cross correlation function, and the predictor matrix was adjusted according to the obtained delay vector. During the simulation, the 4 manipulated variables were randomly perturbed around their nominal values at steady state sequentially, i.e. input variables were perturbed one after the other in fixed time gaps. The resulting PLS-based ARX model's short horizon predictive ability was tested by cross validation with a set of separately calculated simulation sequences with different excitation regimes. For each of these cross validation sequences, the root mean square error of the model based prediction (RMSEP), relative to the simulated plant measurements was calculated for a given short horizon period. In order to make a more representative quantification of the predicting ability of the model, the short horizon start point was moved along the time axis of the data one time step after another thus creating a set of RMSEP measures out of which an average and maximum RMSEP could be calculated. All variables were mean centered and scaled to unit variance prior to processing.

Based on this technique a sensitivity test was performed in order to estimate the required size of the data set needed for reliable process modeling. Figure 2 shows the convergence of RMSEP related to the length of data set used for the PLS model training. This plot is based on averaging 100 multiple simulations and modeling runs for each training length. Sample rate was set to 2 minutes and the simulated process step response time (τ) was set to 4.5 minutes. The prediction horizon was set to 8 samples (i.e., 16 minutes). From this figure one can note that most of the dynamic features are captured by the model in the first 200 minutes of training data, as the mean RMSEP converges to low values. However using training data of up to 600 minutes would improve the model predictions. Notice that these results are not so sensitive to the excitations rate used in the modeling data set (i.e., time between two successive input variable perturbations), as long as this time is in the order of magnitude of the expected variations in process variables. Figure 3 depicts the response of the process model obtained with 520 minutes of training data to the process simulation for a step response in one of the input variables.

All of these models use one lagged output variable (granules median size) in the predictor block and 2 latent variables in the PLS model, which are linear combinations of the time-lagged values of the output variable and the delayed values of the five process variables. Figure 4 shows the prediction abilities of three PLS models obtained using different lengths of data sets for training (120, 220 and 520 minutes long) from a single simulation data, with input variable excitation every 15 minutes. In this example, the predictions of these models are cross-validated using data from a separate simulation run with randomly timed excitations of the inputs. Considering that PLS models captures covariance in X and Y, it is possible to calculated the percentage of variance captured by each of their latent variables by dividing the variance predicted by the latent variables to the total variance in the original data. The percentage of variance captured by the abovementioned 3 PLS models (from the training data) is detailed in Table 1. It is noticeable that the longer the training set used, more fine details of the process dynamics are captured by the models, confirming Figure 2 results. Notice that high values of explained variance do not guarantee good prediction of validation data by these models. The robustness of the PLS based models to measurement noise is demonstrated in Figure 5 and Table 2, where the simulated process was subject to 5% white noise on the output and input variable measurements.

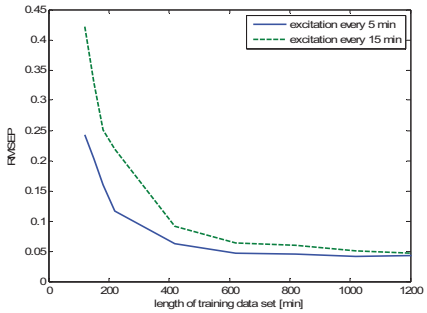


Fig. 2. Root Mean Square Error of Prediction (in validation simulation) versus length of training data set (based on modeling simulation), at different excitation rates.

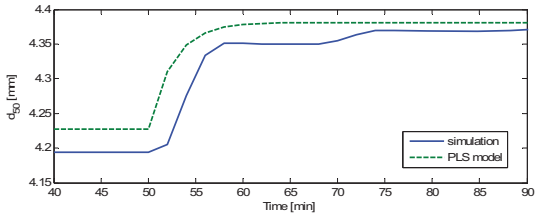


Fig. 3. Step response to a 1.2% step change in Binder feed flow - PLS model based on 520 minutes training data vs. process simulation.

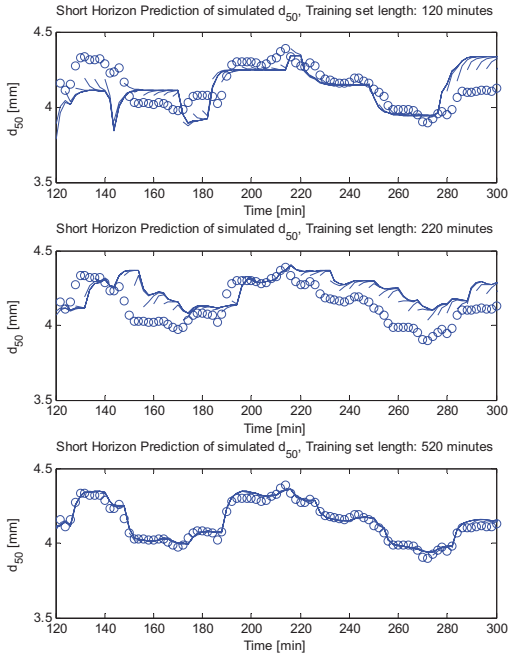


Fig. 4. PLS based dynamic model validation for different training set length. Circles represents simulation results, lines represent 8 point horizon prediction.

Table 1: Percent variance captured by PLS models based on different lengths of training data sets.

Training set length	LVs	X Block		Y Block	
		This	Total	This	Total
120min	1	76.04	76.04	82.44	82.44
	2	7.66	83.7	11.42	93.85
220min	1	59.45	59.45	93.05	93.05
	2	14.21	73.66	2.8	95.85
520min	1	41.35	41.35	97.88	97.88
	2	17.29	58.64	1.04	98.92

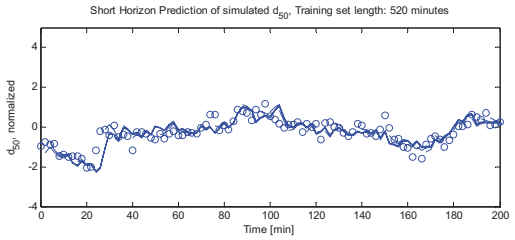


Fig. 5. PLS based dynamic model validation, simulated process with 5% white measurement noise. Circles represents simulation results, lines represent 8 point horizon prediction.

Table 2: Percent variance captured by PLS model, simulated process with 5% white measurements noise.

Training set length	LVs	X Block		Y Block	
		This	Total	This	Total
520 min	1	33.93	33.93	82.75	82.75
	2	33.10	67.03	1.61	84.37

3. DYNAMIC PLS MODELLING OF AN INDUSTRIAL PROCESS PLANT

3.1 Process plant description

In this section, we report on some preliminary studies on granulation model identification for a Procter & Gamble (P&G) industrial granulation process using normalized process data.

A complex industrial granulation process, such as the flowsheet shown in Fig. 6, was subjected to a series of (designed) random perturbations in a number of input parameters. This plant is equipped with an on-line granule size measurement system that measures particle size based on image analysis of 2-D camera images. The analysis constructed size distributions on the basis of the measured cross sectional area of the 2-D images. Granule size data along with all other plant variables were then sent to the UCSB team for modeling.

There are notable distinctions between the P&G study and the one reported in Section 2.2. The P&G process study is not meant to be used as a direct comparison (or validation) for the process simulation studies in Section 2.2. Rather, we present both as separate case studies to demonstrate the feasibility of using PLS methodology as an empirical modeling tool for granulation process control.

The low-shear drum-granulation pilot plant that was used to design the simulator in Section 2.2 produced particles with d_{50}'s of several mm; on the other hand, the medium-high shear mixer-granulation process shown in Fig 6 typically produced particles with d_{50}'s less than 1 mm. While the underlying physical mechanisms of growth and consolidation may be similar, the flow and shear fields are very different for the two processes (the drum granulator is relatively low shear, compared to the medium-high shear mixer-granulation process), the process layouts and control handles are different, as are the material properties. As such, the choices of process variables (manipulated and measured) are unique for each process.

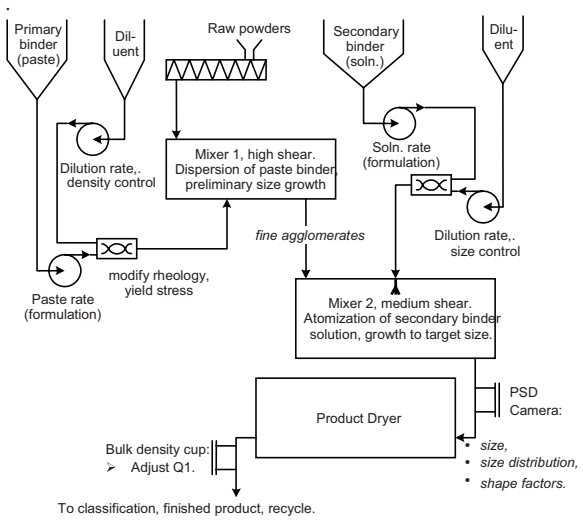


Figure 6. Representative P&G process flow diagram for mixer-granulator (Mort et al. 2001). For simplicity, this diagram omits the usual operations for classification and recycle.

3.2 PLS modeling of real plant data- Case I

Figure 7 describe the dynamic PLS model fitting obtained for the granules median size. The data set obtained from the plant originally contained 147 sampling points, each consist of 81 process variables, together with granules size measurements. Sampling time was 0.4 times the process characteristic time τ. During this time period 4 manipulating variables were subjected to random perturbations around their nominal values at steady states (Fig. 8) in a similar way to the simulation work described earlier (adjusted to the process τ), while other adjustments were continuously made to other plant variables (i.e. normal plant operations). Granules median size (d_{50}) was selected as the output variable. Nine out of the 81 process variables were chosen as predictor variables for the PLS model, based on engineering judgment, GA based variable selection (PLS Toolbox 5.0 by Eigenvector research incorporated), and trial and error. The

output lag time and process variables delay times were evaluated using the auto and cross correlation functions, respectively, as described in section 2.2. The process model uses two latent variables, which are linear combinations of the time-lagged values of the output variable and the delayed values of the nine process measurements. For an independent cross validation of the model, the above data was divided to two sections – the first half was used to train the PLS model, and the second used to test model predictions, yielding RMSEP value of 0.26. These results, as shown in Fig. 9. and Table 3, are very similar to the fitting obtained for the simulation data of the same training length to τ ratio (Fig. 2 and Fig 4. upper plot).

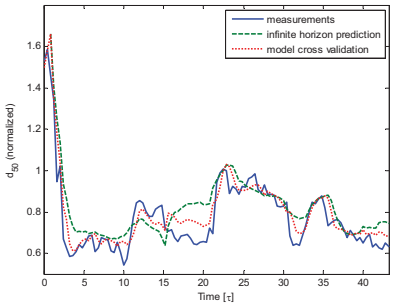


Fig. 7. Dynamic PLS model fitting to plant d_{50} data.

Table 3: Percent variance captured by PLS model, plants' d50 data.

Training set length	LVs	X Block		Y Block	
		LVi	Total	LVi	Total
60τ	1	24.97	24.97	79.34	79.34
	2	25.09	50.06	9.04	88.38
36τ	1	39.46	39.46	72.13	72.13
	2	23.31	62.77	16.8	88.92

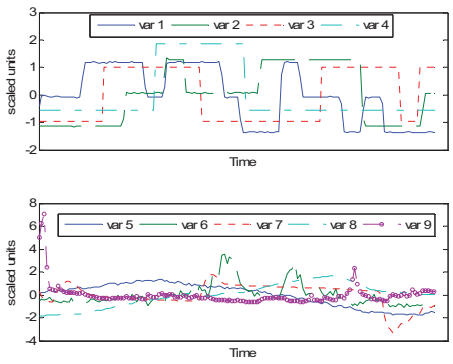


Fig. 8. Values of the 9 predictor variables used in Case I PLS model – 4 manipulated variables (top) and 5 additional process variables (bottom)

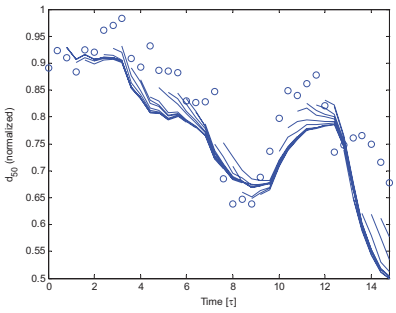


Fig. 9. Cross validation of the dynamic PLS model for plant d_{50} data. Circles represents measurements, lines represent 8 point horizon prediction.

3.3 PLS modeling of real plant data- Case II

In a separate test, the granules distribution width as a function of selected process variables was modeled. This analysis was performed on two limited sets of data, each from a different operating day. A series of step tests were performed on one of the manipulating variables. As in the previous case, other adjustments were continuously made to other plant variables to maintain normal plant operation. The standard deviation of the granules measured area was used as the output variable to be modeled. A dynamic PLS model was built using 3 input variables and one lagged output variable, based on the first data set, and then validated using the second data set, resulting in an excellent fit (Figure 10). In this case, as well, the process model uses two latent variables.

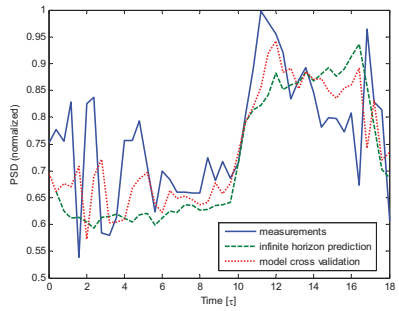


Fig. 10. Actual process data (granules area standard deviation): Cross validation of model based on first data set, tested on second data set

On the scores plot (Figure 11) it is clear, however, that these two sets represent different and distinct operating conditions. If one further examines the contribution of each variable for these two sets (Figures 12) we can see that the main difference is that on the validation set, much lower values of variable 2 were used, compared to the modeling set. It is also interesting to note that the only outlier of the modeling set also exhibits the same low value on variable 2.

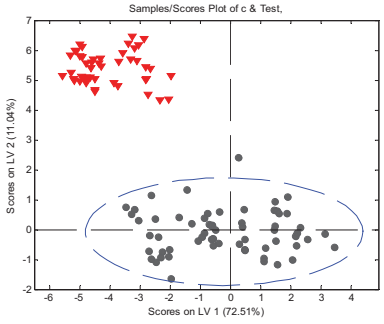


Fig. 11. Score Plots for the PSD width model. Circles are samples from the modeling set; Triangles are samples from the validation set. The ellipse marks the 95% confidence limit for the model.

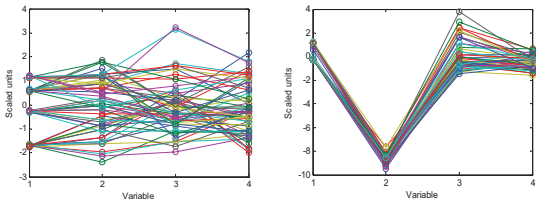


Figure 12: Values of input variables used in the modeling set (left) and in the validation set (right).

Although these results looks promising with respect to the ability to analyze and predict the granulation process variables, a quick look at the high values of Q residuals and Hotelling T^2 (Figure 13) indicates that this model is far from describing the whole complexity of the process, and many more measurements should be done in order to characterize the different operating regimes of this process.

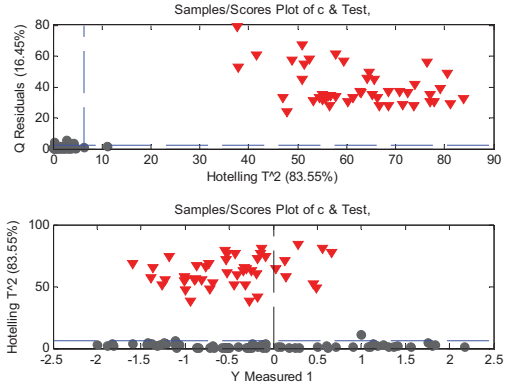


Figure 13: Q Residuals and Hotelling T^2 Values for the PSD width model

4. CONCLUSIONS

Dynamic PLS modeling was proven to be an effective tool in modeling key process variables in an industrial granulation process. Our future work will explore methods to capture the additional dynamics that remain in the plant data. We are also planning longer plant runs with larger input variable excitation to improve the model identification. Longer term goals are to develop a model-based controller for plant testing.

Funding for this work from IFPRI (International Fine Particle Research Institute) is gratefully acknowledged.

REFERENCES

Dayal, B.S., MacGregor, J.F., Taylor, P.A., Kildaw, R. and Marcikic, S. (1994). Application of Feedforward Neural Networks and Partial Least Squares Regression for Modeling Kappa Number in a Continuous Kamyr Digester, *Pulp and Paper Canada*, 95, 26-32.

García, C.E., Prett, D.M. and Morari, M. (1989), Model Predictive Control: Theory and Practice—a Survey, *Automatica*, 25, 335–348.

Glaser, T., Sanders, C.F.W., Wang, F.Y., Cameron, I.T., Litster, J.D., Poon, J.M.H., Ramachandran, R., Immanuel, C.D. and Doyle, F.J. (2008), Model Predictive Control of Continuous Drum Granulation, *Journal of Process Control*, in press.

Iveson, S.M., Litster, J.D., Hapgood, K. and Ennis, B.J. (2001), Nucleation, Growth and Breakage Phenomena in Agitated Wet Granulation Processes: A Review, *Powder Technology*, 117, 3-39.

Juricek, B.C., Seborg, D.E., and Larimore, W. E. (2001), Identification of the Tennessee Eastman Challenge Process with Subspace Methods, *Control Engineering Practice*, 9, 1337-1351.

Kasper, M.H. and Ray, W.H. (1993), Dynamic PLS Modelling for Process Control, *Chemical Engineering Science*, 20, 3447-3461.

MacGregor, J.F. and Kourti, T. (1995), Statistical Process Control of Multivariate Processes, *Control Engineering Practice*, 3, 403-414.

Mort, P.R., Capeci, S.W. and Holder J.W. (2001), Control of Agglomerate Attributes in a Continuous Binder-Agglomeration Process, *Powder Technology*, 117, 173-176.

Wise, B.M., Gallagher, N.B., Bro, R., Shaver, J.M., Windig W. and Koch, R.S. (2006), *PLS_Toolbox 4.0*, Eigenvector Research Incorporated, Wenatchee WA.

Control and Estimation of Distillation Systems

Oral Session

Geometric Estimation of Binary Distillation Columns

Carlos Fernandez and Jesus Alvarez*

Universidad Autónoma Metropolitana-Iztapalapa
Departamento de Ingeniería de Procesos e Hidráulica.
Apdo. 55534, 09340 México, D.F. MEXICO
**(e-mail: jac@xanum.uam.mx)*

Abstract: The problem of designing the estimation structure to perform a specific (entire profile, single/two-effluent, and so son) concentration estimation task in a binary distillation column with two temperature measurements is addressed within a geometric estimation (GE) framework. The structure design involves the choices of: (i) the measurement locations, (i) the complete or truncated estimation model, and (iii) the innovated-noninnovated state partition of the model. First, a structural analysis is performed on the basis of detectability measures in the light of the column characteristics, yielding candidate structures for a specific estimation task. Then, the behavior of the structures is assessed with estimator functioning and dimension considerations, yielding conclusive structural results. The proposed methodology is applied to a representative case example with experimental data.

Keywords: Nonlinear estimator, geometric estimator, nonlinear system realization, distillation column estimation.

1. INTRODUCTION

The study of the concentration estimation problem for distillation columns is motivated by the development of advanced monitoring and control schemes. The estimator design involves decisions on: (i) the structure (sensor number and locations), and (ii) the kind of (EKF, Luenberger, Geometric, etc) algorithm. In the distillation column estimation field: (i) the EKF has been, by far, the most widely employed and accepted algorithm, (ii) the sensor structure has been decided with heuristics (Oisiovici and Cruz, 2000) or observability measures (van der Berg et al. 2000; Singh and Hahn, 2005), (ii) only in a few studies (Alonso et al, 2004; Bian and Henson 2006; Venkateswarlu and Kumar, 2006; Kadu et al, 2008) the measure-based sensor location results have been tested with estimator functioning and is not clear to what extent the results depend on tuning. In principle, the resolution of this drawback requires a unified framework to address the algorithm, sensor location, and tuning aspects.

Recently (Alvarez and Fernandez, 2009), the general-purpose nonlinear geometric estimator (GE) (Alvarez and Lopez, 1999; Alvarez 2000) has been redesigned with: (i) an implementation in terms of model Jacobian matrices (and not of the cumbersome or intractable Lie derivation-based gain of the nonlinear Luenberger observer), (ii) the structure (sensors, complete/truncated model, innovation scheme) as a key design degree of freedom, and (iii) a simple tuning procedure based on a robust convergence criterion, regardless of the structure, and (iv) testing with an experimental binary distillation column. This adjustable-structure GE methodology and the associated nonlinear detectability measures (Lopez and Alvarez, 2004) constitute the methodological points of departure for the present study.

In this work, the problem of designing the best estimation structure (in terms of reconstruction speed, robustness, and algorithm simplicity) to perform a specific (entire profile or two-effluent) concentration estimation task in a binary distillation column with two temperature measurements is addressed, with structure meaning the choices of: (i) the *measurement locations*, (i) the (complete or truncated) *estimation model*, and (iii) the innovated-noninnovated model *state partition*. The proposed methodology is applied to a representative case example with experimental data.

2. ESTIMATION PROBLEM

2.1 Column system and model

Consider an N-stage binary distillation column, with molar feed flow F in tray n_f at (light-component) mole fraction c_F, and bottoms (or distillate) flow B (or D) at composition c_B (or c_D), and a total condenser. Under standard (constant pressure, stage equilibrium, fast holdup dynamics with perfect mixing, evaporator level control, constant molar flow, saturated feed, and adiabatic system) assumptions the *N-composition column model* is given by (Luyben, 1990)

$$\dot{c}_1 = \{(R + F)(c_2 - c_1) - V[\varepsilon(c_1) - c_1]\}/M_1 := f_1, \qquad c_B = c_1$$

$$\dot{c}_i = \{(R + F)(c_{i+1} - c_i) - V[\varepsilon(c_i) - \varepsilon(c_{i-1})]\}/\eta^{-1}(R + F) := f_i,$$
$$1 \le i \le n_f - 1, \qquad \eta(M) = a_\eta(M - M_o)^{b_\eta}$$

$$\dot{c}_{nf} = \{(R + F)(c_{nf+1} - c_{nf}) - V[\varepsilon(c_{nf}) - \varepsilon(c_{nf-1})] + F(c_F - c_{nf})\}/\eta^{-1}(R + F) := f_{nf}$$

$$\dot{c}_i = \{R(c_{i+1} - c_i) - V[\varepsilon(c_i) - \varepsilon(c_{i-1})]\}/\eta^{-1}(R) := f_i, \quad n_f + 1 \le i \le n - 1$$

$$\dot{c}_N = \{R[\varepsilon(c_N) - c_N] - V[\varepsilon(c_N) - \varepsilon(c_{N-1})]\}/\eta^{-1}(R) := f_N, \; c_D = \varepsilon(c_N)$$
$$y_1 = T_1 = \beta(c_{l1}), \qquad y_2 = T_2 = \beta(c_{l2})$$

c_i is the i-th stage mol fraction of light component, V (or R) is the vapor (or reflux) flow, M_i is the i-th stage molar holdup, and T_i is the temperature measurement at the l_1-th stage, ε, β, and η are the liquid-vapor, bubble point, and (Francis weir equation) hydraulics functions, respectively. Assuming the feed composition is fixed at $\bar{c}_e$, the *preceding N-composition column model* is written by

$$\dot{c} = f_c(c, u), \quad c(0) = c_o, \quad y = h(c) := [\beta(c_{l_1}), \beta(c_{l_2})] \qquad (1)$$
$$c = (c_1,\ldots, c_N)', \quad u = (F, R, V)', \quad \dim (c, y, u) = (N, m, 3, 1)$$

In virtue of the afore stated modeling assumptions, the *actual column dynamics* are given by

$$\dot{c} = f_c(c, u) + \tilde{f}_c(c, \xi, d), \; c(t_o) = c_o, \; y = h(c) + \tilde{h}_c(c, \xi) + e_y \quad (2a)$$
$$\dot{\xi} = f_\xi(c, \xi, d, \dot{d}), \qquad \xi(t_o) = \xi_o, \; d = (u, e_u, d_e), \; \dim \xi = n_\xi \quad (2b)$$

with concentration (or unmodeled) state x (or ξ), actuator error e_u, and unmodeled exogenous input d_e. The unmodeled dynamics (2b) have slow and fast components due to the modeling-measurement errors, including holdup and enthalpy QSS assumptions. Thus, the N-composition model (1) is the actual system (2) with the modeling assumption $(\tilde{f}_c, \tilde{h}_c, e_y) = 0$.

2.2 *Adjustable*-structure estimation model

Following a previous binary distillation column GE (Alvarez and Fernandez, 2009) study with the estimation model regarded as design degree of freedom, rewrite the actual column dynamics (1) in terms of $n \le N$ modeled compositions (x):

$$\dot{x} = f(x,u) + \tilde{f}(x, \chi, d, u), \; x(t_o) = x_o, \; y = h(x) + \tilde{h}(x, \chi) + e_y \quad (3a)$$
$$\dot{\chi} = f_\chi(x, \chi, d, \dot{d}), \; \chi(t_o) = \chi_o, \; (x, x_\varnothing) = I_c c, \; \chi = (x_\varnothing, \xi) \qquad (3b)$$

where $x_\varnothing$ are the unmodeled concentrations and χ is the augmented unmodeled state. The vector f depends only on the modeled concentrations x, due to a key modeling assumption made for estimator decentralization purposes (Alvarez and Fernandez, 2009): (i) f is the part of f_c that describes x, and is calculated with the unmodeled state at an average constant value ($\bar{x}_\varnothing$), as the related error can be effectively compensated by the GE integral action when the estimation structure is adequately chosen. In terms of κ_1 (or κ_2) *innovated sates* x_1 (or x_2) and $(n - \kappa)$ *noninnovated states* (x_ι), the *actual dynamics* (2) are given by

$$\dot{x}_1 = f_1(x_1, u) + \tilde{f}_1(x, \chi, d, u), \; y_1 = h_1(x_1) + \tilde{h}_1(x, \chi) + e_1 \quad (4a)$$
$$\dot{x}_2 = f_2(x_2, u) + \tilde{f}_2(x, \chi, d, u), \; y_2 = h_2(x_2) + \tilde{h}_2(x, \chi) + e_2 \quad (4b)$$
$$\dot{x}_v = f_v(x_v, x_1, x_2, u) + \tilde{f}_v(x, \chi, d, u), \quad x_v(t_o) = x_{vo} \quad (4c)$$
$$\dot{\chi} = \tilde{f}_\chi(x, \chi, d, \dot{d}), \quad \chi(t_o) = \chi_o, \; x_1(t_o) = x_{1o}, \quad x_2(t_o) = x_{2o} \quad (4d)$$

where $(x_1, x_2, x_v, x_\varnothing) = I_c c$, $\dim (x_1, x_2, x_v) = (\kappa_1, \kappa_2, \kappa_v)$
$\kappa_1 + \kappa_2 = \kappa_\iota$, $\kappa_1 + \kappa_2 + \kappa_v = n \le N$, $(x_1, x_2) = x_\iota$, $\dim (x_\iota) = \kappa_\iota$

f_1 (or f_2) corresponds to x_1 (or x_2), and is calculated with some (average) constant value [$\bar{x}_1$(or $\bar{x}_2$), $\bar{x}_v$]. From the specialization of the general-purpose definition of model structure [Alvarez and Fernandez, 2009] to the binary column case, the definition of *column model structure* follows

$$\sigma = (\kappa, x_\iota\text{-}x_v), \; \kappa = (\kappa_1, \kappa_2), \quad \kappa_1 + \kappa_2 = \kappa_\iota \le n, \; x_\iota = (x_1^\iota, x_2^\iota)' \quad (5)$$

where κ is the *estimation order* vector, x_ι-x_v is the innovated-noninnovated state partition, and x_1 (x_2) are κ_1 (or κ_2) adjacent concentrations associated with the measurement y_1 (or y_2). Thus, from the enforcement of the modeling assumption

$$(\tilde{f}_1, \tilde{f}_2, \tilde{f}_v) = 0, \qquad (\tilde{h}_1, \tilde{h}_2) = 0, \qquad (e_1, e_2) = 0 \qquad (6)$$

upon the actual subsystem (4a-c), the *estimation model*, with estimation structure σ (5), follows:

$$\dot{x}_1 = f_1(x_1, u), \qquad x_1(t_o) = x_{1o}, \qquad y_1 = h_1(x_1) \qquad (7a)$$
$$\dot{x}_2 = f_2(x_2, u), \qquad x_2(t_o) = x_{2o}, \qquad y_2 = h_2(x_2) \qquad (7b)$$
$$\dot{x}_v = f_v(x_v, x_1, x_2, u), \quad x_v(t_o) = x_{vo} \qquad (7c)$$

2.3 Adjustable-structure geometric estimator (GE)

In virtue of the σ-detectability property of the N-composition staged model (7), the (possibly truncated) model state (x) can be on-line robustly estimated by the *geometric estimator (GE) with structure* σ:

$$\dot{\hat{x}}_1 = f_1(\hat{x}_1, u) + O_1^{-1}(\hat{x}_1, u)\{\pi_1 \hat{\iota}_1 + k_1(\zeta_1, \omega_1)[y_1 - h_1(\hat{x}_1)]\},$$
$$\hat{x}_1(t_o) = \hat{x}_{1o}; \qquad \dot{\hat{\iota}}_1 = \omega_1^{\kappa 1+1}[y_1 - h_1(\hat{x}_1)]\}, \qquad \hat{\iota}_1(0) = \hat{\iota}_{1o} \qquad (8a)$$
$$\dot{\hat{x}}_2 = f_2(\hat{x}_2, u) + O_2^{-1}(\hat{x}_2, u)\{\pi_2 \hat{\iota}_2 + k_2(\zeta_2, \omega_2)[y_2 - h_2(\hat{x}_2)]\},$$
$$\hat{x}_2(t_o) = \hat{x}_{2o}; \qquad \dot{\hat{\iota}}_2 = \omega_2^{\kappa 2+1}[y_2 - h_2(\hat{x}_2)], \qquad \hat{\iota}_2(0) = \hat{\iota}_{2o} \qquad (8b)$$
$$\dot{\hat{x}}_v = f_v(\hat{x}_\iota, \hat{x}_1, \hat{x}_2, u) \qquad \hat{x}_v(t_o) = \hat{x}_{vo} \qquad (8c)$$

where: $k_i(\zeta_i, \omega_i) = [a_1^i(\zeta_i)\omega_i, \ldots, a_{\kappa i}^i(\zeta_i)\omega_i^{\kappa i}]'$ (9c)
$O_i^\iota(x_i, u) = [\beta'(c_{l_i})]e_i[I, A_i(x_i, u), \ldots, A_i^{\kappa i+1}(x_i, u)]$, $i = 1, 2$ (9a)
$A_i(x, u) = \partial_x f(x_i, u)$, $\pi_i = (0, \ldots, 0, 1)'$, $\dim \pi_i = \kappa_i$ (9b)
$|\beta'(c_{l_1})| \ge \varepsilon_T, \quad |\beta'(c_{l_2})| \ge \varepsilon_T$ (10a-b)

ζ_i (or ω_i) is the damping factor (or characteristic frequency) of the prescribed linear, noninteractive, pole assignable (*LNPA*) output error dynamics

$$\tilde{y}_i^{(\kappa i+1)} + a_1^i(\zeta_i)\,\omega_i \tilde{y}_i^{(\kappa i)} + \ldots + a_1^i(\zeta_i)\,\omega_i \tilde{y}_i^{(1)} + \omega_i^{\kappa i+1}\tilde{y}_i = 0, \; i = 1, 2 \quad (11)$$

with coefficient sets $\{a_1,\ldots, a_{\kappa_i}\}_i$ determined by pole placement (Lopez and Alvarez, 1999). The invertibility of O_1 and O_2 is ensured by the tridiagonal state dependency of f_c and the sensor location condition (10) which amounts placing each sensor at a tray with temperature gradient larger than a minimum value (say, two degrees). For any model structure σ (5) (Alvarez and Fernandez, 2009): (i) the afore stated nonlocal robustness convergence feature holds with respect to the N-composition model (1) with decentralization,

truncation, and actuator-measurement errors, and (ii) the rather simple GE tuning scheme applies to any structure. Thus, the adjustable structure-algorithm GE methodological framework offers the means to fairly compare the behavior of different structures, in the sense that the behavior differences are due to the structures itself and not to the tuning.

2.4 Estimation structure design problem

In view of the preceding adjustable-structure column GE approach, *our present problem* consists in, given a specific estimation objective, determining the two-measurement structure, which yields the best estimator behavior in terms of reconstruction speed, robustness and dimensionality. Technically speaking, the problem amounts to choosing the (complete or truncated) estimation model (7) and its σ-detectability structure (5), or equivalently: (i) the location l_1 (or l_2) of the temperature measurement y_1 (or y_2), (ii) the corresponding innovated concentrations x_1 (or x_2), and (iii) the noninnovated state (x_v).

2.5 Structure search methodology

From the perspective of a general-purpose mixed-integer optimization approach, in our 12-stage distillation column example, the N-composition model offers 527, 345 structural possibilities with 4,095 observable (or passive) ones structures, and the number of possibilities grows even more when model truncation is considered. Leaving aside the implementation complexity and difficulties of an optimization-based search method, in the spirit of the constructive control (Sepulchre, 1997) and GE (Alvarez and Fernandez, 2009) approaches, here the structural search will performed by exploiting the column staged feature in the light of the easy to compute version (Alvarez and Fernandez, 2009) of the GE detectability measures (Lopez and Alvarez, 2004), in two steps: (i) first, detectability measures will be used to draw candidate structures for a given estimation objective, and (ii) then, conclusive structural results will be obtained in terms of GE functioning.

2.6 Experimental case example

The proposed methodology will be illustrated and tested with experimental case example employed before to illustrate and test the theoretically drawn features and capabilities of the general-purpose GE approach (Alvarez and Fernandez, 2009): a methanol-water mixture feed F = 40 ml/min, at light component composition c_e = 0.2 and temperature 57°C. Initially, the column was at a steady-state with low reflux ratio (R/D = 0.2) and poor separation ($c_B \approx 0.0$, $c_D \approx 0.57$). Then, at time t = 0, a feed concentration step increase (c_e: 0.2 $\rightarrow$ 0.4) was introduced, yielding: (i) an overall composition response that settled (≈ 40 min) at an intermediate separation steady-state ($c_B \approx 0.01$, $c_D \approx 0.79$), and (ii) a distillate (or bottoms) composition settling time of ≈ 15 (or 40) min. Finally, at t = 40 min, a reflux step increase (R/D: 0.2 $\rightarrow$ 1.5) was introduced, yielding: (i) an overall response that settled (≈ 60 min.) at a high-separation steady-state ($c_B \approx 0.15$, $c_D \approx$ 0.98), and (ii) a distillate (or bottoms) composition settling time of ≈ 20 (or 50) min. The experimental data can be seen in (Alvarez and Fernandez, 2007 and 2009).

3. STRUCTURAL ANALYSIS

In this section, the dependency of the GE detectability measures (Lopez and Alvarez, 2004; Alvarez and Fernandez, 2007) over sensor location and innovated state dimension are analyzed to draw candidate structures for complete profile and two-effluent estimation purposes.

3.1 Detectability measures

To account for the effect of the decentralization-truncation performed in the passage from the complete N-composition (1) to the truncated-decentralized n-composition estimation model (7) with structure σ (5), the detectability measures (12) for the next N-composition model (13) with innovated-noninnovated state partition will be employed (Alvarez and Fernandez, 2009):

$$s_\iota = 1/\mathrm{msv}(O), \quad c_\iota = \mathrm{cn}(O) \qquad (12a\text{-}b)$$

$$s_v = \mathrm{msv}(F), \quad c_v = \mathrm{cn}(F); \quad \lambda_v = \tfrac{1}{2}\,\mathrm{lev}(F + F') < 0 \qquad (12c\text{-}d)$$

$$\dot{x}_\iota = f_\iota(x_\iota, x_v, u), \quad x(0) = x_{\iota o}, \; y = h(x_\iota), \; \dim(x_\iota) = N \qquad (13a)$$

$$\dot{x}_v = f_v(x_\iota, x_v, u), \quad x_v(t_o) = x_{vo}, \quad \dim(x_v) = N\text{-}n \qquad (13b)$$

where $O(x, u) = \mathrm{bd}(E_1', E_2')'(x, u)$, $A_i(x_i, u) = \partial_{x_\iota} f_\iota(x, u)$

$F(x, u) = [\partial_{x_v} f_v + \partial_{x_\iota} f_v O^{-1} D](x, u)$, $A_v(x, u) = \partial_{x_v} f_v(x, u)$

$E_i'(x, u) = \beta'(c_{l_i}) e_i(I, A_i, \ldots, A_i^{\kappa i-1})(x, u)$, $i = 1, 2$

$D'(x, u) = \beta'(c_{l_i}) e_i(I, A_v, \ldots, A_v^{\kappa i-1})(x, u)$, $(x_1', x_2')' = x_\iota$

s_ι (or s_v) is the singularity measure equal to the inverse of the minimum singular value (msv) of the matrix O (or F), c_ι (or c_v) is the illconditioning measure equal to condition number (cn) of the matrix O (or F), and λ_v is the dominant frequency of the noninnovated dynamics, or equivalently, the negative of the smallest eigenvalue (lev) of the matrix (F + F')/2, O is the estimation matrix of the σ-structure model (13), and F is the Jacobian matrix of the noninnovated dynamics (13b). The illconditioning value c_ι (or c_v) measures the overshoot response of the innovated (or noninnovated) state estimation error to an initial estimate error, and the singularity value s_ι (or s_v) measures the asymptotic offset of the innovated (or noninnovated) state error due to persistent modeling errors, and the number λ_v measures the convergence rate of the noninnovated state error dynamics. In general, these measures can be taken over a column motion x(t) (Lopez and Alvarez, 2004). In our column case, the detectability measures will be computed at the intermidiate steady state (reached after ≈ 40 min).

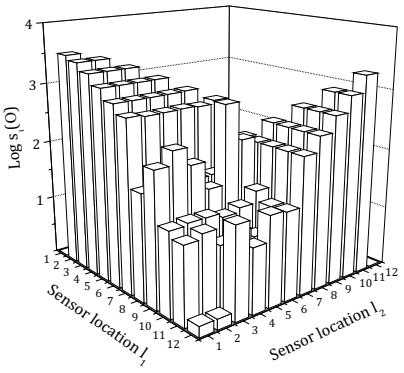
3.2 Measurement locations

In Figure 1 are presented the singularity (s_ι) and ill conditioning (c_ι) measures (12) of the estimation matrix O (9a) as function of the sensor location pair (l_1, l_2), for a completely observable structure σ (5) with estimation order pair $\kappa = (\kappa_1, \kappa_2) = (6, 6)$ and $\kappa_\iota = \kappa_1 + \kappa_2 = 12 = n = N$ (i.e., complete model with observable structure), showing that: (i) the largest singularity and illconditioning values are obtained with the sensor stage location pair (l_1, l_2) ≈ (1 to 2, 12), (ii) the smallest singularity and illconditioning values are

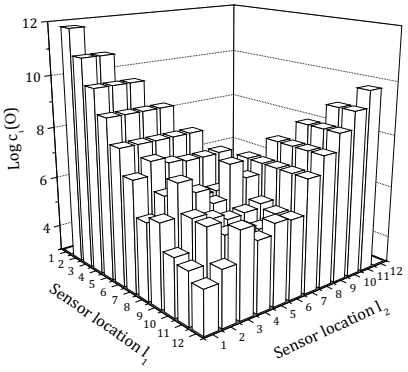
obtained with the sensor location $(l_1, l_2) \approx (1, 2)$ [or (11, 12)] with two adjacent sensors in the bottom (or top) of the column, followed by the sensor location $(l_1, l_2) \approx (5, 6)$ with two adjacent sensors above or below the feed tray (5). These consideration lead to the *following conclusions*: (i) the two sensors should not be in the same section, and (ii) the best location pair for complete profile estimation is given by

$$(l_1, l_2) = (2, 12) := (l_s, l_e) \tag{14}$$

meaning one sensor in stage $l_s = 2$ (tray one) [or $l_e = 12$ tray 10] of the stripping (or enriching section), precisely in the stage with the largest temperature and concentration stage-to-stage gradient.



a)



b)

Fig. 1: Dependency of the singularity s_t and ill conditioning c_t of the estimation matrix O on the sensor location pair (l_1, l_2), for comple observable structure σ (5) estimation order pair $\kappa = (\kappa_1, \kappa_2)$ and $\kappa_1 + \kappa_2 = 12$.

These location results are in agreement with location criteria employed in: (i) two-point temperature PI control of distillation columns (Tolliver, 1980; Castellanos-Sahagun et al., 2005), and (ii) previous distillation column studies with EKF (Baratti et al., 1995; Oisiovici and Cruz, 2000).

3.3 Innovated state dimension pair

Next, the sensor pair location (14) determined in the last section for the complete model (1), with observable structure κ (5), is kept fixed, and the dependency of the illconditioning c_t (12a) and the speed parameter λ_v (12c) of the noninnovated dynamics upon the innovated state dimension (or equivalently, estimation order) pair $\kappa = (\kappa_s, \kappa_e)$ is examined, with κ_s (or κ_e) being the number of adjacent innovated states x_s (or x_e) associated with the measurement y_s (or y_e) of the stripping (or enriching) section. The resulting measure c_t (or λ_v) is presented in Figure 2a (or 2b), showing that the illconditioning measure c_t remains within a reasonable bound $(1 \le c_t \le 100)$ for all the estimation order pairs with at most three innovated states per measurement $[(\kappa_s, \kappa_e) \le (3, 3)]$. As expected (Lopez and Alvarez, 2004): (i) the passive structure $(\kappa_s, \kappa_e) = (1, 1)$ yields the smallest possible value $c_t = 1$, and (ii) the speed parameter λ_v of the noninnovated dynamics is minimum ($\lambda_v = 0$) when the structure is observable with $(\kappa_s, \kappa_e) = (6, 6)$. These results are consistent with the general-purpose GE approach (Alvarez y Fernandez, 2009): (i) as the number of innovated states grows, the reconstruction speed grows and the robustness decreases, and (ii) the maximum robustness is obtained with the passive structure $(\kappa_s, \kappa_e) = (1, 1)$, and (iii) the maximum reconstruction speed is obtained with the observable structures $(\kappa_s, \kappa_e) = (6, 6)$.

3.4 Candidate structures

From the preceding structural analysis the next results follow. For *complete profile estimation*, the candidate models are decentralized versions of the N-concentration model (1) with structure σ (5):

$$\dot{c}_s = f_s(c_s, u), \qquad \dot{c}_e = f_e(c_e, u), \quad y = h(c) \tag{15a}$$
$$(\kappa_s, \kappa_e) \le (3, 3), \quad (l_s, l_e) = (2, 12) \tag{15b}$$

For *two-effluent estimation*, the candidate model is the truncated-decentralized model with passive structure σ (5):

$$\dot{c}_1 = \{(R + F)(c_2 - c_1) - V[\varepsilon(c_1) - c_1]\}/M_1, \tag{16a}$$
$$\dot{c}_2 = \{(R + F)(\bar{c}_3 - c_2) - V[\varepsilon(c_2) - \varepsilon(c_1)]\}/\eta^{-1}(R + F), \; y_s = \beta(c_2)$$
$$\dot{c}_{12} = \{R[\varepsilon(c_{12}) - c_{12}] - V[\varepsilon(c_{12}) - \varepsilon(\bar{c}_{11})]\}/\eta^{-1}(R), \quad y_e = \beta(c_{12})$$
$$(\kappa_s, \kappa_e) = (1, 1), \qquad (l_s, l_e) = (2, 12) \tag{16b}$$

4. STRUCTURAL RESULTS

Having as point of departure the suggestive structural results (15, 16) of the last section, in the present section conclusive results are obtained by assessing the candidate structures in terms of reconstruction speed and robustness.

4.1 Estimator tuning and convergence

The column (or holdup) dominant (or fastest) frequency ω_c (or ω_η) was determined from the experimental data and the detailed model (1), the estimator frequency ω is written as n_ω times ω_n, and the adjustable constants (20) are listed next:

$$(\omega_c, \omega_\eta) \approx (1/15, 1) \text{ min}^{-1}; \; \zeta_s, \zeta_e, \omega_s = \omega_e = \omega = n_\omega \omega_c \tag{17-18}$$

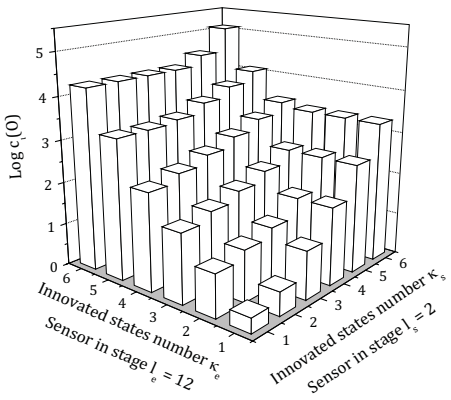
Thus: (i) there are three adjustable gains (ζ_s, ζ_e, n_ω), and (ii) the limit upper ω^+ is related to ω_η. From the specialization to

the column case of Lemma and Proposition 1 in (Alvarez and Fernandez, 2009): (i) the GE error dynamics is robustly stable if the stabilizing term λ_s dominates the potentially destabilizing one λ_d (λ_s and λ_d defined in Alvarez and Fernandez, 2009), according to inequality (19), or equivalently, if: (i) the related threshold equation (20) has two strictly positive and sufficiently separated roots (ω^- and ω^+) for ω, and (ii) the gain frequency ω (18) of the prescribed LNPA output error dynamics (11) is chosen so that the low-high gain conditions (21) are met:

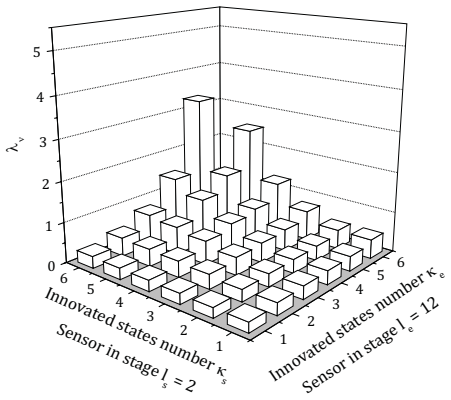
$$\lambda_s(\omega, \zeta, \sigma) - \lambda_d(\omega, \zeta, \sigma) := \lambda(\omega, \zeta, \sigma) > 0, \quad \zeta = (\zeta_s, \zeta_e) \qquad (19)$$

$$\lambda(\omega, \zeta, \sigma) = 0 \quad \Rightarrow \quad \exists \quad \omega = \omega^-(\zeta, \sigma), \quad \omega^+(\zeta, \sigma) \qquad (20)$$

$$0 < \omega^-(\zeta, \sigma) < \omega < \omega^+(\zeta, \sigma) \qquad (21)$$



a)



b)

Fig. 2: Dependency of the singularity (or noninnovated-dynamics speed parameter) s_t (or c_t) of the estimation matrix O (or Jacobian matrix F) on the dimension $\kappa = (\kappa_s, \kappa_e)$ of the innovated state pair x_s-x_e, for the sensor location $(l_s, l_e) = (2, 12)$

The meaning of these conditions and their dependency on κ_t (= κ_1, κ_2, κ_3 with $\kappa_1 < \kappa_2 < \kappa_3$) are depicted in the Figure 1b of Alvarez and Fernandez (2009): (i) the fulfilment ($\kappa_t = \kappa_1$) or violation ($\kappa_t = \kappa_2, \kappa_3$) of the conditions, and (ii) as the estimation order κ_t grows, the convergence gain (ω^+, ω^-) decreases, and eventually vanishes. In our column problem, we shall be interested in the interplay between structure, behavior, and tuning (ζ_s, ζ_e, ω, ω^-, ω^+, $\Delta\omega$).

4.2 Entire profile estimation

First, the GE estimator (8) was run with the candidate structures (κ_s, κ_e) = (1, 1), … (6, 6), finding that the best behavior was attained with (κ_s, κ_e) = (3, 3), followed by (2, 2). In Figure 3 are presented the results for (κ_s, κ_e) = (1, 1), (3, 3), and (6, 6), and the corresponding gain tuning limit results are listed in Table 1. The structure with three innovated states per measurement yields the best speed versus robustness behavior, with a reasonable gain interval (ω^+, ω^-). In agreement with the convergence-tuning theoretical derivations (Alvarez and Fernandez, 2009): (i) the passive (or observable) structure yields the slowest (or fastest) reconstruction rate with the largest (smallest) robustness, or equivalently, the largest (or smallest) gain interval $\Delta\omega$, and (iii) to avoid oscillatory response, the damping factor $\zeta_{s/e} = 2^{1/2}$ (or 1.5) is used for observable (or passive) innovation (Alvarez and Lopez, 1999).

4.3 Two-effluent estimation

In this case, the three-state model with two decoupled subsystems and passive innovation candidate structure (15) was implemented as well as several other neighbouring structures, finding that, the candidate structure yielded the best behavior with the least number of states, followed closely by some neighboring structures. The corresponding tuning and behavior are listed in Table 2 and Figure 4, respectively. As it can be seen in the Figure 4, for effluent estimation purposes, the truncated model outperforms the complete one, and this verifies the effectiveness of setting the model dimension as design degree of freedom. Comparing with the complete model-based estimation cases, the truncated model with single-stage innovation per measurement yields faster and more robust effluent estimates.

Table 1. Tuning for entire profile estimation.

y	κ	ζ	ω^-	ω	ω^+	$\Delta\omega$	n_ω
T_2, T_{12}	6, 6	$2^{1/2}$	1/15	2/5	8/15	7/15	6
T_2, T_{12}	3, 3	1	1/15	2/3	4/5	11/15	10
T_2, T_{12}	1, 1	3/2	1/15	4/5	14/15	13/15	12

Table 2. Tuning for two-effluent estimation.

n	κ	ζ	ω^-	ω	ω^+	$\Delta\omega$	n_ω
12	1, 1	3/2	1/15	4/5	14/15	13/15	12
3	1, 1	3/2	1/15	14/15	1	14/15	14

5. CONCLUSIONS

The problem of drawing the structure for best estimator behavior with respect to a specific concentration estimation task has been resolved for a binary distillation column with two temperature measurements and experimental data. It was found that: (i) the (12-concentration) profile must be estimated with the complete model, six innovated concentrations (three per measurement), and a 6-concentration open-loop observer module, and (ii) the two-effluent concentration must be estimated with a three-stage truncated model, two innovated concentrations, (one per

measurement), and one noninnovated module. In the complete (or two-effluent) estimation case, the GE consists of 5 (or 14) ODE's, which are considerably less than the 72 ODEs required by an EKF implementation.

The proposed approach: (i) resolves the structure-algorithm estimation design problem in a way that is more effective and simpler than the ones of previous studies, and (ii) is a point of departure to address the multi-component case.

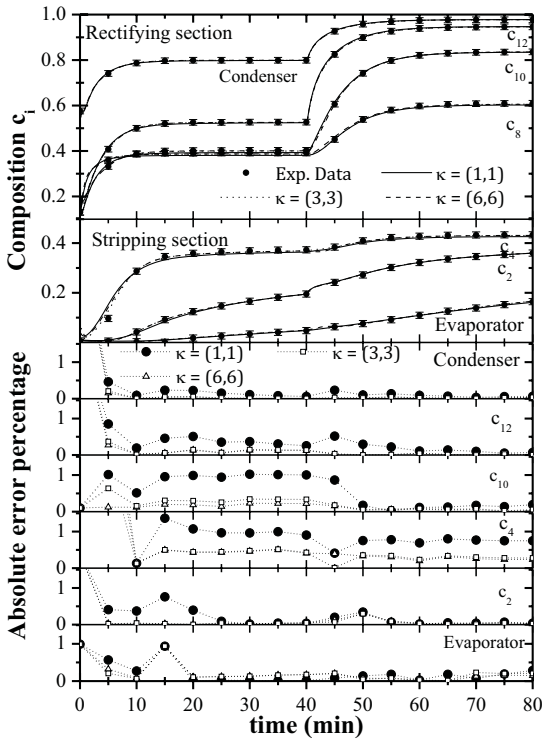


Fig. 3. Profile estimation with two sensors (l_s = 2, l_e = 12) and complete model.

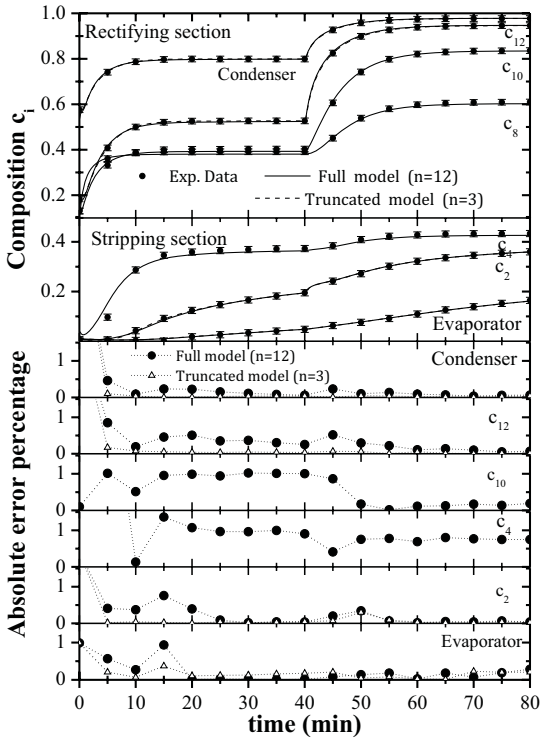


Fig. 4. Two-effluent estimation with two sensors (l_s = 2, l_e = 12) and truncated model.

REFERENCES

Alonso, A. A., Kevrekidis, I. G., Banga, J. R. and Frouzakis, C. E. (2004). Optimal sensor location and reduced order observer design for distributed process systems, *Comput. Chem. Eng.*, 28(1): 27-35

Alvarez, J. and Lopez, T. (1999). Robust Dynamic State Estimation of Nonlinear Plants, *AIChE J.*, 45(1): 107-123

Alvarez, J. (2000). Nonlinear state estimation with robust convergence, *J. Process Control*, 10(1): 59-71

Alvarez, J. and Fernandez, C. (2007). Constructive estimation design for binary distillation columns, *IFAC-DYCOPS 2007*, Preprints (II): 75-80

Alvarez, J. and Fernandez, C. (2009). Geometric estimation of nonlinear process systems, *J. Process Control*, 19(2): 247-260.

Baratti, R., Bertucco, A., Da Rold, A. and Morbidelli, M. (1995). Development of a Composition Estimator for Binary Distillation Columns: Application to a Pilot Plant, *Chem. Eng. Sci.*, 50: 1541-1550

Bian, S. and Henson, M. A. (2006). Measurement selection for on-line estimation of nonlinear wave models for high purity distillation columns, *Chem. Eng. Sci.*, 61(10): 3210-3222

Castellanos-Sahagun, E., Alvarez-Ramírez, J. and Alvarez, J. (2005). Two-point temperature control structure and algorithm design for binary distillation columns, *Ind. Eng. Chem. Res.* 44: 142-152

Kadu, S. C., Bhushan, M. and Gudi, R. (2008). Optimal sensor network design for multirate systems, *J. Process Control*, 18: 594-609

Lopez, T. and Alvarez, J. (2004). On the effect of the estimation structure in the functioning of a nonlinear copolymer reactor estimator, *Journal of Process Control*, 14: 99-109

Luyben, W. L. (1990). *Process modeling, simulation, and control for chemical engineers*. Nueva York, McGraw-Hill.

Oisiovici, R. M. and Cruz, S. L. (2000). State estimation of batch distillation columns using an extended Kalman filter, *Chem. Eng. Sci.*, 55: 4667-4680

Sepulchre, R., M. Janković, P. R.Kokotović. (1997). *Constructive Nonlinear Control*, Communications and Control Engineering Series, Springer - Verlag, London

Singh, A. K. and Hahn, J. (2005). State estimation for high-dimensional chemical processes, *Comput. Chem. Eng.*, 29(11): 2326-2334

Tolliver, T. L. and McCune, L. C. (1980). Finding the Optimum Temperature Control Trays for Distillation Columns, *Instrum. Sci. Technol.* 27(9): 75-80

van den Berg, F. W. J., Hoefsloot, H. C. J., Boelens, H. F. M. and Smilde, A. K. (2000). Selection of optimal sensor position in a tubular reactor using robust degree of observability criteria, *Chem. Eng. Sci.*, 55: 827-837

Venkateswarlu, C. and Jeevan Kumar, B. (2006). Composition estimation of multicomponent reactive batch distillation with optimal sensor configuration, *Chem. Eng. Sci.*, 61: 5260-5274

Distributed Optimization for Predictive Control of a Distillation Column with Output and Control-Input Constraints ⋆

Helton F. Scherer * **Eduardo Camponogara** * **Agustinho Plucenio** *

* *Department of Automation and Systems Engineering, Federal University of Santa Catarina, Florianópolis, SC 88040-900 Brazil (e-mails: scherer@das.ufsc.br, camponog@das.ufsc.br, plucenio@das.ufsc.br)*

Abstract: A distributed predictive control framework based on a state-space model with constraints on the output and control-input is proposed. By a benchmark process, the performance of this framework is analyzed and compared with centralized control strategies in a regulation problem of a distillation column.

Keywords: Model predictive control, distributed optimization, distillation column, convex optimization, interior-point method.

1. INTRODUCTION

Advanced techniques for multivariable control like model predictive control (MPC) have become widespread in the industry, but they are still complex, time-consuming to set-up, and consequently expensive. Further, the centralized approach may not be suitable to the operation of large dynamic networks, either by the communication difficulty between sensors and the central unit, or by the computational limitation to solve optimization problems. Some petrochemical plants are examples of large systems composed by distributed, however coupled subsystems.

An alternative is distributed predictive control (Camponogara et al., 2002), which breaks the static optimization problem into smaller sub-problems to be solved by a network of control/optimization agents. It aims to solve the sub-problems in the most simple form while the final performance is preserved or even improved.

Many studies about distributed formulations are being developed. Mercangöz and Doyle III (2007) propose a distributed formulation that ensures self-sufficient state estimation in each node. Motee and Jadbabaie (2006) present a study of receding horizon control applied to physically decoupled systems with input and state constraints, where the couplings appear through the finite horizon cost function. Li et al. (2005) and Giovanini and Balderud (2006) propose MPC strategies based on Nash optimality to decoupled sub-systems.

Besides this, many algorithms to ensure convergence of distributed problems are being proposed. Dunbar (2007) presents distributed algorithms for dynamically coupled nonlinear systems subject to decoupled input constraints. An iterative procedure based on cooperation that ensures convergence to the global optimum for linear systems with constraints on the local controls is presented by Venkat et al. (2008). Camponogara and Talukdar (2007) present synchronous and asynchronous solutions of optimization problems, proposing a high level optimization framework and safety margins for meeting constraints. Recently, distributed predictive control was specialized to linear dynamic networks and applied to traffic light control, in which the dynamics of the sub-systems are coupled and the constraints are on the local controls (Camponogara and de Oliveira, 2009).

This paper proposes a problem decomposition and distributed algorithm for predictive control of linear networks with dynamic couplings and restrictions on output and control-input signals. Further, it reports on a comparison with existing approaches for the control of a distillation column.

The end result of this research is a distributed predictive control technique for programming control agents which can be deployed to perform regulatory control of linear dynamic networks. Each agent would be responsible for solving a problem of control action, exchanging its local sensor data and control actions with the other agents. The resulting control action is obtained after resolving conflicts with neighboring agents.

Because the algorithms embedded in the agents are much simpler than the centralized one, the distributed approach makes it simpler to modify and reconfigure the plant. Instead of modifying the complex centralized algorithm, it would suffice to add new agents and update only the nearby agents with whom the new agents would have relation. Further, maintenance would be facilitated since the distributed agents are much simpler.

2. DYNAMIC MODEL AND CONTROL PROBLEM

A linear dynamic network is obtained by interconnecting decoupled sub-systems that have local dynamics and controls. The couplings arise from the dynamic interconnections and the constraints on the network's output equations. $\mathscr{M} = \{1, \ldots, M\}$ denotes the set of sub-systems. Each sub-system m is governed by the following discrete-time linear dynamic equation:

$$\mathbf{x}_m(k+1) = A_m \mathbf{x}_m(k) + B_m \mathbf{u}_m(k) \quad (1)$$

⋆ This research has been funded in part by Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) under grant PRH-34/aciPG/ANP and by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under grant 473841/2007-0.

where $\mathbf{x}_m \in \mathbb{R}^{n_m}$ is the state, $\mathbf{u}_m \in \mathbb{R}^{p_m}$ is the control input, and A_m and B_m are matrices of appropriate dimensions. The output from sub-system m depends on the state of the sub-systems in set $I(m) \subseteq \mathscr{M}$ such that $m \in I(m)$:

$$\mathbf{y}_m(k) = \sum_{i \in I(m)} C_{m,i}\mathbf{x}_i(k) \tag{2}$$

and is subject to the output constraints:

$$\mathbf{y}_m^{\min} \leq \mathbf{y}_m(k) \leq \mathbf{y}_m^{\max} \tag{3}$$

with $C_{m,i}$ being matrices of suitable dimensions.

The regulation problem for the overall system subject to output and control-input constraints is:

$$\min \frac{1}{2}\sum_{m=1}^{M}\sum_{k=0}^{\infty}\left[\mathbf{y}_m(k+1)'Q_m\mathbf{y}_m(k+1) + \mathbf{u}_m(k)'R_m\mathbf{u}_m(k)\right] \tag{4a}$$

$$\text{S.to : For } m = 1,\ldots,M,\ k = 0,\ldots,\infty:$$

$$\mathbf{x}_m(k+1) = A_m\mathbf{x}_m(k) + B_m\mathbf{u}_m(k) \tag{4b}$$

$$\mathbf{y}_m(k+1) = \sum_{i \in I(m)} C_{m,i}\mathbf{x}_i(k+1) \tag{4c}$$

$$\mathbf{u}_m^{\min} \leq \mathbf{u}_m(k) \leq \mathbf{u}_m^{\max} \tag{4d}$$

$$\mathbf{y}_m^{\min} \leq \mathbf{y}_m(k+1) \leq \mathbf{y}_m^{\max} \tag{4e}$$

where Q_m are symmetric positive semi-definite and R_m are symmetric positive definite matrices.

Model predictive control solves an optimization problem that approximates the regulation problem for a finite-time horizon. Given the state $\mathbf{x}(k) = (\mathbf{x}_1,\ldots,\mathbf{x}_M)(k)$ of the system at time k, the MPC regulation problem is defined as:

$$P: \min f = \frac{1}{2}\sum_{m=1}^{M}\sum_{j=0}^{T-1}\left[\hat{\mathbf{y}}_m(k+j+1|k)'Q_m\hat{\mathbf{y}}_m(k+j+1|k) + \hat{\mathbf{u}}_m(k+j|k)'R_m\hat{\mathbf{u}}_m(k+j|k)\right] \tag{5a}$$

$$\text{S.to : For } m = 1,\ldots,M,\ j = 0,\ldots,T-1:$$

$$\hat{\mathbf{x}}_m(k+j+1|k) = A_m\hat{\mathbf{x}}_m(k+j|k) + B_m\hat{\mathbf{u}}_m(k+j|k) \tag{5b}$$

$$\hat{\mathbf{y}}_m(k+j+1|k) = \sum_{i \in I(m)} C_{m,i}\hat{\mathbf{x}}_i(k+j+1|k) \tag{5c}$$

$$\mathbf{u}_m^{\min} \leq \hat{\mathbf{u}}_m(k+j|k) \leq \mathbf{u}_m^{\max} \tag{5d}$$

$$\mathbf{y}_m^{\min} \leq \hat{\mathbf{y}}_m(k+j+1|k) \leq \mathbf{y}_m^{\max} \tag{5e}$$

$$\hat{\mathbf{x}}_m(k|k) = \mathbf{x}_m(k) \tag{5f}$$

where $\hat{\mathbf{u}}_m(k+j|k)$ is the prediction for the control input to sub-system m at time $(k+j)$ as predicted at time k, and similarly $\hat{\mathbf{y}}_m$ and $\hat{\mathbf{x}}_m$ are output and state predictions respectively. The variable T is the length of the prediction and control horizons, that have the same length to make the developments simpler.

The term "$|k$" is dropped from all variables for the sake of simplification. Before continuing with the MPC formulation, some terminology will be introduced to simplify the representation. First, it is possible to obtain the state of sub-system m at time $(k+t)$ by using the initial state and the past controls. The future states and outputs are represented as:

$$\hat{\mathbf{x}}_m(k+t) = A_m^t\mathbf{x}_m(k) + \sum_{j=1}^{t} A_m^{j-1}B_m\hat{\mathbf{u}}_m(k+t-j)$$

$$\hat{\mathbf{y}}_m(k+t) = \sum_{i \in I(m)} C_{m,i}\left(A_i^t\mathbf{x}_i(k) + \sum_{j=1}^{t} A_i^{j-1}B_i\hat{\mathbf{u}}_i(k+t-j)\right)$$

By defining the vectors $\hat{\mathbf{x}}_m$, $\hat{\mathbf{u}}_m$, and $\hat{\mathbf{y}}_m$ to represent the predictions over the entire horizon of the states, controls, and outputs, respectively, and the matrices $\overline{CA}_{m,i}$ and $\overline{CB}_{m,i}$ for the dynamics:

$$\hat{\mathbf{x}}_m = \begin{bmatrix}\hat{\mathbf{x}}_m(k+1)\\ \vdots \\ \hat{\mathbf{x}}_m(k+T)\end{bmatrix}, \quad \hat{\mathbf{u}}_m = \begin{bmatrix}\hat{\mathbf{u}}_m(k)\\ \vdots \\ \hat{\mathbf{u}}_m(k+T-1)\end{bmatrix}$$

$$\hat{\mathbf{y}}_m = \begin{bmatrix}\hat{\mathbf{y}}_m(k+1)\\ \vdots \\ \hat{\mathbf{y}}_m(k+T)\end{bmatrix}, \quad \overline{CA}_{m,i} = \begin{bmatrix}C_{m,i}A_i\\ C_{m,i}A_i^2\\ \vdots \\ C_{m,i}A_i^T\end{bmatrix}$$

$$\overline{CB}_{m,i} = \begin{bmatrix} C_{m,i}B_i & 0 & \cdots & 0 \\ C_{m,i}A_iB_i & C_{m,i}B_i & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ C_{m,i}A_i^{T-1}B_i & C_{m,i}A_i^{T-2}B_i & \cdots & C_{m,i}B_i \end{bmatrix}$$

the prediction of the outputs over the entire horizon is written in a compact form as:

$$\hat{\mathbf{y}}_m = \sum_{i \in I(m)}\left(\overline{CA}_{m,i}\mathbf{x}_i(k) + \overline{CB}_{m,i}\hat{\mathbf{u}}_i\right) \tag{6}$$

Defining the matrices $\bar{Q}_m$ and $\bar{R}_m$ with proper dimensions:

$$\bar{Q}_m = \begin{bmatrix} Q_m & 0 & \cdots & 0 \\ 0 & Q_m & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & Q_m \end{bmatrix}, \quad \bar{R}_m = \begin{bmatrix} R_m & 0 & \cdots & 0 \\ 0 & R_m & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & R_m \end{bmatrix}$$

and the vectors $\hat{\mathbf{u}}_m^{\min}$, $\hat{\mathbf{u}}_m^{\max}$, $\hat{\mathbf{y}}_m^{\min}$, and $\hat{\mathbf{y}}_m^{\max}$:

$$\hat{\mathbf{u}}_m^{\min} = \begin{bmatrix}\mathbf{u}_m^{\min}\\ \vdots \\ \mathbf{u}_m^{\min}\end{bmatrix}, \quad \hat{\mathbf{u}}_m^{\max} = \begin{bmatrix}\mathbf{u}_m^{\max}\\ \vdots \\ \mathbf{u}_m^{\max}\end{bmatrix}$$

$$\hat{\mathbf{y}}_m^{\min} = \begin{bmatrix}\mathbf{y}_m^{\min}\\ \vdots \\ \mathbf{y}_m^{\min}\end{bmatrix}, \quad \hat{\mathbf{y}}_m^{\max} = \begin{bmatrix}\mathbf{y}_m^{\max}\\ \vdots \\ \mathbf{y}_m^{\max}\end{bmatrix}$$

problem P is expressed as:

$$P: \min \frac{1}{2}\sum_{m=1}^{M}\left(\hat{\mathbf{y}}_m'\bar{Q}_m\hat{\mathbf{y}}_m + \hat{\mathbf{u}}_m'\bar{R}_m\hat{\mathbf{u}}_m\right) \tag{7a}$$

$$\text{S.to : For } m = 1,\ldots,M:$$

$$\hat{\mathbf{y}}_m = \sum_{i \in I(m)}\left(\overline{CA}_{m,i}\mathbf{x}_i(k) + \overline{CB}_{m,i}\hat{\mathbf{u}}_i\right) \tag{7b}$$

$$\hat{\mathbf{u}}_m^{\min} \leq \hat{\mathbf{u}}_m \leq \hat{\mathbf{u}}_m^{\max} \tag{7c}$$

$$\hat{\mathbf{y}}_m^{\min} \leq \hat{\mathbf{y}}_m \leq \hat{\mathbf{y}}_m^{\max} \tag{7d}$$

The next step is to represent problem P using only the current state and the control predictions. First, let us define f_m as the portion of problem P for a specific m and replace (6) in the objective function:

$$
\begin{aligned}
f_m = & \frac{1}{2}\left(\sum_{i\in I(m)}\left(\overline{CA}_{m,i}\mathbf{x}_i(k)+\overline{CB}_{m,i}\hat{\mathbf{u}}_i\right)\right)'\bar{Q}_m \\
& \cdot\left(\sum_{i\in I(m)}\left(\overline{CA}_{m,i}\mathbf{x}_i(k)+\overline{CB}_{m,i}\hat{\mathbf{u}}_i\right)\right)+\frac{1}{2}\hat{\mathbf{u}}_m'\bar{R}_m\hat{\mathbf{u}}_m \\
= & \frac{1}{2}\left(\sum_{i\in I(m)}\overline{CA}_{m,i}\mathbf{x}_i(k)\right)'\bar{Q}_m\left(\sum_{i\in I(m)}\overline{CA}_{m,i}\mathbf{x}_i(k)\right) \\
& +\left(\sum_{i\in I(m)}\overline{CA}_{m,i}\mathbf{x}_i(k)\right)'\bar{Q}_m\left(\sum_{i\in I(m)}\overline{CB}_{m,i}\hat{\mathbf{u}}_i\right) \\
& +\frac{1}{2}\left(\sum_{i\in I(m)}\overline{CB}_{m,i}\hat{\mathbf{u}}_i\right)'\bar{Q}_m\left(\sum_{i\in I(m)}\overline{CB}_{m,i}\hat{\mathbf{u}}_i\right) \\
& +\frac{1}{2}\hat{\mathbf{u}}_m'\bar{R}_m\hat{\mathbf{u}}_m
\end{aligned}
\tag{8}
$$

Defining vectors $\mathbf{g}_{m,i,j}$, matrices $H_{m,i,j}$, and a constant c_m to represent the terms of f_m:

$$\mathbf{g}_{m,i,j} = \overline{CB}_{m,i}'\bar{Q}_m\overline{CA}_{m,j}\mathbf{x}_j(k) \text{ for } i,j\in I(m) \tag{9a}$$

$$H_{m,m,m} = \overline{CB}_{m,m}'\bar{Q}_m\overline{CB}_{m,m}+\bar{R}_m \tag{9b}$$

$$H_{m,i,j} = \overline{CB}_{m,i}'\bar{Q}_m\overline{CB}_{m,j} \quad \text{for } i,j\in I(m), i\neq m \text{ or } j\neq m \tag{9c}$$

$$c_m = \frac{1}{2}\sum_{i\in I(m)}\sum_{j\in I(m)}\mathbf{x}_i(k)'\overline{CA}_{m,i}'\bar{Q}_m\overline{CA}_{m,j}\mathbf{x}_j(k) \tag{9d}$$

it is possible to redefine problem P as:

$$P: \min\frac{1}{2}\sum_{m=1}^{M}\sum_{i\in I(m)}\sum_{j\in I(m)}\left[\hat{\mathbf{u}}_i'H_{m,i,j}\hat{\mathbf{u}}_j+\mathbf{g}_{m,i,j}'\hat{\mathbf{u}}_i\right]+\sum_{m=1}^{M}c_m \tag{10a}$$

S.to : For $m=1,\ldots,M$:

$$\hat{\mathbf{u}}_m^{\min}\leq\hat{\mathbf{u}}_m\leq\hat{\mathbf{u}}_m^{\max} \tag{10b}$$

$$\hat{\mathbf{y}}_m^{\min}\leq\sum_{i\in I(m)}\left(\overline{CA}_{m,i}\mathbf{x}_i(k)+\overline{CB}_{m,i}\hat{\mathbf{u}}_i\right)\leq\hat{\mathbf{y}}_m^{\max} \tag{10c}$$

This quadratic programming formulation will be used to solve the problem of control calculation in the centralized approach.

2.1 Logarithmic Barrier Method

The logarithmic barrier method is an *interior-point* method for solving convex optimization problems with inequality constraints (Boyd and Vandenberghe, 2004),

$$\text{minimize } f(\mathbf{x}) \tag{11a}$$

$$\text{subject to } A\mathbf{x}\leq\mathbf{b}, \tag{11b}$$

It is assumed that the problem is solvable, *i.e.*, an optimal solution $\mathbf{x}^\star$ exists, and the constraints delimit a closed set. A *barrier function* is any function $B(\mathbf{x}):\Re^n\to\Re$ that satisfies

- $B(\mathbf{x})\geq 0$ for all $\mathbf{x}$ that satisfy $A\mathbf{x}<\mathbf{b}$, and
- $B(\mathbf{x})\to\infty$ as $\mathbf{x}$ approaches the boundary of $\{\mathbf{x}|A\mathbf{x}\leq\mathbf{b}\}$

Being $\mathbf{a}_i'$ the i-th row of A, the idea of the method is to treat the constraints using a logarithmic barrier function as follows:

$$\phi(\mathbf{x}) = -\sum_{i=1}^{m}\log\left(b_i-\mathbf{a}_i'\mathbf{x}\right) \tag{12}$$

where the domain of ϕ is $\mathbf{dom}\ \phi=\{\mathbf{x}|A\mathbf{x}<\mathbf{b}\}$.

With (12), problem (11) can be approximated by:

$$P(\varepsilon):\ \min\ g(\mathbf{x}) = f(\mathbf{x})+\varepsilon\phi(\mathbf{x}) \tag{13}$$

where $\varepsilon>0$ is a parameter that sets the accuracy of the approximation. As ε decreases, more accurate the approximation $P(\varepsilon)$ becomes, whose optimal solution is $\mathbf{x}(\varepsilon)$. The optimal solution is reached by solving (13) for a decreasing sequence of $\varepsilon\to 0^+$, *i.e.*, $\lim_{\varepsilon\to 0^+}\mathbf{x}(\varepsilon)=\mathbf{x}^\star$. The pseudo-code of the barrier method for solving problem (11) appears in Algorithm 1.

Newton's method can be used to compute the optimal solution to $P(\varepsilon)$ using the gradient and the Hessian of $\phi(\mathbf{x})$ and $f(\mathbf{x})$. Algorithm 2 shows how to use Newton's method with a backtracking line search to choose the step size of each iteration. The gradient and the Hessian of the logarithmic barrier function ϕ are given by:

$$\nabla\phi(\mathbf{x}) = \sum_{i=1}^{M}\frac{1}{(b_i-\mathbf{a}_i'\mathbf{x})}\mathbf{a}_i,\ \nabla^2\phi(\mathbf{x}) = \sum_{i=1}^{M}\frac{1}{(b_i-\mathbf{a}_i'\mathbf{x})^2}\mathbf{a}_i\mathbf{a}_i'$$

Algorithm 1: Barrier method

given: strictly feasible $\mathbf{x}$, $\varepsilon:=\varepsilon^0$, $0<\mu<1$, tolerance $e>0$
repeat
  compute $\mathbf{x}(\varepsilon)$ by minimizing $g(\mathbf{x})$, starting at $\mathbf{x}$;
  update $\mathbf{x}:=\mathbf{x}(\varepsilon)$; $\varepsilon:=\mu\varepsilon$;
until $\varepsilon\leq e$;

Algorithm 2: Newton's method

given: a starting point $\mathbf{x}\in\mathbf{dom}\ g$, tolerance $e>0$
repeat
  compute the Newton step: $\Delta\mathbf{x}_{nt}:=-\nabla^2 g(\mathbf{x})^{-1}\nabla g(\mathbf{x})$;
  choose step size
    given: a descent direction $\Delta\mathbf{x}_{nt}$, $\alpha\in(0,0.5)$, $\beta\in(0,1)$
    $t:=1$;
    while $g(\mathbf{x}+t\Delta\mathbf{x}_{nt})>g(\mathbf{x})+\alpha t\nabla g(\mathbf{x})'\Delta\mathbf{x}_{nt}$ **do**
      $t:=\beta t$;
  end
  update: $\mathbf{x}:=\mathbf{x}+t\Delta\mathbf{x}_{nt}$;
  compute the decrement: $\lambda^2:=\nabla g(\mathbf{x})'\nabla^2 g(\mathbf{x})^{-1}\nabla g(\mathbf{x})$;
until $\lambda^2/2\leq e$;

By approximating problem P given in (10) to the equivalent unconstrained form $P(\varepsilon)$, where the constraints on outputs and controls are put together in $\phi(\hat{\mathbf{u}})$, unconstrained minimization algorithms like those described in this section can be used to solve the problem.

3. DISTRIBUTED OPTIMIZATION AND CONTROL

This paper focuses now on the distributed solution of P, discussing how to perform a decomposition of the problem into a network of coupled sub-problems P_m that will be solved by a network of distributed agents (Camponogara and Talukdar, 2005, 2007), and the use of a distributed iterative algorithm to solve these sub-problems.

Each agent will compute a control vector $\hat{\mathbf{u}}_m$. So, for a perfect decomposition, each agent m must have all the information on problem P that depends on $\hat{\mathbf{u}}_m$. Before giving the decomposition, let us define some special sets:

- $O(m)=\{i: m\in I(i), i\neq m\}$ to represent the set of *output neighbors* of m;

- $C(m) = \{(i,j) \in I(m) \times I(m) : i = m \text{ or } j = m\}$ for the sub-system pairs of quadratic terms in the cost function of sub-system m that depend on $\hat{\mathbf{u}}_m$;
- $O(m,k) = \{(i,j) \in I(k) \times I(k) : i = m \text{ or } j = m\}$ for the pairs of quadratic terms in the cost function of sub-system k, $k \in O(m)$, that depend on $\hat{\mathbf{u}}_m$;
- $N(m) = (I(m) \cup O(m) \cup \{i : (i,j) \in O(m,k), k \in O(m)\}) - \{m\}$, which defines the *neighborhood* of agent m, including input and output *neighbors*;
- $\hat{\omega}_m = (\hat{\mathbf{u}}_i : i \in N(m))$, for the set of control signals of the neighbors of agent m;
- $\hat{\mathbf{z}}_m = (\hat{\mathbf{u}}_i : i \in \mathscr{M} - N(m) \cup \{m\})$, for the set of all control signals that are not in $\hat{\omega}_m$ and $\hat{\mathbf{u}}_m$.

The problem P from the view of agent m is defined as:

$$\begin{aligned} P_m: \min\ & \frac{1}{2} \sum_{(i,j) \in C(m)} \hat{\mathbf{u}}_i' H_{mij} \hat{\mathbf{u}}_j + \sum_{i \in I(m)} \mathbf{g}_{mmi} \hat{\mathbf{u}}_m + c_m \\ & + \frac{1}{2} \sum_{k \in O(m)} \sum_{(i,j) \in O(m,k)} \hat{\mathbf{u}}_i' H_{kij} \hat{\mathbf{u}}_j \qquad (14a) \\ & + \sum_{k \in O(m)} \sum_{(m,j) \in O(m,k)} \mathbf{g}_{kmj} \hat{\mathbf{u}}_m \\ \text{S.to}: & \\ & \hat{\mathbf{u}}_m^{\min} \le \hat{\mathbf{u}}_m \le \hat{\mathbf{u}}_m^{\max} \qquad (14b) \\ & \hat{\mathbf{y}}_i^{\min} \le \sum_{j \in I(i)} \left(\overline{CA}_{i,j} \mathbf{x}_j(k) + \overline{CB}_{i,j} \hat{\mathbf{u}}_j \right) \le \hat{\mathbf{y}}_i^{\max}, \qquad (14c) \\ & \text{for all } i \in O(m) \cup \{m\} \end{aligned}$$

It is possible to simplify the representation of problem P_m by grouping some terms as follows:

$$H_m = H_{mmm} + \sum_{k \in O(m)} H_{kmm} \qquad (15a)$$

$$\begin{aligned} \mathbf{g}_m = & \frac{1}{2} \sum_{(i,m) \in C(m): i \neq m} \left(H_{mim}' + H_{mmi} \right) \hat{\mathbf{u}}_i + \sum_{i \in I(m)} \mathbf{g}_{mmi} \\ & + \frac{1}{2} \sum_{k \in O(m)} \sum_{(i,m) \in O(m,k): i \neq m} \left(H_{kim}' + H_{kmi} \right) \hat{\mathbf{u}}_i \qquad (15b) \\ & + \sum_{k \in O(m)} \sum_{(m,j) \in O(m,k)} \mathbf{g}_{kmj} \hat{\mathbf{u}}_m \end{aligned}$$

Using the terms defined above, problem P_m is represented as:

$$P_m(\hat{\omega}_m): \min f_m(\hat{\mathbf{u}}_m) = \frac{1}{2} \hat{\mathbf{u}}_m' H_m \hat{\mathbf{u}}_m + \mathbf{g}_m' \hat{\mathbf{u}}_m + c_m \qquad (16a)$$

$$\begin{aligned} \text{S.to}: & \\ & \hat{\mathbf{u}}_m^{\min} \le \hat{\mathbf{u}}_m \le \hat{\mathbf{u}}_m^{\max} \qquad (16b) \\ & \hat{\mathbf{y}}_i^{\min} \le \sum_{j \in I(i)} \left(\overline{CA}_{i,j} \mathbf{x}_j(k) + \overline{CB}_{i,j} \hat{\mathbf{u}}_j \right) \le \hat{\mathbf{y}}_i^{\max}, \qquad (16c) \\ & \text{for all } i \in O(m) \cup \{m\} \end{aligned}$$

This quadratic form will be used by each agent m to compute the control signal $\hat{\mathbf{u}}_m$ in the distributed approach.

Some properties about the decomposition are:

- $P_m(\hat{\omega}_m)$ consists of problem P with all the objective terms and constraints that depend on $\hat{\mathbf{u}}_m$;
- each sub-problem $P_m(\hat{\omega}_m)$ is convex.

3.1 Distributed Algorithm

This section describes briefly the distributed algorithm for the agent network to reach a solution. Let $P_m(\varepsilon)$ be the centering problem for $P_m(\hat{\omega}_m)$ with a given ε:

$$P_m(\varepsilon): \min f_m(\hat{\mathbf{u}}_m) + \varepsilon \phi_m(\hat{\mathbf{u}}_m) \qquad (17)$$

where $\phi_m(\hat{\mathbf{u}}_m)$ is the logarithmic barrier function of the constraints given in (16). It is important to note that the problem to be solved by each agent is much simpler than the one used in the centralized formulation. So, the distributed solution must encompass a sequence of steps before the optimal control sequence is reached (de Oliveira and Camponogara, 2008).

First, define the vector $\hat{\mathbf{u}}^k = \left(\hat{\mathbf{u}}_1^k, \ldots, \hat{\mathbf{u}}_M^k\right)$ with the set of all control variables of P at iteration k. For a given ε, all agents have to *negotiate* to find a solution for each $P_m(\varepsilon)$ in the network. And this process is repeated for a decreasing sequence of $\varepsilon \to 0^+$. The convergence to a stationary solution is ensured by respecting two assumptions:

(1) *Synchronous Work:* if agent m revises its decisions at iteration k, then:
   (a) agent m uses $\hat{\omega}_m^k$ to produce an approximate solution of $P_m(\varepsilon)$;
   (b) all the neighbors of agent m keep their decisions at iteration k: $\hat{\mathbf{u}}_i^{k+1} = \hat{\mathbf{u}}_i^k$ for all $i \in N(m)$.

(2) *Continuous Work:* if $\hat{\mathbf{u}}^k$ is not a stationary point for problems $P_i(\varepsilon)$, $i \in \mathscr{M}$, then at least one agent m for which $\hat{\mathbf{u}}_m^k$ is not a stationary point for $P_m(\varepsilon)$ produces a new iterate $\hat{\mathbf{u}}_m^{k+1}$.

Condition (a) of Assumption 1 and Assumption 2 hold by arranging the agents to iterate repeatedly in a sequence $\langle S_1, \ldots, S_r \rangle$, where $S_i \subseteq \mathscr{M}$, $\cup_{i=1}^r S_i = \mathscr{M}$, and all distinct pairs $m, n \in S_i$ are non-neighbors for all i. The pseudo-code of the distributed barrier method for solving the problem network $\{P_m(\varepsilon)\}$ is given in Algorithm 3.

Algorithm 3: Distributed barrier method

given: strictly feasible $\hat{\mathbf{u}}^0 = (\hat{\mathbf{u}}_1^0, \ldots, \hat{\mathbf{u}}_M^0)$, $\varepsilon = \varepsilon^0 > 0$, $0 < \mu < 1$, tolerance $e > 0$, and a sequence $\langle S_1, \ldots, S_r \rangle$

repeat
- Define the initial group of decoupled agents, $i := 1$;
- Define a flag for the stationary test, $\delta := false$;
- **repeat**
  - each agent $m \in S_i$ receives $\hat{\omega}_m^k$ and computes $\hat{\mathbf{u}}_m^{k+1}$ by solving $P_m(\varepsilon)$ starting at $\hat{\mathbf{u}}_m^k$;
  - each agent j computes $\hat{\mathbf{u}}_j^{k+1} = \hat{\mathbf{u}}_j^k$ for $j \notin S_i$;
  - $\hat{\mathbf{u}}^{k+1} := (\hat{\mathbf{u}}_1^{k+1}, \ldots, \hat{\mathbf{u}}_M^{k+1})$;
  - **if** $\hat{\mathbf{u}}^{k+1}$ *is a stationary point for* $P_i(\varepsilon)$, $\forall i \in \mathscr{M}$ **then**
    - $\delta := true$;
  - **else**
    - $k := k + 1$;
    - $i := (i \mod r) + 1$;
- **until** $\delta = true$;
- $\varepsilon := \mu \varepsilon$;

until $\varepsilon < e$;

4. COMPUTATIONAL ANALYSIS

4.1 Distillation Column

This section presents the application of model predictive control to a benchmark problem, comparing the performance of the centralized and distributed approaches. The model of the heavy oil fractionator utilized is referred in the literature as the Shell Oil's heavy oil fractionator (Prett and Morari, 1987; Camacho and Bordons, 2004). It relates the controlled variables y_1, y_2, and y_3 that correspond to the top endpoint composition, side end composition, and bottom reflux temperature, respectively, with the manipulated variables u_1, u_2, and u_3, corresponding to top draw, side draw, and bottom reflux duties. The discrete model is obtained with a sampling time of 4 minutes, and it is possible to obtain a state space representation in the form:

$$\begin{bmatrix} \mathbf{x}_{11}(k+1) \\ \mathbf{x}_{21}(k+1) \\ \mathbf{x}_{31}(k+1) \\ \mathbf{x}_{12}(k+1) \\ \vdots \\ \mathbf{x}_{33}(k+1) \end{bmatrix} = \begin{bmatrix} A_{11} & & & \\ & A_{21} & \emptyset & \\ & \emptyset & \ddots & \\ & & & A_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{11}(k) \\ \mathbf{x}_{21}(k) \\ \vdots \\ \mathbf{x}_{33}(k) \end{bmatrix}$$

$$+ \begin{bmatrix} B_{11} & 0 & 0 \\ B_{21} & \vdots & \vdots \\ B_{31} & 0 & \\ 0 & B_{12} & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & B_{33} \end{bmatrix} \begin{bmatrix} u_1(k) \\ u_2(k) \\ u_3(k) \end{bmatrix}$$

$$\begin{bmatrix} y_1(k+1) \\ y_2(k+1) \\ y_3(k+1) \end{bmatrix} = \begin{bmatrix} C_{11} & 0 & 0 & C_{12} & \cdots & 0 \\ 0 & C_{21} & 0 & 0 & \ddots & \vdots \\ 0 & 0 & C_{31} & 0 & \cdots & C_{33} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{11}(k) \\ \mathbf{x}_{21}(k) \\ \vdots \\ \mathbf{x}_{33}(k) \end{bmatrix}$$

where each group $\mathbf{x}_{ij}$, A_{ij}, B_{ij}, and C_{ij} represents the couplings among the outputs and controls of the transfer function located in line i and column j of the transfer function matrix (Camacho and Bordons, 2004).

This formulation makes the distillation column a special case of the theory developed in the previous sections. For this special case, because the column is fully coupled, the set $I(m)$ is equal to $\{1,2,3\}$ for any m, and the *neighborhood* is equal to the *output* for each agent m. Other sets are given in Table 1.

Table 1. Sets used in the problem decomposition

m	$O(m)$	$C(m)$	$O(m,k)$
1	{2,3}	{(1,1),(1,2),(2,1),(1,3),(3,1)}	$O(1,2) = O(1,3) = C(1)$
2	{1,3}	{(1,2),(2,1),(2,2),(2,3),(3,2)}	$O(2,1) = O(2,3) = C(2)$
3	{1,2}	{(1,3),(3,1),(2,3),(3,2),(3,3)}	$O(3,1) = O(3,2) = C(3)$

Three different algorithms were used to solve the problem for the purpose of comparison:

- *Centralized Quadratic Programming* (cent_{QP}): the solution of P is obtained using a specific solver in Matlab® for problems in the quadratic form with constraints;
- *Centralized Barrier* (cent_{Br}): the solution of P is reached using the logarithmic barrier method and the centralized formulation;

Table 2. Numerical results

	dist_{Br}		cent_{Br}		cent_{QP}	
T	time	objective	time	objective	time	objective
1	0.0332	5.6188	0.0133	5.6188	0.0099	5.6188
2	0.0386	11.5680	0.0200	11.5680	0.0096	11.5680
5	0.0553	40.5596	0.0243	40.5596	0.0096	40.5596
10	0.0857	73.1869	0.0423	73.1868	0.0109	73.1868
15	0.1417	85.9801	0.0670	85.9797	0.0157	85.9797
20	0.2527	91.2292	0.1211	91.2284	0.0197	91.2284
25	0.3880	93.7129	0.2004	93.7115	0.0361	93.7115
30	0.6258	95.0680	0.2657	95.0660	0.0302	95.0660

- *Distributed Barrier* (dist_{Br}): each problem P_m is solved by a different agent and the constraints are treated with the logarithmic barrier method.

The criteria of convergence in Newton's method, which is used in cent_{Br} and dist_{Br} to solve centering problems, is $\lambda^2/2 \leq 10^{-5}$. The convergence criterion for the logarithmic barrier method is $\varepsilon \leq 10^{-4}$, while the stationary test is satisfied with $e \leq 10^{-4}$. The weights on control action and output deviation were set equal because their adjustment is not the focus of this work, but it is clear that the choice of weights affects directly the compromise between performance and robustness.

Ten different feasible start points were chosen at random in the experiments and the initial ε was defined as 10^3. The previous solution was not used as an initial approximation for reoptimization to induce worst-case scenarios. The analyses considered each type of algorithm, different lengths of prediction horizon (T), and rate of decrease $\mu \in \{0.05, 0.1, 0.3, 0.5\}$ for the interior-point methods.

4.2 Numerical Results

The objective function given in (7) is used for comparison as the cost of the computational experiments. Table 2 has the results of the accumulated cost obtained with the ten start points and the four different values for μ. The cost difference between dist_{Br} and the centralized approach is less than 3×10^{-3}, proving that solving the set of distributed problems, $\{P_m\}$, and solving the centralized problem, P, is equivalent. The gap in cost is due to the acceptance range of the convergence criteria.

Table 2 also contains the results about the time spent in the experiments, comparing cent_{QP}, cent_{Br} and dist_{Br}. A computer with an AMD Turion™ 64x2 1.60 GHz processor and 2048 MB of memory was used to perform the experiments. Time is given in seconds and represents the mean time spent to compute the control actions of forty different experiments with each algorithm. As expected, the greater the value of T, more time is necessary to reach the solution, but the time used was always less than one second, which is a very good result, considering real applications.

The iterations between agents to exchange information were also counted. Fig. 1 depicts the average number of iterations for the experiments with varying prediction horizon and decrease rate of the barrier method.

The convergence criterion can be relaxed to minimize the number of iterations between agents. In this case, due to the strong couplings among the variables, the information exchange is considerable. More scattered models will be used in future experiments, which are more suitable for a distributed approach.

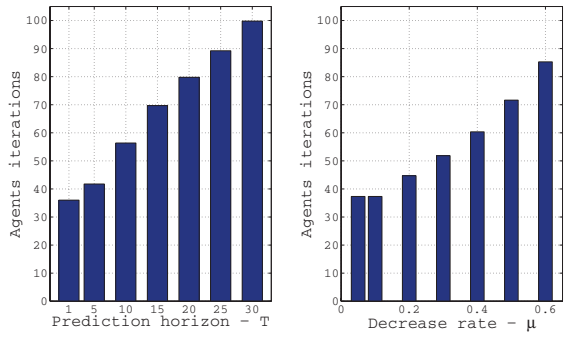


Fig. 1. Algorithm iterations varying T and μ, respectively.

Ending the analyses, it can be said that a network of distributed agents can solve the set of sub-problems $\{P_m\}$, rather than having a single agent solve P in centralized MPC, without incurring great loss of performance. Other adjustments can be made in the algorithm, such as the limit on the number of iterations, to guarantee the fulfillment of the deadlines, but all modifications have a compromise between speed and quality.

5. CONCLUSION

This work presented a distributed MPC framework with constraints on output and control-input signals. The methodology of problem decomposition was outlined and the barrier method was used to deal with the constraints, replacing the constrained minimization problem by a sequence of unconstrained minimization problems. Centralized MPC algorithms and the distributed MPC algorithm were applied to solve a regulation problem in a distillation column model.

The performance of distributed MPC in the distillation column scenario was comparable to the performance obtained with centralized MPC. The computational cost necessary to solve each agent problem was less than the centralized case. This advantage might allow the use of the distributed algorithm in machines with less computational resources.

It is worth emphasizing that distributed predictive control is more appropriate for multivariable problems where the couplings are scattered, which does not happen with the model of the distillation column.

Future works will focus on the implementation of reference tracking and state observers. As well as the whole study of how to introduce these extensions in a distributed algorithm. Another goal is to look into other scenarios that can be more appropriate for the application of distributed algorithms.

REFERENCES

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization*. Cambridge University Press.

Camacho, E.F. and Bordons, C. (2004). *Model Predictive Control*. Springer-Verlag, London.

Camponogara, E. and de Oliveira, L.B. (2009). Distributed optimization for model predictive control of linear dynamic networks. Accepted by *IEEE Transactions on Systems, Man, and Cybernetics – Part A*.

Camponogara, E., Jia, D., Krogh, B.H., and Talukdar, S.N. (2002). Distributed model predictive control. *IEEE Control Systems Magazine*, 22(1), 44–52.

Camponogara, E. and Talukdar, S.N. (2005). Designing communication networks to decompose network control problems. *INFORMS Journal on Computing*, 17(2), 207–223.

Camponogara, E. and Talukdar, S.N. (2007). Distributed model predictive control: synchronous and asynchronous computation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A*, 37(5), 732–745.

de Oliveira, L.B. and Camponogara, E. (2008). Multi-agent model predictive control for split signaling in urban traffic networks. In *Proceedings of the 5th Workshop on Agents in Traffic and Transportation*, 21–28. Estoril, Portugal.

Dunbar, W.B. (2007). Distributed receding horizon control of dynamically coupled nonlinear systems. *IEEE Transactions on Automatic Control*, 57(7), 1249–1263.

Giovanini, L. and Balderud, J. (2006). Game approach to distributed model predictive control. In *Proceedings of the International Control Conference*. Glasgow, UK.

Li, S., Zhang, Y., and Zhu, Q. (2005). Nash-optimization enhanced distributed model predictive control applied to the Shell benchmark problem. *Inf. Sci.*, 170(2-4), 329–349.

Mercangöz, M. and Doyle III, F.J. (2007). Distributed model predictive control of an experimental four-tank system. *Journal of Process Control*, 17(3), 297–308.

Motee, N. and Jadbabaie, A. (2006). Distributed receding horizon control of spatially invariant systems. In *Proceedings of the American Control Conference*, 731–736.

Prett, D.M. and Morari, M. (1987). *Shell Process Control Workshop*. Butterworths, Boston.

Venkat, A.N., Hiskens, I.A., Rawlings, J.B., and Wright, S.J. (2008). Distributed MPC strategies with application to power system automatic generation control. *IEEE Transactions on Control Systems Technology*, 16(6), 1192–1206.

Comparison of discrete and continuous-discrete observers for composition estimation in distillation columns

A. Aguilera-González *, **A.C. Téllez-Anguiano** *, **C.M. Astorga-Zaragoza** *, **M. Adam-Medina** *

* *Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira s/n, Col. Palmira, A.P. 5-164, C.P. 62490, Cuernavaca, Mor., México*

Abstract: In this paper we present a high-gain observer implemented in its discrete and continuous-discrete versions in order to compare the performance of both algorithms. The comparison is made considering the sampling time used to perform the observer's correction stage in order to establish that the continuous-discrete observer is the best option when a low sampling time is used. Under this condition the continuous discrete observer can process data performing a reliable on-line estimation of the system (a slow dynamics of the process is required). We apply both algorithms to a distillation column that uses the Ethanol-Water binary mixture.

Keywords: High-gain observer, discrete, continuous-discrete, distillation columns.

1. INTRODUCTION

The knowledge of state variables is often required in order to apply the advanced concepts of control and fault diagnosis to practical applications, specially in the chemical process industry. A method to obtain such variables, consists of combining a priori knowledge about physical systems with experimental data to provide an on-line estimator (observer).

The main control problems of distillation columns are caused due to tight interactions between the process variables, nonlinearities of the process, process and measurement delays and the large number of variables involved (see Murray-Gunther (2003)). For these reasons, a significant amount of effort has been devoted to develop algorithms that provide accurate parameter identification and state estimation (state observers) to reconstruct the product composition dynamics by secondary measurements (e.g., temperatures and flows) (see Quintero-Mármol et al. (1991), Deza et al. (1991)), and most recently in Bahar et al. (2006)Jana et al. (2006). The proposed observers are used to estimate unmeasured state variables from on-line and/or off-line measurements, see e.g. Bakir et al. (2005), Hammouri et al. (2006), Yildiz et al. (2005), Astorga et al. (2002) and Nadri et al. (2004).

Estimators are generally dynamic systems obtained from a nominal model by adding a correction term which is proportional to some output deviation. In other words, given a nominal model:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(\mathbf{t}) = \mathbf{h}(x(t)) \end{cases} \tag{1}$$

The state $x(t)$ belongs to an open subset $\mathbf{V}$ of $\mathbb{R}^n$, the input u(t) belongs to a Borelian subset $\mathbf{U}$ of $\mathbb{R}^m$ and the output $y(t) \in \mathbb{R}^p$. An observer for the system represented by Eq. (1) is generally a dynamic system of the form:

$$\begin{cases} \dot{\hat{x}}(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t)) + \mathbf{k}(\mathbf{t})[h(\hat{\mathbf{x}}(\mathbf{t})) - \mathbf{y}(\mathbf{t})] \\ \dot{\mathbf{r}} = \mathbf{F}(\mathbf{r}(\mathbf{t}), \mathbf{u}(\mathbf{t}), \mathbf{y}(\mathbf{t}), \hat{\mathbf{x}}(\mathbf{t})) \\ \mathbf{k}(\mathbf{t}) = \varphi(\mathbf{r}(\mathbf{t})) \end{cases} \tag{2}$$

$r(t)$ and $k(t)$ are called indifferently the gain of the observer (see Hammouri et al. (2002)). An interesting class of nonlinear systems consists of those systems which are observable for every input, called *uniformly observable systems*. For this class of nonlinear systems, we can design an observer whose gain does not depend on the inputs (see Bornard and Hammouri (1991) and Gauthier et al. (1992)). For such systems a canonical (called triangular) form is designed in order to develop an observer. To ensure mathematical convergence, a particular high-gain is required (see Hammouri et al. (2002)).

However, using a high-gain observer may generate the so-called peak phenomena (overshoot problem); moreover, the estimator becomes noise sensitive. Due to nonlinearity of the system, the choice of the gain which gives the best compromise between fast convergence, the noise rejection and the attenuation of the peak phenomena becomes a difficult task, and only simulations allow to determine a possible gain. This paper aims to present a high-gain observer status in its discrete and continuous-discrete versions. We apply this algorithms to a binary distillation column that uses the binary mixture Ethanol-Water.

2. THE BINARY DISTILLATION COLUMN MODEL

The binary distillation column model is derived from the binary distillation column scheme shown in Fig. 1. There are three principal stages considered for a distillation column (condenser, tray and boiler); for every stage, the balances of energy and material should be formulated as well as the equilibrium conditions of the mixture.

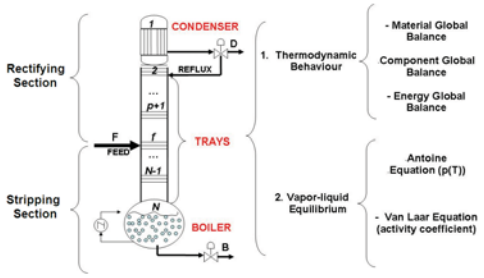


Fig. 1. Distillation column

2.1 General Aspects

The scheme of a typical distillation column is shown in Fig. 1. The binary feeding mixture (molar flow rate F) is introduced toward the middle of the column (the feeding tray). The distilled product (molar flow rate D), which mainly contains the light component, is removed from the top of the column. The bottom product (molar flow rate B), which contains the heavy component in greater concentration, is removed from the bottom of the column. Part of the overhead product is returned into the column to improve purity. Column stages are labeled with an ascendent numeration from the condenser to the boiler: $i = 1, ..., N$. The sections of the column are: the condenser, tray $i = 1$; the rectifying section, trays $i = 2, ..., f - 1$; the feeding tray, tray $i = f$; the stripping section, trays $i = f + 1, ..., N - 1$; the boiler, tray $i = N$.

2.2 Physic Behaviour

Due to its physical structure, a distillation column can be modeled as a set of interconnected stages using the mass balance and vapor-liquid equilibrium relation at every stage. The algebraic and differential equations of the model are formulated to calculate the light component composition of the mixture. The liquid phase and the vapor phase of the light component are designated by x and y, respectively.

Assumptions: The following assumptions, taken from Luyben (1992) and Halvorsen and Skogestad (2000) are considered in the model formulation:(**A1**) Constant pressure; (**A2**) Ideal Liquid-Vapor Equilibrium; (**A3**) Liquid-properties behave as a non-ideal mixture; (**A4**) Negligible molar vapor holdup compared to the molar liquid holdup; (**A5**) Boiler as a theoretical tray; (**A6**) Total condenser; (**A7**) Constant liquid volumetric hold up.

Vapor-liquid equilibrium: If a vapor and a liquid are in intimate contact for a long period of time, equilibrium is attained between the two phases. This concept of vapor-liquid equilibrium is fundamental to model distillation columns. If the vapor-liquid equilibrium exists, then the vapor composition y_i and the liquid composition x_i can be computed by means of correlating equations of the form $(y_i^{eq}, x_i^{eq}) = K_i(T_i, P_T)$, where T_i is the temperature, P_T is the total pressure, y_i^{eq} and x_i^{eq} are the vapor and liquid composition at the equilibrium phase respectively. The equilibrium constant K_i depends on the thermodynamical properties of the mixture.

The non-ideality of a binary mixture is due to different causes, the most frequent is the non-ideality of the liquid phase. In consequence, specially designed models are used to represent these non-idealities. For low pressure systems, the equation that represents the vapor composition is:

$$y_{i,j} P_T = P_{i,j}^{sat} x_{i,j} \gamma_{i,j} \tag{3}$$

where $j = 1$ if the component is ethanol and $j = 2$ if the component is water; $\gamma_{i,j}$ is the activity coefficient for every stage, it is a correction factor highly dependent on the concentration. In this work the vapor composition is calculated as a function of the light component. One method to determine this coefficient in every component of the mixture uses the Van Laar equation (see Perry (1999)).

Mass transfer effects: In order to deal with the mass transfer effects, the Murphree's efficiencies are introduced. The Murphree stage efficiency E_i is the ratio between the current change in vapor composition between two stages and the change that will occur if the vapor is in equilibrium with the liquid leaving the stage (see Murray-Gunther (2003)).

The molar flow rates: In the rectifying section the vapor molar flow V_R and the liquid molar flow L_R are :

$$\begin{cases} V_R = V_S + (1 - q_F)F, & i = 1, ..., f \\ L_R = (1 - R)V_R, & i = 1, ..., f - 1 \end{cases} \tag{4}$$

where

$$q_F = 1 + \frac{C_p(T_b - T_F)}{\lambda} \tag{5}$$

q_F describes the feeding condition. C_{p_j} is the specific heat, T_b is the boiling temperature, T_F is the feeding temperature and λ is the vaporization enthalpy for ethanol and water respectively.

F is the molar flow of the feeding stream:

$$F = F_V[\rho_1 w_1 + \rho_2(1 - w_1)] \left(\frac{x_f}{M_{W_1}} + \frac{1 - x_f}{M_{W_2}} \right) \tag{6}$$

where F_V is the volumetric flow of the feeding stream, ρ_j is the density of the component j, M_{W_j} is the molecular weight, w_1 is the weight fraction of the light component given by:

$$w_1 = \frac{x_f \rho_1}{x_f \rho_1 + (1 - x_f)\rho_2} \tag{7}$$

and x_f is the molar composition of the feeding stream given by:

$$x_f = \left(\frac{V_1 \rho_1 / M_{W_1}}{V_1 \rho_1 / M_{W_1} + V_2 \rho_2 / M_{W_2}} \right) \tag{8}$$

where V_j is the initial volume of the component j on the feeding container.

The distilled product flow rate, $D = 0$ if the three-way ON-OFF reflux valve shown in 1) is totally closed ($r_v = 0$) and $D = V_R$ if this valve is totally open ($r_v = 1$)

In the stripping section (subindex $[\cdot]_S$ is used), the vapor molar flow V_S and the liquid molar flow L_S are, respectively

$$\begin{array}{rcll} V_S & = & \dfrac{Q_b}{\Delta H_1^{vap} x_{1,N} + \Delta H_2^{vap}(1 - x_{1,N})}, & i = f+1, ..., N \\ L_S & = & L_R + q_F F, & i = f, ..., N \end{array} \tag{9}$$

where Q_b is the heating power on the boiler. Finally, the molar flow rate of the bottom product is:

$$B = (L_S - V_S) b_v$$

b_v is a binary variable representing the bottom-valve opening, *i.e.*: $b_v = 0$ if a batch distillation is performed and $b_v = 1$ if the bottom product is totally withdrawn from the boiler.

Molar hold-up: The molar hold-up for every stage must be determined from the distillation plant features and properties of the mixture. This quantity can be approximated as:

$$M_i = v_i \frac{1}{\dfrac{x_1 M_{W_1}}{\rho_1} + \dfrac{(1 - x_2) M_{W_2}}{\rho_2}} \tag{10}$$

2.3 The dynamic model

Tray and distilled product compositions are estimated by using the dynamic model based on material, component and energy balances. Taking into account the assumptions (**A1**) to (**A7**), a set of differential equations can be derived for the light component material balance as follows:

$$\begin{cases} \frac{dM_1}{dt} = V_2 - L_1 - D \\ \frac{dM_i}{dt} = V_{i+1} - L_i - V_i + L_{i-1} + \delta(i) F \\ \frac{dM_N}{dt} = L_{N-1} - V_N - B \end{cases} \tag{11}$$

for $i = 2, 3, ..., N-1$. M_i is the molar holdup of the boiler. The component balances for every stage are given by:

$$\begin{cases} \frac{d(M_1 x_1)}{dt} & = V_2 y_2 - L_1 x_1 - D x_1 \\ \frac{d(M_i x_i)}{dt} & = V_{i+1} y_{i+1} - L_i x_i - V_i y_i + L_{i-1} x_{i-1} + \delta(i) F x_F \\ \frac{d(M_N x_N)}{dt} & = L_{N-1} x_{N-1} - V_N y_N - B x_N \end{cases} \tag{12}$$

where $\delta(i) = 1$ if $i = f$ and $\delta(i) = 0$ if $i \neq f$

The enthalpies of the process are considered constants, therefore the energy balance is not taken into account to develop this model. The state model presented in the following section is based on this dynamic model for the distillation column.

2.4 The state model

The distillation column is a process that belongs to a class of multi-variable nonlinear systems. The process inputs are the heating power applied on the boiler, and the opening period of the reflux valve, this is $\mathbf{u}(t) = [Q_b(t)\ r_v(t)]^T$. F y x_F are considered perturbations in the system, this is $\mathbf{d} = [F, x_F, b_v]$. A state representation can be obtained from Eqs. (11) and (12). Additionally, the nonlinear model has the following triangular form:

$$\begin{cases} \dot{\zeta}_1 = f_1(\zeta_1, \zeta_2, \mathbf{u}) \\ \dot{\zeta}_i = f_i(\zeta_1, ..., \zeta_i, \zeta_{i+1}, \mathbf{u}); \quad (i = 2, ..., f-2) \\ \dot{\zeta}_{f-1} = f_{f-1}(\zeta_1, ..., \zeta_{f-1}, \zeta_f, \mathbf{u}) \\ \dot{\zeta}_f = f_f(\zeta_{f-1}, ..., \zeta_N, \mathbf{u}, \mathbf{d}) \\ \dot{\zeta}_i = f_i(\zeta_{i-1}, ..., \zeta_N, \mathbf{u}, \mathbf{d}); \quad (i = f+1, ..., N-1) \\ \dot{\zeta}_N = f_N(\zeta_{N-1}, \zeta_N, \mathbf{u}, \mathbf{d}) \end{cases} \tag{13}$$

where, ζ represent the states of the process (the liquid compositions of the light component). Subindex f represents the feeding tray number in Eqs. (11) and (12). The model allows to calculate the flows B, D, V_i, L_i and T_i, x_i from inputs Q_b, r_v, b_v, F_V, x_F.

3. OBSERVER DESIGN FOR A CLASS OF NONLINEAR TRIANGULAR SYSTEMS

A special class of nonlinear systems consists of those which are observable for every input, called uniformly observable systems (see Hammouri et al. (2002)). For this class of nonlinear systems, an observer whose gain does not depend on the inputs can be designed. For such systems a canonical (triangular) form is used in order to design an observer. Due to the nonlinearity of the system, it is important to select the gain which gives the best compromise between fast convergence and accuracy. Consider the follow triangular system that can be rewritten in a compact form:

$$\begin{cases} \dot{\zeta}^1 = \mathbf{f^1}(\boldsymbol{\zeta}(t), \mathbf{u}(t)) \\ \dot{\zeta}^2 = \mathbf{f}^2(\boldsymbol{\zeta}(t), \mathbf{u}(t), \mathbf{d}(t), \varepsilon(t)) \\ \boldsymbol{\varrho(t)} = (\varrho_1(t), \varrho_2(t))^T = (\mathbf{C}_{n_1} \boldsymbol{\zeta^1}(t), \mathbf{C}_{n_2} \boldsymbol{\zeta^2}(t))^T \end{cases} \tag{14}$$

where the states $\boldsymbol{\zeta}(t) = \left[\boldsymbol{\zeta}^1(t), \boldsymbol{\zeta}^2(t)\right]^T \in \mathbb{R}^n$ and $n = n_1 + n_2$; $\zeta^j = \left[\zeta_1^j, \zeta_2^j, \ldots, \zeta_{n_j}^j\right]^T \in \mathbb{R}^{n_j}$ for $j = 1, 2$; $y_j = \mathbf{C}_{n_j} \zeta^j = \zeta_1^j$ the first component of ζ^j; $\mathbf{C}_{n_j} = [1, 0, \ldots, 0]$; the input $\mathbf{u} \in \mathbb{R}^m$, and $\varepsilon(t)$ is an unknown and bounded function. The following assumptions are considered in order to design the observer:

- (**A8**) $\mathbf{f}^j$ is globally Lipschitz *w.r.t.* ζ.
- (**A9**) The state variables $\boldsymbol{\zeta}(t)$ are bounded

Considering the following notations:

i) $\mathbf{C}_{n_j} = [1, \ldots, 0] \in \mathbb{R}^{n_j}$ where n_j is the size for every state vector ζ^j.
ii)

$$\mathbf{A}_{n_j}(t) = \begin{bmatrix} 0 & a_1(t) & 0 & 0 \\ \vdots & & a_2(t) & 0 \\ & & & \ddots \\ 0 & & & a_{n_j - 1}(t) \\ 0 & \ldots & 0 & 0 \end{bmatrix},$$

where $a_k(t)$, $k = 1, \ldots, n_{j-1}$ are bounded and unknown functions satisfying the following assumption:

- (**A10**) There are two finite real numbers α, β with $\alpha > 0$, $\beta > 0$ such that $\alpha \leq a_k(t) \leq \beta$.

Lemma 1 *Under assumptions (**A8**) and (**A10**) exist a symmetric positive definite (S.P.D.) matrix* $\mathbf{S}_{n_j}$ *and a constant* $\mu > 0$ *s.t.:*

$$\forall t, \mathbf{S}_{n_j}\mathbf{A}_{n_j}(t) + \mathbf{A}_{n_j}^T(t)\mathbf{S}_{n_j} \leq -\mu\mathbf{I}_d \tag{15}$$

where $\mathbf{I}_d$ *is the identity matrix.*

Then,

$$\mathbf{S}_{n_j} = \begin{bmatrix} s_{11} & s_{12} & 0 & & & 0 \\ s_{12} & s_{22} & \ddots & & & \vdots \\ 0 & \ddots & & \ddots & & 0 \\ \vdots & & & \ddots & & s_{(n_j-1)n_j} \\ 0 & \ldots & 0 & s_{(n_j-1)n_j} & & s_{n_j n_j} \end{bmatrix},$$

Assume that the system given in Eq. (14) satisfies hypothesis (**A8**) to (**A10**). Then the observer:

$$\begin{cases} \dot{\hat{\boldsymbol{\zeta}}}^1 = \mathbf{f}^1(\hat{\boldsymbol{\zeta}}, \mathbf{u}) - r_1\boldsymbol{\Delta}_{\theta^{\delta_1}}\mathbf{S}_{n_1}^{-1}\mathbf{C}_{n_1}^T(\mathbf{C}_{n_1}\hat{\boldsymbol{\zeta}}^1 - \varrho_1) \\ \dot{\hat{\boldsymbol{\zeta}}}^2 = \mathbf{f}^2(\hat{\boldsymbol{\zeta}}, \mathbf{u}, \mathbf{d}) - r_2\boldsymbol{\Delta}_{\theta^{\delta_2}}\mathbf{S}_{n_2}^{-1}\mathbf{C}_{n_2}^T(\mathbf{C}_{n_2}\hat{\boldsymbol{\zeta}}^2 - \varrho_2) \end{cases} \tag{16}$$

is an estimator for the system given in Eq. (14), where $r_1 > 0$, $r_2 > 0$; $\theta > 0$; $\boldsymbol{\Delta}_{\theta^{\delta_j}} = diag(\theta^{\delta_j}, \theta^{2\delta_j}, ..., \theta^{n_j\delta_j})$; $\delta_1 > 0$, $\delta_2 > 0$; $\mathbf{S}_{n_1}$ is given by **Lemma 1**.

The following theorem is given:

Theorem 1: *Denote by* ε *the upper bound of* $|\varepsilon(t)|$ *i.e.* $\varepsilon = \sup_{t>0}|\varepsilon(t)|$, *then for* $r_1 > 0$, $r_2 > 0$, $\theta > 0$ *sufficiently large and* $\forall\delta_1 > 0$, $\delta_2 > 0$ *s.t.*

$$\frac{2n_1 - 1}{2n_2 + 1}\delta_1 < \delta_2 < \frac{2n_1 + 1}{2n_2 - 1}\delta_1; \tag{17}$$

$$|\hat{\zeta}(t) - \zeta(t)\| \leq \lambda e^{-\mu t} + \lambda'\varepsilon; \tag{18}$$

for some constants $\lambda > 0$, $\mu > 0$ *and* $\lambda' > 0$. *Moreover,* $\mu \to +\infty$ *as* $\theta \to 0$.

Remark 1. *If* $\varepsilon = 0$, *the system given in Eq. (16) becomes an exponential observer for the system given in Eq. (14). A proof of this result is given in Hammouri et al. (2002).*

3.1 Application of the designed observer to a distillation column

In the previous sections Eqs. (14) to (16) describe the high-gain observer designed for a distillation column, as well as the appropriate model in which this observed is based. In this section, an observer synthesis to the following class of nonlinear systems is developed, which contains the model of binary distillation columns, considering the following notations:

$$\begin{cases} \zeta_i = \zeta_i^1; \ \ 1 \leq i \leq f-1 \\ \zeta_{N-i+1} = \zeta_i^2; \ \ 1 \leq i \leq N-f+1 \\ \zeta_F = \zeta_{N-f+2}^2 \end{cases} \tag{19}$$

then, the system given in Eq. (13) can be represented in the following compact form:

$$\begin{cases} \dot{\zeta}^1(t) = \mathbf{f}^1(\boldsymbol{\zeta}(t), \mathbf{u}(t), D(t)) \\ \dot{\zeta}^2(t) = \mathbf{f}^2(\boldsymbol{\zeta}(t), \mathbf{u}(t), \mathbf{d}(t), \varepsilon(t), B(t)) \\ \boldsymbol{\varrho}(t) = (\zeta_1^1, \zeta_1^2)^T = (\zeta_1, \zeta_N)^T \end{cases} \tag{20}$$

where $\varepsilon(t)$ is a bounded and unknown function *s.t.* $\dot{\zeta}_F = \varepsilon(t)$. The liquid flow rates in the stripping section and vapor flow rates in the rectifying section are not known variables. The holdup in the condenser, the boiler and trays (where it is assumed to be constant) is calculated . The weighted error between the estimated and the measured compositions is feed back to the model equation in order to correct the liquid flow rates in the stripping section and the vapor flow rates in the rectifying section (see Halvorsen and Skogestad (2000)).

The hypothesis (**A9**) and (**A10**) must be verified:

- (**A9**) is fulfilled since the flow rates are physically bounded.
- (**A10**) is satisfied because the liquid compositions $x_i \in [0, 1]$

Applying **Theorem 1**, the high-gain observer for the distillation column is:

$$\begin{cases} \dot{\hat{\boldsymbol{\zeta}}}^1 = \mathbf{f}^1(\hat{\boldsymbol{\zeta}}, \mathbf{u}, D(t)) - \mathbf{Q}_{1\theta}(\mathbf{C}_{n_1}\hat{\boldsymbol{\zeta}}^1 - \varrho_1) \\ \dot{\hat{\boldsymbol{\zeta}}}^2 = \mathbf{f}^2(\hat{\boldsymbol{\zeta}}, \mathbf{u}, B(t), \mathbf{d}, \varepsilon(t)) - \mathbf{Q}_{2\theta}(\mathbf{C}_{n_2}\hat{\boldsymbol{\zeta}}^2 - \varrho_2) \end{cases} \tag{21}$$

where $\mathbf{Q}_{j\theta} = r_j\boldsymbol{\Delta}_{\theta^{\delta_j}}\mathbf{S}_{n_j}^{-1}\mathbf{C}_{n_j}^T$, for $j = 1, 2$. The constant parameters are the same described in Eq. (16).

4. EXTENSION OF THE HIGH-GAIN OBSERVER TO THE CONTINUO-DISCRETE CASE

There are processes where the measurement of their variables are performed using long sampling times due to their slow dynamics. In Bahar et al. (2006) it is demonstrated that certain restrictions in the dimension of the sampling period used by a purely discrete observer exist. An alternative of solution for this problem is the use of continuous-discrete observers (see Hammouri et al. (2002)).

Consider a non-linear uniformly observable systems of the form:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)) \\ \mathbf{y}(\mathbf{t}) = \mathbf{h}(x(t)) \end{cases} \tag{22}$$

where $x(t) \in \mathbb{R}^n$, $u = (u_1, ..., u_n) \in \mathbb{R}^m$, are measurable inputs and $y \in \mathbb{R}$ is a measurable output. Using the model of the system with $u(t), y(t)$ as known measurements it is possible to estimate the state line in $x(t)$ of the system represented by Ec. 22, this task is performed by a recursive algorithm with the following structure:

i. A prediction period in the time interval $t \in [t_k, t_{k+1}]$*:*

$$\dot{\hat{x}}_k(t) = \mathbf{f}(\hat{\mathbf{x}}(t), \mathbf{u}(t)) \tag{23}$$

ii. A correction period in the time $t = t_{k+1}$*:*

$$\dot{\hat{x}}_{k+1}(t) = \hat{\mathbf{x}}_{\mathbf{k+1}}(-) - r\Delta_\theta S_\theta^{-1}C^T(C\hat{x}_{k+1}(-) - y_{k+1}) \tag{24}$$

To make the extension of the high-gain observer to continuous-discrete case, it is assumed that the observations are made at the time $k\triangle t$, where $\triangle t$ is the time between measurements and k is the instant in which the sample is taken. In this case, it is not considered a coordinates change because the triangular structure of the model studied in section 3 is used.

As the observer gain is constant, it is true that:

$$\forall t > 0, \mathbf{A}_k^T(t)\mathbf{S}_k + \mathbf{S}_k\mathbf{A}_k(t) - \rho\mathbf{C}_k^T\mathbf{C}_k \leq\leq -\mu\mathbf{I}_k \quad (25)$$

where S_k is a symmetric positive definite matrix with the following structure:

$$\mathbf{S}_k = \begin{bmatrix} s_{11} & s_{12} & 0 & & 0 \\ s_{12} & s_{22} & \ddots & & \vdots \\ 0 & \ddots & & \ddots & 0 \\ \vdots & & \ddots & & s_{(k-1)k} \\ 0 & \dots & 0 & s_{(k-1)k} & s_{kk} \end{bmatrix},$$

C_k is denoted as a vector of k elements:

$$\mathbf{C}_{n_j} = [1, \dots, 0] \quad (26)$$

and A_k is given by:

$$\mathbf{A}_k(t) = \begin{bmatrix} 0 & a_1(t) & 0 & 0 \\ \vdots & & a_2(t) & 0 \\ & & \ddots & \\ 0 & & & a_{k-1}(t) \\ 0 & \dots & 0 & 0 \end{bmatrix},$$

where the terms a_k may be unknown and satisfy the hypothesis **A9**.

5. EXPERIMENTAL VALIDATION OF THE OBSERVER

The distillation pilot plant, located at the Process Control Laboratory of the National Center of Technological Research and Development (CENIDET) in Cuernavaca, Morelos, México, was used to carry out the required experiments. It has twelve trays, where temperature measurements are available through 8 RTD's Pt-100 located at trays 1, 2, 4, 6, 7, 9, 11 and 12. Using these measurements and considering the equilibrium relation (see Section 2.2.2), the respective liquid compositions can be obtained.

The mixture used in these experiments was Ethanol(EtOH)-Water(H_2O) which is considered as a non-ideal mixture. The experimental validation of the observers is done considering: EtOH volume of 2000 ml, H_2O volume of 2000 ml and process total pressure of 105.86 kPa. The specifications of every component of the mixture can be found in Perry (1999). The experiment lasts 82 minutes, once it has reached the stable state. In minute 27 the system goes from total reflux to partial reflux. In minute 54, a change in the input Q_b is applied.

In the discrete observer the sampling time is used to estimate and correct. In the continuous-discrete observer the prediction and correction times can be different in order to use less data, therefore less processing time, performing a reliable online estimation. The observer estimates the liquid composition of the light component (EtOH) for every stage by having the temperature measurements on stage 1 (condenser) and stage 12 (boiler) only.

The high-gain observer is obtained by fixing $r_1 = r_2 = 25$; $\theta = 0.09$ and satisfying (16) with $\delta_1 = 1.2$, $\delta_2 = \frac{1}{2}\left[1 + \left(\frac{2n_1+1}{2n_2-1}\right)^2\right]\delta_1 = 0.0983$(where $n_1 = f - 1 = 6, n_2 = n - f + 2 = 6$). Finally Lemma 1 gives:

$$\mathbf{S}_{n_1} = \mathbf{S}_{n_2} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1.5 & 0 & 0 & 0 \\ 0 & -1.5 & 4 & -2 & 0 & 0 \\ 0 & 0 & -2 & 8 & -3 & 0 \\ 0 & 0 & 0 & -3 & 10.5 & -4 \\ 0 & 0 & 0 & 0 & -4 & 15.5 \end{bmatrix}$$

Fig. 2 shows a comparison between the experimental dat and the estimation performed by the discrete observer in plate 12 (boiler) using a sampling time of $3s$. If the sampling time is slightly increased to $5.4s$ the discrete observer can not perform an adequate estimation, as can be seen in Fig. 3.

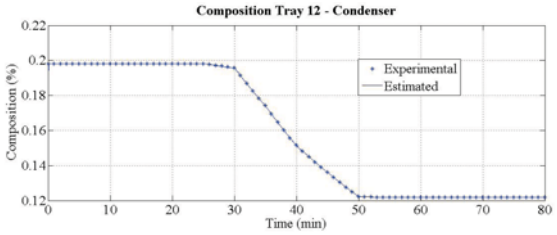


Fig. 2. Composition estimation of tray 12 by the discrete observer using a sampling time of 3 sec

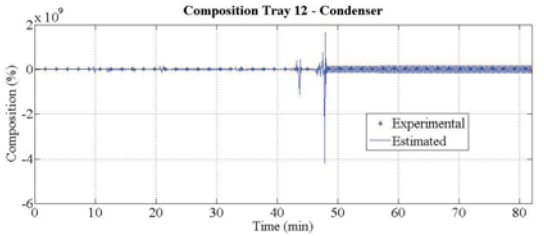


Fig. 3. Composition estimation of tray 12 by the discrete observer using a sampling time of 5.4 sec

In the continuous-discrete observer, a fixed prediction time of $3s$ is used, but different correction times are used in order to validate its performance. Figs. 4 to 6 show the experimental and estimated data when the correction time is 15 seconds, 30 seconds and 1 minute, respectively. In these figures it can be seen the good tracking and quickly convergence of the observer to the experimental data. The observer estimates the compositions of the plant adequately, under different conditions of correction time, having a maximum error of 0.03 and a minimum error 0.0001 between the estimated and the experimental data (the Euclidean norm was used to estimate the error).

6. CONCLUSIONS

In order to validate the performance of the high-gain observer versions: discrete and continuous-discrete, some experiments were conducted under similar conditions. First,

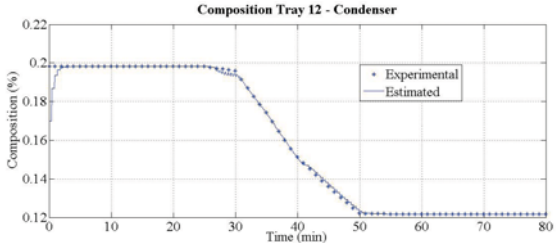


Fig. 4. Composition estimation of tray 12 by the continuous-discrete observer using a correction time of 15 sec

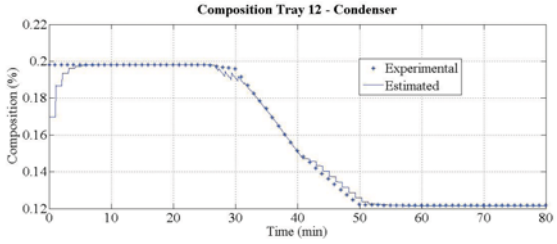


Fig. 5. Composition estimation of tray 12 by the continuous-discrete observer using a correction time of 30 sec

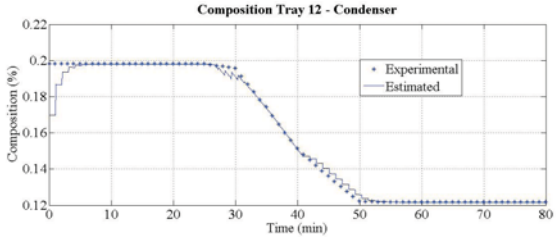


Fig. 6. Composition estimation of tray 12 by the continuous-discrete observer using a correction time of 60 sec

the purely discrete observer was validated, in order to perform, later, an adequate comparison with the continuous-discrete observer and analyze their response. Both validation use the same component specifications and same experimental inputs.

As can be seen in the presented figures the continuous-discrete observer presents a good tracking and quickly convergence to the experimental data, in spite of the sampling time used in the correction stage unlike the purely discrete observer, where the sampling time affects considerably its performance. Therefore, it can be assumed that the continuous-discrete observer is a suitable option to estimate the desired variables when the measurements of the system are performed using a long sampling time due to the slow dynamics of the process, which is the case of a distillation column where the compositions of the light component in a binary mixture of Ethanol-Water are estimated accurately.

REFERENCES

Astorga, C.M., Othman, N., Othman, S., Hammouri, H., and McKenna, T. (2002). Nonlinear continuous-discrete observers: application to emulsion polymerization reactors. *Control Engineering Practice*, 10(1), 3–13.

Bahar, A., Guner, E., Ozgen, C., and Halici, U. (2006). Design of state estimators for the inferential control of an industrial distillation column. In IEEE (ed.), *International Joint Conference on Neural Networks*. Vancouver.

Bakir, T., Othman, S., Puel, F., and Hammouri, H. (2005). Continuous-discrete observer for crystal size distribution of batch crystallization process. In IEEE (ed.), *Proceedings of the 44th IEEE Conference on Decision and Control, and the European Control Conference*, 6240–6244. Seville, Spain.

Bornard, G. and Hammouri, H. (1991). A high gain observer for a class of uniformly observable sytems. In *Proc. of the 30rd IEEE Conf. on Decision and Control*, 130–140. Brighton, UK.

Deza, F., Busvelle, E., and Gauthier, J.P. (1991). Exponentielly converging observers and internal stability using dynamic output feedback for distillation columns. *Chem. Eng. Science*, 47, 3935–3941.

Gauthier, J.P., Hammouri, H., and Othman, S. (1992). A simple observer for nonlinear systems, applications to bioreactors. *IEEE Transactions on Automatic Control*, vol. 37(6), pp. 875–880.

Halvorsen, I.J. and Skogestad, S. (2000). Distillation theory. *Encyclopedy of Separation Science*.

Hammouri, H., Nadri, M., and Mota, R. (2006). Constant gain observer for continuous-discrete time uniformly observable systems. In IEEE (ed.), *Proceedings of the 45th IEEE Conference on Decision and Control*, 5406–5411. San Diego, CA, USA.

Hammouri, H., Targui, B., and Armanet, F. (2002). High gain observer based on a triangular structure. *International Journal of Robust and Nonlinear Control*, 12, 497–518.

Jana, A., Samantab, A.N., and Gangulyb, S. (2006). Observer-based control algorithms for a distillation column. *Chemical Engineering Science*, 61, 4071 4085.

Luenberger, D. (1971). An introduction to observers. *IEEE Transactions on Automatic Control*, vol. 16(6), pp. 596–602.

Luyben, W.L. (1992). *Practical Distillation Control*. Van Nostrand Reinhold, New York, USA.

Murray-Gunther, A. (2003). *Dynamic Mathematical Model of a Distillation Column*. Ph.D. thesis, University of Tenessee at Chattanooga.

Nadri, M., Hammouri, H., and Astorga, C. (2004). Observer design for continuous-discrete time state affine systems up to output injection. *European Journal of Control*, 10(3), .

Perry, R. (1999). *Perrys chemical engineers handbook*. 7 edition.

Quintero-Mármol, E., Luyben, W.L., and Georgakis, C. (1991). Application of an extended luenberger observer to the control of multicomponent batch distillation. *Ind. Eng. Chem. Res.*, 30, 1870–1880.

Yildiz, U., Gurkan, U., Ozgen, C., and Leblebicioglu, K. (2005). State estimator design for multicomponent batch distillation columns. *Chemical Engineering Research and Design*, 83, 433–444.

Composition estimation of a six-component distillation column with temperature measurements

Andrea Frau, * **Roberto Baratti,** * **Jesús Alvarez** **

* *Dipartimento di Ingegneria Chimica e Materiali, Università degli Studi di Cagliari, Piazza d'Armi, 09123 Cagliari*
** *Universidad Autónoma Metropolitana-Iztapalapa, Depto. De Ingegneria de Procesos e Hidraúlica, Apdo. 55534, 09340 México D.F. México*

Abstract: The problem of jointly designing the estimation structure and algorithm to infer all or some composition in a six-component distillation column with temperature measurements is addressed. The structure design involves the choices of: (i) modeled and unmodeled compositions, (ii) the number of measurements and their location, and (iii) the innovated-noninnovated state partition. The algorithm is the dynamic data processor that performs the estimation task. The application of the geometric estimation approach (GE), in the light of the column characteristics, yields a tractable procedure to draw the solution of the estimation structure-algorithm design problem, with an estimation scheme that is considerably simpler than previous ones with extended Kalman Filter (EKF). The proposed methodology is applied to a representative six-component case example through simulations, finding that the estimation task can be performed with a three-component reduced model.

Keywords: Distillation column, nonlinear, estimators, geometric estimators, multicomponent.

1. INTRODUCTION

Distillation is an important energy-intensive industrial operation where many substances are separated and purified. The development of estimation schemes with temperature measurements for multicomponent distillation columns is motivated by: (i) the need of developing monitoring systems and (supervisory, advisory, and feedback) controllers with applicability-oriented requirements on reliability, tractability and maintenance cost, and (ii) the availability of reasonable and reliable multicomponent distillation models, including commercial packages.

The extended Kalman Filter (EKF) (Jazwinsky (1970)) has been by far the most widely used estimation technique in chemical process systems engineering in general, and in multicomponent distillation columns in particular, with successful simulations and experimental implementations for continuous and batch column operations, mostly for binary (Baratti et al. (1995); Yang and Lee (1997)) and ternary systems (Baratti et al. (1998)), and up to four-component systems (Venkateswarlu and Kumar (2006)). Even though adequate multicomponent column models are available, the related EKF construction and implementation for multicomponent systems still rises reliability, complexity, and development-maintenance cost concerns among industrial practitioners, because: (i) the dimensionality of the EKF grows rapidly with the number of stages and components ($\frac{n(n+1)}{2}+n$), (ii) the tuning of the covariance gains of the Riccati equation is a rather complex task, and (iii) due to the lack of formal connections between estimator tuning and functioning, the implementation of the EKF requires extensive testing.

Recently, the geometric estimation (GE) approach (Alvarez (2000)), which does not require the on-line integration of Riccati equations, has been redesigned with the following features (Alvarez and Fernandez (2008)): (i) the obstacle of high order Lie derivations has been removed and replaced by Jacobian model-based gain computations, enabling the consideration of staged systems, (ii) the equivalence between the GE and the EKF has been identified, (iii) the estimation model, its (innovation-noninnovation) detectability structure, and the sensor locations are regarded as structural design degrees of freedom, and (iv) a simple tuning scheme is coupled with a robust convergence criterion. The general-purpose GE approach has been successfully tested with experimental binary columns with complete (Tronci et al. (2005); Fernandez and Alvarez (2007)) or reduced models (Alvarez and Fernandez (2008)), and ternary columns (Pulis et al. (2006)), yielding estimation schemes which are considerably simpler than the EKF-based ones. These considerations motivate the present study on the six-component distillation column problem.

In this work, the problem of simultaneously designing the estimation structure and algorithm to infer all or some composition in a six-component distillation column with temperature measurements is addressed. By structure design we mean the choices of: (i) modeled and unmodeled compositions, (ii) the number of measurements and their locations, and (iii) the innovated-noninnovated composi-

tion state partition which, in conjunction with the model-sensor choice, determines the data assimilation versus error propagation mechanism. By algorithm it is meant the dynamic data processor that performs the estimation task, according to the estimation structure and a suitable gain tuning scheme. The application of afore discussed GE approach, in the light of the six-component column characteristics, yields a tractable procedure to solve the structure-algorithm problem, with an estimation scheme that is considerably simpler than previous EKF-based ones. The proposed methodology is applied to a representative six-component case example through simulations, finding that the estimation task can be performed with a three-component reduced model, a single-stage innovation with passive structure, and without the need of online integrating Riccati equations. The study can be seen as an inductive step towards the consideration of columns with more than six components.

2. ESTIMATION PROBLEM

2.1 Six-component distillation column

Consider a continuous multicomponent column with N stages and C components. Under standard assumptions (energy balance neglected on each tray, constant vapor and liquid flows, holdup dynamics neglected, tight reboiler and condenser level control, and stage liquid-vapor equilibrium), the column dynamics are described by the following set of nonlinear differential equations (Skogestad (1997); Baratti et al. (1998)):

Reboiler ($i = 1$, $j = 1, \dots, C-1$)

$$\dot{c}_1^j = \frac{(R+F)c_2^j - V\epsilon_j(c_1, P_1) - Bc_1^j}{M_1} = f_1^j(c_1, c_2^j) \quad (1a)$$

Stripping section ($2 \le i \le N_F - 1$, $j = 1, \dots, C-1$)

$$\dot{c}_i^j = \frac{(R+F)(c_{i+1}^j - c_i^j) - V(\epsilon_j(c_i, P_i) - \epsilon_j(c_{i-1}, P_{i-1}))}{M_i} \quad (1b)$$

$$= f_i^j(c_{i-1}, c_i, c_{i+1}^j)$$

Feed tray ($i = N_F$, $j = 1, \dots, C-1$)

$$\dot{c}_{N_F}^j = \frac{R(c_{N_F+1}^j - c_{N_F}^j) + F(c_F^j - c_{N_F}^j)}{M_{N_F}} - \frac{V(\epsilon_j(c_{N_F}, P_{N_F}) - \epsilon_j(c_{N_F-1}, P_{N_F-1}))}{M_{N_F}} \quad (1c)$$

$$= f_{N_F}^j(c_{N_F-1}, c_{N_F}, c_{N_F+1}^j, c_F)$$

Enriching section ($N_F + 1 \le i \le N - 1$, $j = 1, \dots, C-1$)

$$\dot{c}_i^j = \frac{R(c_{i+1}^j - c_i^j) - V(\epsilon_j(c_i, P_i) - \epsilon_j(c_{i-1}, P_{i-1}))}{M_i} \quad (1d)$$

$$= f_i^j(c_{i-1}, c_i, c_{i+1}^j)$$

Condenser ($i = N$, $j = 1, \dots, C-1$)

$$\dot{c}_N^j = \frac{V\epsilon_j(c_{N-1}, P_{N-1}) - Rc_N^j - Dc_N^j}{M_N} = f_N^j(c_{N-1}, c_N^j) \quad (1e)$$

Temperature measurements ($i = 1, \dots, m$)

$$T_{s_i} = \beta(c_{s_i}, P_{s_i}) \quad (1f)$$

where m is the number of sensors along the column and s_i is the location of the i-th sensor, c_i^j is the composition of the component j at i-th stage, $c_i = [c_i^1 \ \dots \ c_i^{C-1}]^T$ is the composition vector at i-th stage, T_{s_i} and P_{s_i} are respectively the temperature and the pressure at s_i-th stage, F is the feed flow rate with composition $c_F = [c_F^1 \ \dots \ c_F^{C-1}]^T$, D, B, R, and V are respectively the distillate, bottom, reflux and vapor flow rate (V is proportional to reboiler duty Q through the heat of vaporization λ), M_i is the holdup at i-th stage, ϵ_j is the liquid-vapor equilibrium function that determines the vapor composition of the component j, and β is the bubble-point implicit function that sets the temperature. The components c_i^C are determined by the (mass conservation) condition $\sum_{j=1}^{C} c_i^j = 1$, where $i = 1, \dots, N$. Henceforth, column system (1), will be referred to as the *complete six-component column system*, which in compact vector notation is written as follows:

$$\dot{x}_P = f_P(x_P, u_P, d_P) \qquad y_P = h_P(x_P) \quad (2)$$

where $x_P = [{c_1}^T \dots {c_N}^T]^T$, $c_i = [c_i^{C2} \, c_i^{C3} \, c_i^{IC4} \, c_i^{NC4} \, c_i^{IC5}]^T$, $u_P = [R \, V]^T$, $d_P = [F \, {c_F}^T]^T$, and $y_P = [T_{s_1} \ \dots \ T_{s_m}]^T$ are respectively the states, the inputs, the disturbances, and the outputs. The disturbances d_P are assumed to be constant and known.

2.2 Estimation problem

The estimation problem consists in jointly designing the estimation structure (i.e. estimator model, sensor location, innovated states and data assimilation mechanism), and the estimation algorithm (i.e., the dynamic data processor), to infer some of or all the effluent compositions of the six-component distillation column (2) on the basis of a reduced model (to be designed) in conjunction with temperature measurements, according to a specific estimation objective. In virtue of the general-purpose GE approach (Alvarez and Fernandez (2008)) and its applications to binary (Fernandez and Alvarez (2007); Alvarez and Fernandez (2008)) and ternary columns (Pulis et al. (2006)), in the present six-component column estimation study, the emphasis will be placed on: (i) the design of a reduced-component model for estimation, (ii) the employment of a robustness-oriented single-stage innovation scheme with passive structure (Pulis (2007)), (iii) the corresponding decision on the innovated components, meaning the components of the measurement stage with information and error injection, and (iv) the estimation of the effluent (distillate and bottom) impurity compositions as estimation objective.

2.3 Case example

As a representative industrial case example, consider the T110 distillation column located at SARAS refinery (Sarroch, Italy) with $N = 37$ stages and $C = 6$ components: a C3-C4 (propane-butane) splitter fed with propane (C3), iso-butane (IC4), and n-butane (NC4), as well as ethane (C2), iso-pentane (IC5), and n-pentane (NC5) as secondary components (with compositions less than 1%).

The column has a kettle reboiler (1-st stage), a total condenser (37-th stage), 35 nutter float valve trays, the feed is introduced at 19-th stage, the tray spacing is 61 cm, the column diameter is 2 m, and the pressure changes linearly along the column, with the top and bottom pressure being 16.3 Kg·cm^{-2} and 16.6 Kg·cm^{-2}, respectively. This case example represents a sufficiently important class of industrial columns, where two or three main components to be split are present, together with other secondary components in a much smaller amount.

The behavior of the "actual" six-component system (2) was numerically simulated with MATLAB, in the understanding that the same task can be performed with commercial packages (say ASPEN). The thermodynamics was simulated with ideal equilibrium (Reid et al. (1998)). The feed flow and compositions, as well as the reflux flow, and reboiler duty are given in Table 1. In all simulations, the initial conditions for the complete column system correspond to the steady-state determined by the data listed in Table 1.

Table 1. Input values for feed flow and compositions, reflux flow and reboiler duty

F (m^3·h^{-1})	82.9	C2 molar fraction	0.0036
R (m^3·h^{-1})	69.7	C3 molar fraction	0.281
Q (BTU)	19819000	IC4 molar fraction	0.236
		NC4 molar fraction	0.4746
		IC5 molar fraction	0.004
		NC5 molar fraction	0.0008

3. STRUCTURAL ANALYSIS

Motivated by the GE detectability measure-based sensor location criterion employed in previous binary (Tronci et al. (2005); Fernandez and Alvarez (2007); Alvarez and Fernandez (2008)) and ternary columns (Pulis et al. (2006)) as well as by their interpretation in terms of thermodynamic diagrams (Pulis et al. (2006)), in this section the sensor location and innovated composition structure is analyzed on the basis of stage-to-stage temperature gradients and their component-wise contributions.

3.1 Model reduction

The stage-to-stage temperature gradient about a certain operation condition is approximated as follows:

$$\Delta T_i = T_{i+1} - T_i \approx \sum_{j=1}^{C} \left.\frac{\partial T_i}{\partial c_i^j}\right|_{c_i} \Delta c_i^j = \sum_{j=1}^{C} \Delta T_{c_i^j} \qquad (3)$$

where $\Delta c_i^j = c_{i+1}^j - c_i^j$, ΔT_i is the temperature gradient at the i-th stage, and $\Delta T_{c_i^j}$ is the contribution of the i-th gradient due to the component j. The idea which underlies the model reduction criterion is to set a data assimilation scheme with a favorable compromise between data assimilation and error propagation: (i) the stages with large temperature gradients are candidates for robustness-oriented single-stage (i.e. passive) innovation, and (ii) the compositions with large contributions to the overall gradient are candidates for being both modeled states and innovated states.

Table 2. Feed compositions for the reduced model

C3 molar fraction	0.2838
IC4 molar fraction	0.2388
NC4 molar fraction	0.4774

On the basis of the steady-state solution of the complete system (2) in conjunction with the total and per-component temperature gradient formula (3), the diagram presented in Figure 1 was obtained, showing that: (i) with respect to stage-to-stage temperature change, the most sensitive zone is the enriching section around the 32-nd stage, (ii) in general, the temperature gradients are due to C3, IC4 and NC4 composition changes, (iii) the IC5, and NC5 components have a rather small contribution to the temperature gradient, and (iv) at the top of the column (around the condenser stage) the C2 has a rather important influence on temperature.

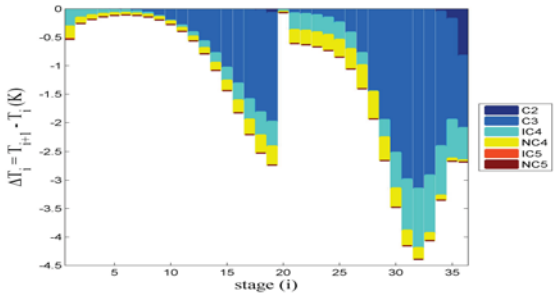


Fig. 1. Complete model-based temperature gradient and its per-component contributions

From the examination of the stage-to-stage temperature gradient and the per-component contributions to such gradient, the following structural conclusions are reached: the reduced model is obtained by retaining the C3, IC4, and NC4 components with appreciable manifestation in the temperature gradient, and discarding the three other ones (C2, IC5, and NC5) with comparatively small manifestation. Thus the reduced model is given by (1) with $C = 3$. In vector notation, the *reduced model* is written as follows:

$$\dot{x} = f(x, u, d) \qquad y = h(x) \qquad (4)$$

where $x = [c_1^{C3} c_1^{IC4} \dots c_N^{C3} c_N^{IC4}]^T, u = [RV]^T$, $d = [Fc_F^{C3} c_F^{IC4}]^T$, and $y = [T_{s_1} \dots T_{s_m}]^T$ are respectively the states, the inputs, the disturbances, and the outputs. The reduced model (4) was set with the feed compositions presented in Table 2.

The behaviors of the reduced three-component model (4) and complete six-component system (2) are presented in Figure 2, showing that the model reduction error has an appreciable (or negligible) manifestation in the distillate (or bottoms) concentrations, and the same is true for the enriching (or stripping) section. This signifies that: (i) the bottom composition can be adequately estimated, without measurement injection, by means of a reduced model-based open-loop observer, and (ii) the distillate compositions can be estimated with a reduced model-based estimator with one temperature measurement, as

the temperature measurement before the top of the column basically reflects C3 and IC4 changes.

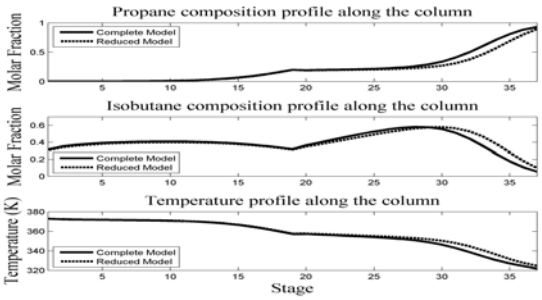


Fig. 2. C3 and IC4 composition and temperature profiles along the column with (six-component) complete and (three-component) reduced models

3.2 Sensor location and innovated states

On the basis of the steady-state solution of the reduced model (4) in conjunction with the total and per-component temperature gradient formula (3), the diagram presented in Figure 3 was obtained, showing that: (i) in the stripping (or enriching) section, the largest temperature gradient, or equivalently, the richest-in-information zone, is located below the feed (or top) tray, and (ii) in the stripping (or enriching) section, the smallest temperature gradient, or equivalently, the poorest-in-information zone, is located above the reboiler (or feed) tray.

From the preceding comments and the findings of Subsection 3.1, the next conclusions on sensor location follow: (i) one temperature measurement should be placed in the stripping section, located at the richest-in-information region (between 29-th and 33-rd stages) with the largest temperature gradient, and (ii) no temperature measurement is needed in the stripping section, as the bottom compositions can be adequately estimated using just the reduced model (without measurements). According to Figure 3, the C3 component has the largest contribution to the temperature gradient in the richest-in-information zone, meaning that the C3 component is an innovated state candidate for a robustness-oriented GE with passive structure (Pulis et al. (2006)).

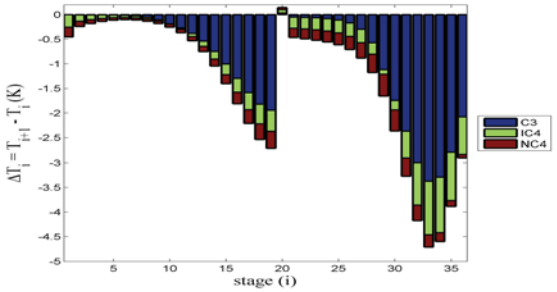


Fig. 3. Reduced model-based temperature gradient and its per-component contributions

3.3 Candidate estimation structures

According to the preceding developments, the two-effluent composition estimation task for the six-component system (2) can be performed using the reduced three-component model with one temperature sensor located between 25-th and 33-rd stages, with the C3 component as innovated state in a robustness-oriented GE with passive structure (Pulis et al. (2006)). To preclude unduly bottom-to-top model and measurement error propagation, no measurement in the stripping section is placed. To have a favorable balance between data assimilation and model-measurement error propagation from the measurement to distillate effluent composition estimate, a measurement should be located in the tray interval between stages 29 and 33, and not in any of the two top trays (stages 35 and 36) as shown in Figure 1. Thus, the preceding considerations lead to the following candidate estimation structure: (i) no sensor in the stripping section, (ii) one sensor in the enriching section between 29-th and 33-rd trays, and (iii) the C3 composition as innovated state. These structural conclusions are suggestive in the sense that: (i) candidate sensor location and innovated C3 compositions around the afore concluded candidates must be examined, and (ii) the conclusive structural assessment will be performed in the next section, on the basis of GE functioning. In order to verify the estimator performance, some different choices for the set of innovated states have been selected by following the considerations above: these innovated state sets will be illustrated and compared in Section 4.

4. STRUCTURE ASSESSMENT WITH ESTIMATOR FUNCTIONING

Having as point of departure the candidate structures identified in the preceding section, in this section the estimation structure for effluent composition estimation aims is assessed on the basis of the structure behavior with a robustness-oriented GE with passive structure (Pulis et al. (2006)). The role of the actual process will be played by the six-component system (2), and the GE will be implemented with the reduced three-component model (4).

4.1 Geometric estimator with passive innovation

Let us recall the adjustable-structure proportional-integral (PI) GE with passive structure (Alvarez and López (1999); Alvarez and Fernandez (2008)):

$$\begin{aligned} \dot{\hat{x}}_I &= f_I(\hat{x},\hat{u}) + \Phi_{x_I}^{-1} P(K_P(y_P - \hat{y}) + z) \\ \dot{\hat{x}}_{II} &= f_{II}(\hat{x},\hat{u}) \qquad \dot{z} = K_I(y_P - \hat{y}) \qquad \hat{y} = h(\hat{x}) \end{aligned} \tag{5}$$

where $\hat{x}$, $\hat{u}$, and $\hat{y}$ are the estimates of x, u, and y, $x_I \in \mathbb{R}^{n_I}$ is the set of the innovated states, $x_{II} \in \mathbb{R}^{n-n_I}$ is the set of the non-innovated states, K_P and K_I are respectively the proportional and integral gain matrices. For this kind of estimator structures, if $x_I = [c_{s_1}^{l_1} \ \ldots \ c_{s_1}^{l_1^{n_1}} \ \ldots \ c_{s_m}^{l_m^1} \ \ldots \ c_{s_m}^{l_m^{n_m}}]^T$, then K_P, K_I, Φ_{x_I}, and P assume the following form:

$$K_P = \begin{bmatrix} K_{P,s_1} & & \\ & \ddots & \\ & & K_{P,s_m} \end{bmatrix} \qquad K_I = \begin{bmatrix} K_{I,s_1} & & \\ & \ddots & \\ & & K_{I,s_m} \end{bmatrix}$$

$$\Phi_{x_I} = \left[\left. \frac{\partial \hat{T}_{s_1}}{\partial \hat{c}_{s_1}^{l_1^1}} \right|_{\hat{c}_{s_1}} \cdots \left. \frac{\partial \hat{T}_{s_1}}{\partial \hat{c}_{s_1}^{l_1^{n_1}}} \right|_{\hat{c}_{s_1}} \cdots \left. \frac{\partial \hat{T}_{s_m}}{\partial \hat{c}_{s_m}^{l_m^1}} \right|_{\hat{c}_{s_m}} \cdots \left. \frac{\partial \hat{T}_{s_m}}{\partial \hat{c}_{s_m}^{l_m^{n_m}}} \right|_{\hat{c}_{s_m}} \right] \cdot \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \qquad P = \begin{bmatrix} 1 \ldots 1 & & \\ & \ddots & \\ & & 1 \ldots 1 \end{bmatrix}^T$$

Note that: (i) the generic composition $c_{s_i}^{l_i^j}$ is innovated only by using temperature at s_i-th stage; (ii) n_i represents the number of innovated components at s_i-th stage and the condition $\sum_{i=1}^{m} n_i = n_I$ holds; (iii) $[x_I \, x_{II} \, z]^T \in \mathbb{R}^{n+m}$; (iv) P is a matrix of 1s and 0s, where for every column $i \in \{1, \ldots, m\}$ there are n_i 1s as shown above. When the integral action state is eliminated, the proportional (P) GE is obtained.

The robust convergence aspects and its connection with the tuning of gains can be seen in Alvarez and Fernandez (2008), and here it suffices to mention that: (i) the convergence criterion is coupled with rather easy-to-apply tuning guidelines, and (ii) the tuning scheme and guidelines apply over all structures. In this way, one has the certainty that the estimator functioning results are due to the structure and not to the tuning scheme.

Next the GE tuning guidelines for the GE PI estimator with passive structure are recalled. Consider an innovated state $c_{s_i}^{l_i^j}$: then, the corresponding temperature used in order to estimate this composition is T_{s_i} and the tuning parameters are:

$$K_{P,s_i} = 2\xi_{s_i}\omega_{s_i} \qquad K_{I,s_i} = {\omega_{s_i}}^2 \tag{6}$$

where ω_{s_i} and ξ_{s_i} are respectively the characteristic frequency and the damping factor of the estimator at s_i-th stage. The characteristic frequency ω_{s_i} must be chosen between 5 and 10 times faster than ω_{o,s_i} (where ω_{o,s_i} is the natural characteristic frequency of composition at s_i-th stage). Following conventional-like filter and control behavior assessments, the estimator functioning will be measured with the IAE index and the steady-state error.

As mentioned before, the estimation task is to infer the effluent impurity compositions (i.e., C3 in the bottom and IC4 in the top). Since the C3 in the reboiler is adequately described by the model without measurement, only the results of the IC4 distillate composition estimates will be presented. Several simulations have been performed, but for sakes of brevity only one of them is reported here. To test the estimator over different structures, a column transient has been induced by some step changes at the reboiler duty ($\Delta Q = +3\%$ at $t = 2$ hrs and $\Delta Q = 0$ at $t = 8$ hrs, with respect to Q value of Table 1)

Two different measures have been employed: (i) the steady-state error (7a), and (ii) the IAE index (7b)

$$e_{i,SS1}^j \quad \text{if} \quad \Delta Q = 0 \quad \text{and} \quad e_{i,SS2}^j \quad \text{if} \quad \Delta Q = +3\% \tag{7a}$$

$$IAE(e_i^j) = \int |e_i^j| dt \tag{7b}$$

where e_i^j is the composition error for the component j at i-th stage. There are two steady-state errors, since two different steady-state conditions are present, as can be seen from Figures 4, 5, and 6. For all the structures considered, the estimator has been tuned with the tuning guidelines sketched in (6), with $\omega_{s_i} = 10\omega_{o,s_i}$ and $\xi_{s_i} = 3$ for $i = 1, \ldots, m$.

4.2 Estimation with one innovated state

On the basis of the structural analysis performed in Section 3, two single-innovated state cases have been considered: $x_I = [c_{29}^{C3}]$ and $x_I = [c_{33}^{C3}]$.

The corresponding GE behavior results are presented in Figure 4 and Table 3, showing that: (i) comparing with the reduced model behavior, in both cases the estimate error undergoes a considerable reduction by measurement injection, and (ii) the case with c_{33}^{C3}-innovation yields a slightly better behavior than with c_{29}^{C3}-innovation. This result is consistent with the conclusion reached in Section 3, as 33-th stage has the largest stage-to-stage gradient and is close to the top stage.

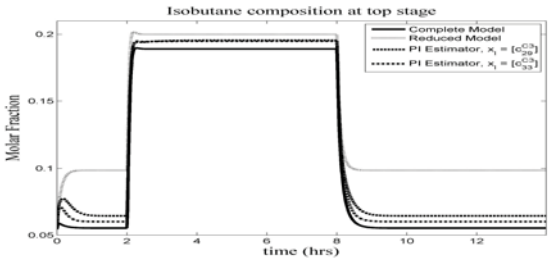


Fig. 4. IC4 distillate composition estimate with one innovated state

4.3 Estimation with two innovated states

Now, let us find out whether the joint consideration of the two previous cases leads to some estimator behavior improvement, this is $x_I = [c_{29}^{C3} \, c_{33}^{C3}]^T$.

The corresponding GE behavior results are presented in Figure 5 and Table 3, showing that there is not an appreciable improvement over the estimation structures with one innovated state discussed in the last subsection (see Figure 4).

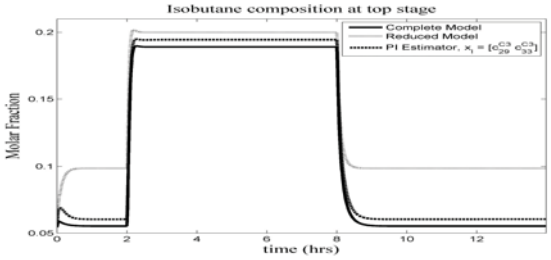


Fig. 5. IC4 distillate composition estimate with two innovated states

4.4 Estimation with three innovated states

Finally, let us investigate if the effluent estimation behavior can be improved by adding one innovated state to the two-innovated state structure presented in Subsection 4.3,

this is $x_I = [c_{29}^{C3}\ c_{31}^{C3}\ c_{33}^{C3}]^T$, with the incorporation of the innovated state c_{31}^{C3} being motivated by the fact that 29-th and 33-rd stages bracket the rich-in-information zone of the enriching section. The corresponding GE behavior results are presented in Figure 6 and Table 3, showing that there is an appreciable improvement over the two-innovated state estimation structure discussed in the last subsection (see Figure 5).

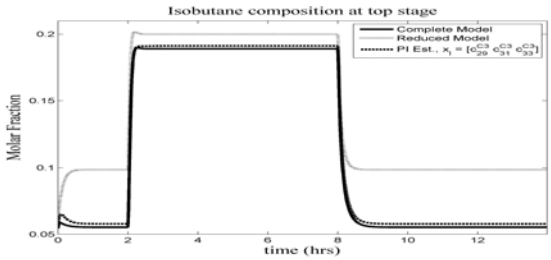


Fig. 6. IC4 distillate composition with three innovated states

Table 3. IAE values and steady-state errors (the subscript T refers to the top stage)

	$e_{T,SS1}^{IC4}$	$e_{T,SS2}^{IC4}$	$IAE(e_T^{IC4})$
Reduced model	−0.0433	−0.0109	20.1124
$x_I = [c_{29}^{C3}]$	−0.0092	−0.0064	5.9306
$x_I = [c_{33}^{C3}]$	−0.0049	−0.0058	3.8622
$x_I = [c_{29}^{C3}\ c_{33}^{C3}]^T$	−0.0052	−0.0054	3.8077
$x_I = [c_{29}^{C3}\ c_{31}^{C3}\ c_{33}^{C3}]^T$	−0.0025	−0.0022	1.7871

4.5 Concluding remarks

The behavior measures of the four estimation structures considered in this section are summarized in Table 3, showing that: the best GE estimator behavior is obtained with the three-innovated state structure, followed by the two-innovated state structure, and by the two single-innovated state structures. It must be pointed out that the steady-state estimation error is smaller than 1% and therefore comparable with typical measurement errors. These results are in agreement with the a priori structural assessments drawn in Section 3. The IC4 distillate composition estimation task can be effectively performed using: (i) a three-component reduced model, (ii) one or more sensors located between stages 29 and 33 in the enriching section, and (iii) passive innovation for component C3.

5. CONCLUSIONS

The problem of jointly designing the estimation structure and algorithm to infer an effluent composition for a six-component distillation column with temperature measurement option has been resolved in a tractable manner within a GE design framework in the light of the staged column system characteristics. The design focused on structural aspects: model reduction, sensor location, and innovation mechanism. The methodology consisted of: (i) one structural analysis step that yielded a few candidate structures and (ii) a step with conclusive structural results on the basis of estimator behavior assessment. The resulting GE (with 75 to 77 nonlinear ordinary differential equations (ODEs), depending on the structure) was considerably simpler than its EKF counterpart (with 2812 ODEs) and of tuning procedure (trial-and-error or optimization for the EKF and well defined for the GE).

Currently, work is underway to apply the proposed approach to estimate the pollutant contents in the outlet streams of an actual industrial column.

ACKNOWLEDGEMENTS

The authors kindly acknowledge the support from Italian and Mexican Foreign Ministries through the Binational Collaboration projectt SAPP01.

REFERENCES

Alvarez, J. (2000). Nonlinear state estimation with robust convergence. *Journal of Process Control*, 10, 59–71.

Alvarez, J. and Fernandez, C. (2008). Geometric estimation of nonlinear process systems. *Journal of Process Control*. doi:10.1016/j.jprocont.2008.04.017.

Alvarez, J. and López, T. (1999). Robust dynamic state estimation of nonlunear plants. *AIChE Journal*, 45(1), 107–123.

Baratti, R., Bertucco, A., Da Rold, A., and Morbidelli, M. (1995). Development of a composition estimator for binary distillation colums. application to a pilot plant. *Chemical Engineering Science*, 50(10), 1541–1550.

Baratti, R., Bertucco, A., Da Rold, A., and Morbidelli, M. (1998). A composition estimator for multicomponent distillation column - development and experimental test on ternary mixtures. *Chemical Engineering Science*, 53(20), 3601–3612.

Fernandez, C. and Alvarez, J. (2007). Constructive estimation design for binary distillation columns. In *International Symposium on Advanced Control of Chemical processes*, 75–80. Cancún, México.

Jazwinsky, A.H. (1970). *Stochastic processes and Filtering Theory*. Academic Press, New York.

Pulis, A. (2007). *Soft sensor design for distillation columns*. Ph.D. thesis, Università degli Studi di Cagliari.

Pulis, A., Fernandez, C., Baratti, R., and Alvarez, J. (2006). Geometric estimation of ternary distillation columns. In *International Symposium on Advanced Control of Chemical processes*, 573–578. Gramado, Brasil.

Reid, R.C., Prausnitz, J.M., and Poling, B.E. (1998). *The properties of gases and liquids*. McGraw-Hill, Singapore, fourth edition.

Skogestad, S. (1997). Dynamics and control of distillation columns: a tutorial introduction. *Trans IChemE, Part A*, 75, 539–559.

Tronci, S., Bezzo, F., Barolo, M., and Baratti, R. (2005). Geometric observer for a distillation column: development and experimental testing. *Ind. Eng. Chem. Res*, 44, 9884–9893.

Venkateswarlu, C. and Kumar, B.J. (2006). Composition estimation of multicomponent reactive batch distillation with optimal sensor configuration. *Chemical Engineering Science*, 61, 5560–5574.

Yang, D.R. and Lee, K.S. (1997). Monitoring of a distillation column using modified extended kalman filter and a reduced order model. *Computers and Chemical Engineering*, 21, 565–570.

Temperature Inferential Dynamic Matrix Control of Reactive Distillation Systems

Deeptanshu Dwivedi and Nitin Kaistha *
Department of Chemical Engineering
Indian Institute of Technology Kanpur India 208016

Abstract: Two-temperature inferential control of the ideal and the methyl acetate double feed reactive distillation (RD) systems operated neat is evaluated using constrained dynamic matrix control (CDMC) and traditional decentralized control. For the ideal RD system, significant improvement in the stripping tray temperature control and the transient deviation in the bottoms purity is observed using CDMC. For the methyl acetate system, CDMC results in significant improvement in the control of the two tray temperatures as well as transient deviations in both the distillate and bottoms purity. Results also show that the magnitude of the maximum through-put change for which the control system fails is noticeably higher using CDMC.

Keywords: Reactive distillation control, dynamic matrix control, temperature inferential control

1. INTRODUCTION

Reactive distillation (RD) is now an established process intensification technology combining reaction and separation in a single column with potentially significant economic savings when the reaction kinetics and component relative volatilities are favorable (Siirola, 1995). When compared to conventional "reactor-separator" processes, the high non-linearity due to direct interaction between reaction and separation combined with fewer control degrees-of-freedom makes the design of an effective control system crucial to the successful implementation of RD technology.

In probably the first paper on RD control, Roat et al (1986) demonstrated that seemingly appropriate control structures succumb to a steady state transition for a moderately large through-put change suggesting the presence of high-non-linearity. Several later articles highlighted the presence of steady state multiplicity in various RD systems (see eg Mohl et al, 1999; Taylor and Krishna, 2000). The presence of steady state multiplicity can result in non-linear dynamic phenomena under open and closed loop operation. Sneesby et al (1997) considered the implications of steady state multiplicity on the operation and control of etherification columns. Kumar and Kaistha (2008a) demonstrated the occurrence of 'wrong' control action and closed loop steady state transition for the hypothetical quaternary ideal RD column studied by Al-Arfaj and Luyben (2000).

Given the high non-linearity in RD systems, the application of non-linear control techniques has been recommended in the literature. Among the prominent non-linear RD control works, Kumar and Daoutiditis (1999) implemented a non-linear inversion based control scheme for an ethylene glycol RD column. Model based gain scheduling has been applied to an ETBE RD column. Gruner et al (2003) report the non-linear control of an industrial RD column operated by Bayer. More recently, Kawathekar and Riggs (2007) have applied a non-linear model predictive control scheme to an ethyl acetate column.

Even as non-linear model based control is widely accepted in the academic community, industrial practice remains strongly biased towards the traditional decentralized PI control and where justifiable, linear model predictive control techniques such as DMC. This is probably due to the difficulty in developing a high fidelity non-linear process model and identifying the model parameters in an industrial setting.

A careful examination of the RD control literature reveals that the control of two-reactant double-feed RD columns operated neat (no excess of a reactant) such as the quaternary ideal RD system and the methyl acetate RD system, is particularly challenging due to the need for stoichiometric balancing of the two fresh feeds. Adjusting one of the fresh feeds to maintain an appropriate tray composition has been shown to be an effective means of maintaining this balance. The use of temperature measurements instead of composition, the former being much more rugged, reliable, cheap and with fast measurement dynamics, however causes the control system to succumb to non-linear dynamic phenomena such as 'wrong' control action in the methyl acetate RD system or a closed loop steady state transition in the quaternary ideal RD system (Kumar and Kaistha, 2008a). Application of linear MPC techniques, which have found industrial acceptance, may significantly improve the performance of a temperature inferential control system for such 'difficult to regulate' processes. This work addresses the same for the quaternary ideal and the methyl acetate RD systems.

2. BASE-CASE COLUMN DESIGN

Figure 1 shows a schematic of the double feed RD columns studied in this work. The reaction $A + B \leftrightarrow C + D$ occurs in the reactive zone. For the methyl acetate system, the components A, B, C and D correspond to methanol, acetic acid, methyl acetate and water respectively. For the ideal RD

system, the component relative volatilities are in the order $\alpha_C > \alpha_A > \alpha_B > \alpha_D$ so that the reactants are intermediate boiling. The reaction kinetics and thermodynamic property models for the methyl acetate and the ideal RD system are taken from Singh et al (2005) and Al-Arfaj and Luyben (2000), respectively. The base-case design and operating conditions for the ideal RD system are taken from Kaymak and Luyben (2006). The internally heat integrated design of the methyl acetate RD system reported in Kumar and Kaistha (2008 b) is studied here. The salient design and operating conditions for the two systems are reported in Table 1.

$$A + B \leftrightarrow C + D$$

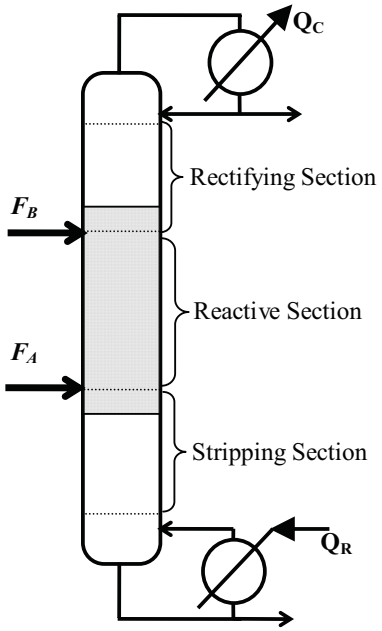


Figure 1. Schematic of a double feed reactive distillation column

Table 1: Design Parameters of Ideal RD system and Methyl Acetate system

	Ideal RD Column	Methyl Acetate RD column
Flow rate of feeds	$F_A = F_B$ = 12.6 mol/s	$F_{HAc}=F_{MeOH}$ =300 kmol/hr
$N_E/N_{RX}/N_S$ design	5/10/5	7/18/10
Feed tray locations	n_{FA} = 9; n_{FB} = 12	n_{FMeOH} =16; n_{FHAc} =28
Catalyst loading per reactive tray	0.7 kmol	300 kg
Pressure	8.5 bar	1.013 bar
Reflux ratio	2.6927	1.4877
Distillate rate	12.6 mol/s	308.63 kmol/h
Reboiler duty	0.8516 MW	3.66387 MW
Product purities	$X_C,D = X_D,B = 0.95$	$X_{MeOAc},D=0.95$ $X_{H2O},B=0.96$

3. TEMPERATURE INFERENTIAL CONTROL

3.1 Control Structures

Figure 2 & 3 plot the tray temperature sensitivities with respect to the two fresh feeds and the reboiler duty at constant reflux ratio for the ideal and methyl acetate columns. The sensitivity plots, suggest two candidate control structures, labeled as CS1 and CS2 for convenience. In CS1, F_A controls a sensitive stripping tray temperature while F_B controls a sensitive reactive tray temperature. The reboiler duty (Q_R) acts as the through- put manipulator. CS2 differs from CS1 in that the Q_R, instead of F_B is used to control a sensitive reactive tray temperature with F_B being the through-put manipulator. The two control structures are schematically depicted in Figure 4. Using bottom-up tray numbering, the temperature of Tray 2 (T_2), a stripping tray, and Tray 12 (T_{12}), a reactive tray, is controlled in the ideal RD column. In the methyl acetate column, the control tray temperature locations are Tray 2 (T_2) and Tray 13 (T_{13}). Note that in the ideal RD system, even as a rectifying tray temperature exhibits higher sensitivity than a reactive tray with respect to F_B (see Figure 2), it is not controlled due an inverse response with respect to F_B (Kaymak and Luyben, 2006) and severe input multiplicity resulting in a steady state transition for moderately large through-put changes (Kumar and Kaistha, 2008a).

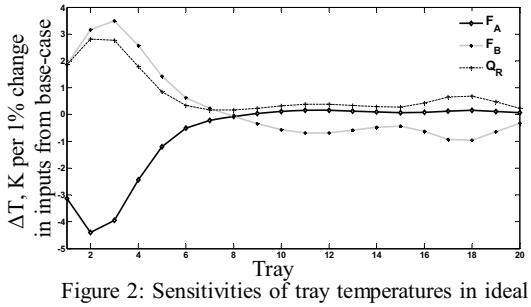


Figure 2: Sensitivities of tray temperatures in ideal RD system with respect to fresh feeds (F_A & F_B) and reboiler (Q_R) duty at fixed reflux ratio

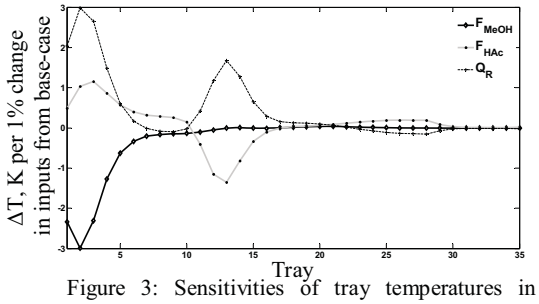


Figure 3: Sensitivities of tray temperatures in methyl acetate system with respect to Methanol feed (F_{MeOH}) and Reboiler Duty (Q_R) at fixed reflux

3.2 Dynamic Simulation and Temperature Controller Details

An in-house dynamic simulator is used to generate the open and closed loop dynamic simulation results for the two RD columns. With the two level controllers and perfect pressure control in place, a 2x2 temperature control system as in structures CS1 and CS2 is implemented. A 1 min lag is applied to the temperature measurements. The performance

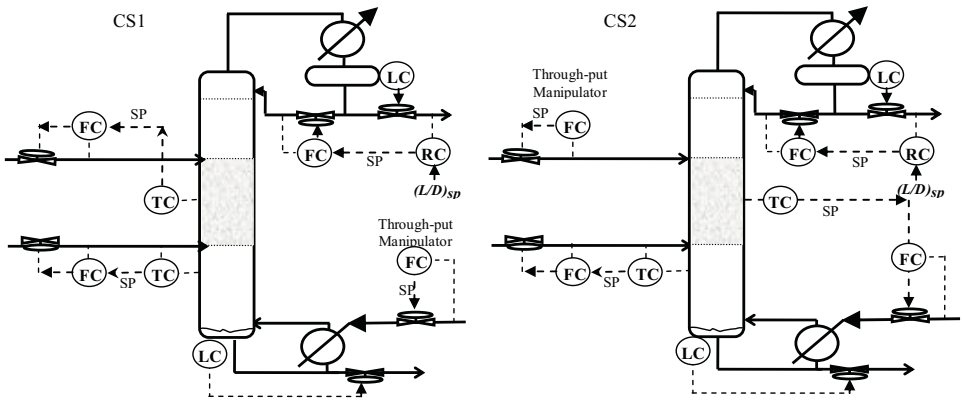


Figure 4. Schematic of used two point temperature inferential control structures

Table 2: CS1 & CS2 controller parameters for the ideal and methyl acetate RD column

Tuning parameters used in CDMC in the two systems						
System	MV	γ	CV	λ	P min	C min
Ideal RD Column with CS1	F_A	1	T_3	5	90	50
	F_B	40	T_{12}	10		
Ideal RD Column with CS2	F_A	1	T_3	5	90	50
	Q_R	4	T_{12}	10		
Methyl Acetate Column with CS1	F_{MeOH}	3	T_2	20	250	125
	F_{HAc}	4	T_{13}	25		
Methyl Acetate Column with CS2	F_{MeOH}	1	T_2	20	250	125
	Q_R	4	T_{13}	25		
Span of temperature measurements = 50K; All valves are 50% open at their design steady state. Slew rate constraints used in CDMC limits the rate of change of manipulated variable from zero to the base-case design value in no less than 2 minutes.						

of a 2x2 decentralized controller is to be compared with that of a 2x2 multivariable constrained dynamic matrix control (DMC) controller. For tuning the two decentralized PI temperature controllers, the relay feedback test is performed to obtain the ultimate gain and ultimate period of the temperature loops. The Tyreus-Luyben controller settings are then applied with appropriate de-tuning, if necessary. In the ideal RD column, both the temperature loops are tuned independently. In the methyl acetate column, sequential tuning is applied where the stripping loop is first tuned. For the DMC controller, appropriate valve saturation and slew rate constraints are applied. The slew rate constraint corresponds to the maximum rate of change of the DMC controller causing the output to saturate in two minutes. The sampling rate of the DMC controller is 0.5 minutes for the ideal RD column and 0.625 minutes for the methyl acetate RD column. The DMC step coefficient matrix is obtained using a +1% step change in the appropriate input. The tuning parameters used for CDMC are shown in Table 2.

4. RESULTS

A through-put change is considered as the primary disturbance to be rejected by the control system. The closed loop performance of the multivariable DMC and decentralized controller is now evaluated for the ideal and the methyl acetate RD columns.

Figure 5 plots the ideal RD column closed loop response to a ±20% through-put change using CS1 as the control structure. The response completion time is about 4 hours for both the controllers. Significantly tighter stripping tray temperature control (T_2) is achieved by the DMC controller while the reactive tray temperature control is comparable. This translates to tighter bottoms purity (x_D,B) control and no appreciable benefit in the distillate purity (x_C,B) control using the DMC controller.

Figure 6 plots the closed loop response to a ±20% through-put change in ideal RD column for CS2 using a decentralized and DMC 2x2 temperature controller. Significantly tighter T_2 and T_{12} temperature control is achieved by the DMC controller.

The closed loop response of the methyl acetate column to a ±20% through-put change is plotted shown in Figure 7. The CDMC temperature controller is far superior to the decentralized controller. The transient deviation in reactive T_{13} control is much smaller for the DMC controller. The tightness of the T_2 control is also better. The tighter temperature control translates to lower transient deviations in both the distillate (x_{MeOAC},D) and bottoms product purity (x_{H2O},B). Notice that with the DMC controller, the two fresh feeds move in tandem for a better stoichiometric feed balance during the transient with consequent improvement in the control performance. Figure 8 plots the closed loop response of CS2 to a ±20% through-put change for the methyl acetate RD column. For this structure also, the DMC achieves much tighter control of the stripping and reactive tray temperatures.

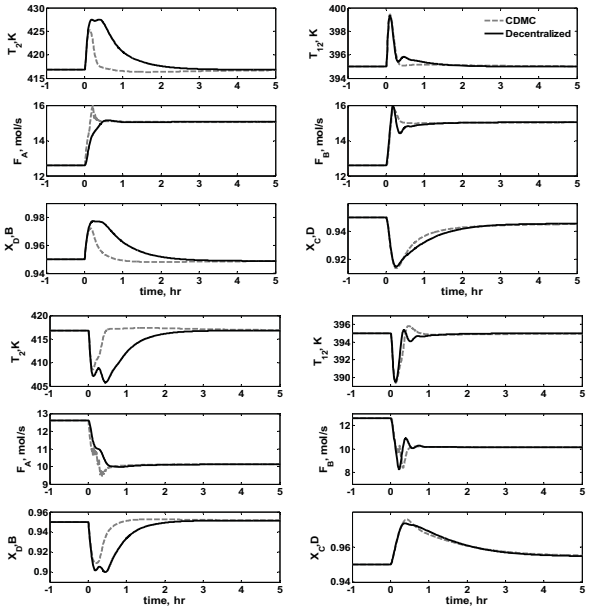


Figure 5. Closed Loop response in Ideal RD system with CS1structure for +20% (Fig A) and -20% (Fig B) change in throughput using CDMC & Traditional decentralized temperature controllers

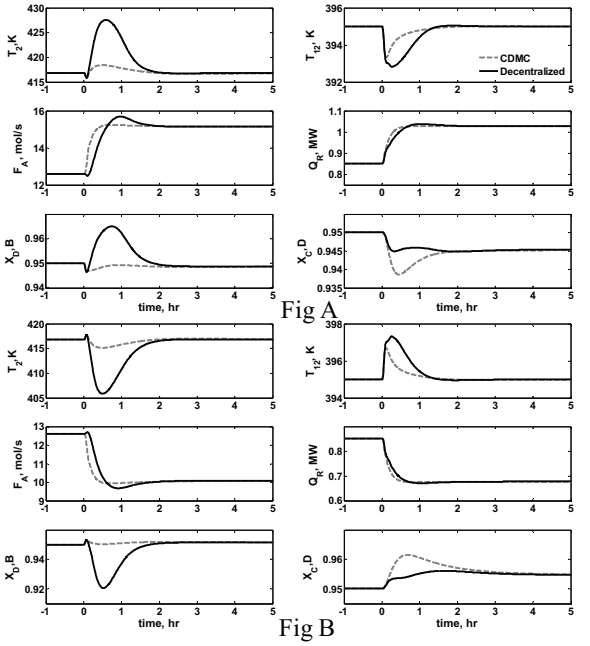


Fig A

Fig B

Figure 6. Closed Loop response in Ideal RD system with CS2 structure for +20% (Fig A) and -20% (Fig B) change in throughput using CDMC & Traditional decentralized temperature controllers

To gain a better perspective on the performance of the controller algorithms and structures, the maximum through-put change that can be handled was studied. In both the ideal and the methyl acetate RD columns, a through-put decrease turns out to be the more severe disturbance with both CS1 and CS2 For the ideal RD column, CS1 using DMC fails for value for the decentralized controller is -45%. CS2 on the a

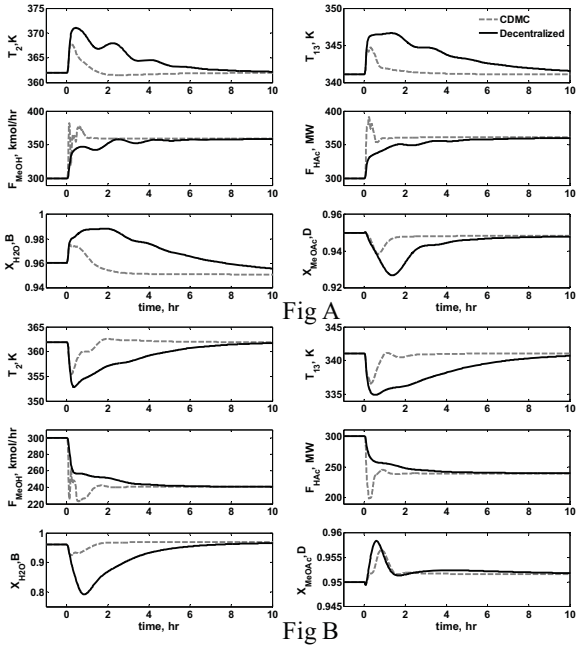


Fig A

Fig B

Figure 7. Closed Loop response in Methyl Acetate RD system with CS1 structure for +20% (Fig A) and -20% (Fig B) change in throughput using CDMC & Traditional decentralized temperature controllers

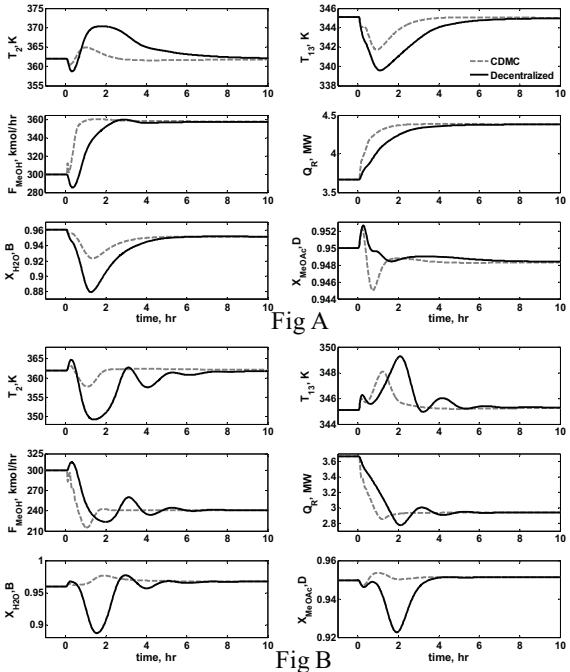


Fig A

Fig B

Figure 8. Closed Loop response in Methyl Acetate RD system with CS2 structure for +20% (Fig A) and -20% (Fig B) change in throughput using CDMC & Traditional decentralized temperature controllers

-70% step change in the through-put. The corresponding other hand fails for -65% and -55% through-put changes using respectively the DMC and decentralized control algorithms. For the methyl acetate system, CS2 with a DMC controller exhibits no improvement in the magnitude of the

maximum through-put decrease handled. Column operation using CS1 with a DMC controller allows for a 60% through-put decrease to be handled where CS1 with decentralized control fails for a 45% through-put decrease.

5 DISCUSSION

The Integral Abslolute Error (IAE) of product purity is plotted in Figure 9 & 10 as the magnitude of the through-put change is increased. Regardless of the control structure and the RD system, the 2x2 DMC control provides tighter product purity controller for large through-put changes.

We have considered a through-put change to be the primary disturbance for the double feed ideal RD systems operated neat. In some situations, variation in the fresh feed composition may also constitute a principal disturbance into the column. To test for the closed loop control performance of the temperature inferential control systems under consideration, we consider a 5 mol% step change in the purity of either feed as a disturbance. For the ideal RD system, component B in F_A and component A in F_B are the feed impurities. For the methyl acetate RD system, water is taken as the impurity in the fresh feeds. Table 3 reports the IAE of the two controlled tray temperatures and the distillate and bottoms purity for the two RD systems using CS1 and CS2. In both the RD systems, the reactive tray temperature control is poorer while the stripping loop temperature is better than decentralized controller. The inferior reactive temperature control is possibly due to the change in the step response coefficients for the altered feed conditions. Inspite of the poorer reactive temperature control, the data in the Table 3 suggests that the deviations in the distillate and bottoms purity for the DMC and decentralized controller are comparable. The DMC controller can thus withstand a feed composition disturbance.

Table 3: Comparison of Controllers for Regulatory Performance with impure fresh feed as disturbance with Integral Absolute Error (IEA's) key control variables

	Disturbance	Control Variable	CS1		CS2	
			CDMC	Decentr-alized	CDMC	Decentr-alized
Ideal RD Column	Pure F_A & Impure F_B with Z_B=0.95, Z_A=0.05	T_{12}	72.85	29.285	43.70	46.504
		T_2	334.83	433.47	266.01	695.89
		Top Purity	0.46	0.63	0.703	0.76
		Bottom Purity	0.59	0.88	0.56	1.27
	Pure F_B & Impure F_A with Z_A=0.95, Z_B=0.05	T_{12}	96.28	45.78	35.50	48.65
		T_2	296.61	202.32	143.31	347.67
		Top Purity	1.33	1.09	0.88	0.581
		Bottom Purity	0.612	0.58	0.47	0.35
Methyl Acetate RD Column	Pure F_{H2O} & Impure F_{HAc} with 5 % water	T_{13}	67.26	58.26	68.86	28.94
		T_2	126.38	460.02	120.77	425.32
		Top Purity	0.047	0.046	0.06	0.036
		Bottom Purity	0.429	1.40	0.41	1.20
	Pure F_{HAc} & Impure F_{H2O} with 5 % water	T_{13}	41.19	167.94	52.65	82.71
		T_2	125.61	535.69	137.49	371.41
		Top Purity	0.49	0.51	0.516	0.51
		Bottom Purity	0.45	2.68	0.350	1.62

The asymmetry in the closed loop results (see e.g. Figure 5) suggests the presence of non-linear effects. To investigate this, Figure 11 & 12 plots the variation in the two tray temperatures with respect to the fresh feeds and the reboiler duty at constant reflux ratio. In the ideal RD system, input multiplicity in the stripping tray temperature (T_2) for excess F_B and lower F_A is evident. The reactive tray temperature (T_{12}) exhibits input multiplicity as F_A and F_B are decreased. For the methyl acetate RD system, even as a crossover with respect to base-case steady state does not occur for the range of variation in the column inputs shown, the reactive tray temperature exhibits gain sign reversal with respect to the fresh acetic acid feed. Also, notice the severe directionality in the steady state reactive tray temperature response with respect to the fresh methanol feed with a very small decrease in temperature as the feed rate is increased and very large increase as it is decreased. This extreme directionality, at least partially, explains the asymmetry. The input-output (IO) relations (Figure 11 & 12) can also be used to understand the control system failure mode to large through-put changes. For example, for CS1 with a decentralized controller, for a -50% through-put change, the F_B valve ends up shutting down in slightly under an hour with F_A maintaining T_2 at its set-point. The control system failure mode likely corresponds to 'wrong' control action due to input multiplicity.

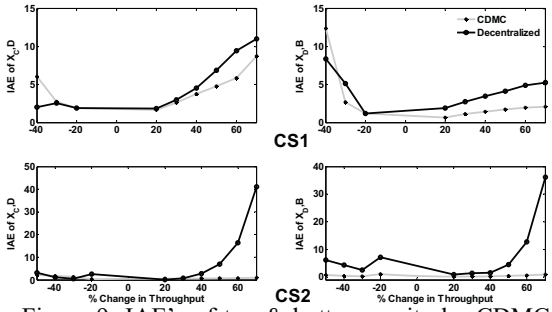


Figure 9: IAE's of top & bottom purity by CDMC and decentralized control using CS1 & CS2 in Ideal RD column for large throughtput changes

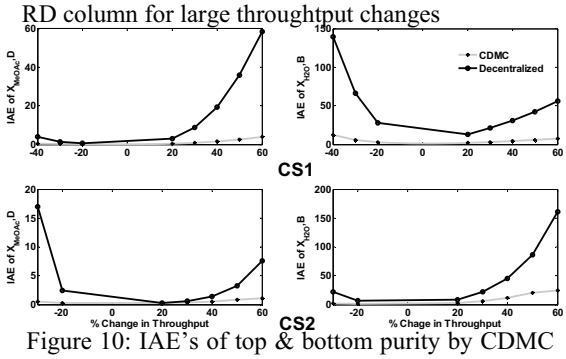


Figure 10: IAE's of top & bottom purity by CDMC and decentralized control by CS1 & CS2 in Methyl Acetate RD column for large throughtput changes

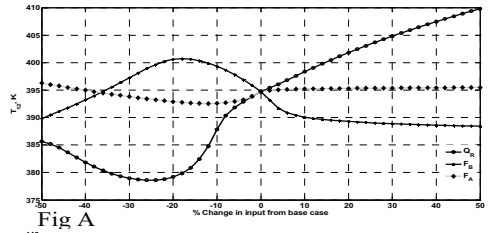


Fig A

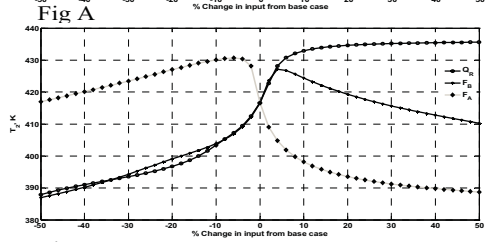


Fig B

Figure 11: I/O relation in Ideal RD system

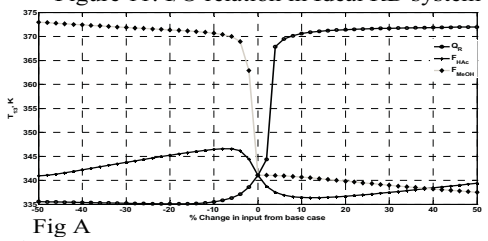


Fig A

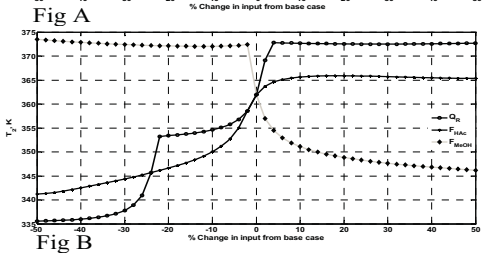


Fig B

Figure 12: I/O relation in Methyl acetate column

It is comforting to note that in spite of the highly non-linear IO relations, a linear control system (decentralized or DMC) effectively rejects such a severe disturbance without succumbing to non-linear dynamic phenomena.

6. CONCLUSIONS

In conclusion, this work demonstrates that the application of constrained dynamic matrix control for two-point temperature inferential control of double feed RD columns operated neat improves the control system performance in terms of the maximum through-put handled and/or the tightness of product purity control achieved. Specifically, in the ideal RD column, significantly tighter bottoms purity control is achieved In the methyl acetate column, tighter control of both the distillate and bottoms purity is achieved using the DMC controller for both the structures. The maximum through-put decrease handled is noticeably higher in CS1 while no such benefit was observed for CS2. These results suggest an overall incentive for the application of linear model predictive control algorithms over conventional decentralized of the highly non-linear RD systems.

REFERENCES

Al-Arfaj, M.A., Luyben, W.L., (2000). Comparison of alternative control structures for an ideal two-product reactive distillation column, *Ind. and Engg. Chem. Res.*, 39, 3298-3307.

Grüner, S., Mohl, K.D., Kienle, A., Gilles, E.D., Fernholz, G., Friedrich, S., (2003). Nonlinear control of a reactive distillation column**,** *Cntl. Engg. Prac., 11,* 915-925.

Kawathekar, R., Riggs, J.B., (2007). Nonlinear model predictive control of a reactive distillation column, *Cntl. Engg. Prac, 15,* 231-239.

Kaymak, D. B., Luyben, W. L., (2006). Evaluation of a Two Temperature Control Structure for a Two-Reactant/Two-Product Type of Reactive Distillation Column, *Chem. Engg. Sc., 61*, 4432-4450.

Kumar, A., Daoutidis, P., (1999). Modeling, analysis and control of ethylene glycol reactive distillation column. *AIChE Journal, 45(1),* 51– 68.

Kumar, M.V.P., Kaistha, N. (2008a). Steady-state multiplicity and its implications on the control of an ideal reactive distillation column. *Ind. and Engg. Chem. Res*, *47*, 2778-2787.

Kumar, M.V.P., Kaistha, N., (2008b), Internal Heat Integration and Controllability of Double Feed Reactive Distillation Columns, 1. Effect of Feed Tray Location, *Ind. and Engg. Chem. Res,* 47, 7294–7303.

Mohl, K., Kienle, A., Gilles, E., Rapmund, P., Sundmacher, K., Hoffmann, U., (1999). Steady-state multiplicities in reactive distillation columns for the production of fuel ethers MTBE and TAME: theoretical analysis and experimental verification, *Chem. Engg. Sc., 54(8)*, 1029-1043.

Roat, S., Downs, J., Vogel, E., Doss, J., (1986). Integration of rigorous dynamic modeling and control system synthesis for distillation columns. In Morari, M., McAvoy, T.J., (Eds), *Chem. Proc. Control; CPC III,* Elsevier: Amesterdam, The Netherlands, 99.

Singh, B.P., Singh, R., Kumar, M.V.P., Kaistha, N., (2005). Steady state analysis of reactive distillation using homotopy continuation method, *Che. Engg. Res. and Des., 83A,* 959-968.

Siirola, J. J., (1995). An industrial perspective on process synthesis, *A.I.Ch.E. Symposium Series, 91 (304),* 222-233.

Sneesby, M.G., Tade, M.O., Smith, T.N. (1997a.). Implications of steady-state multiplicity for operation and control of etherifcation columns. *Distillation and absorption '97, Institute of the Chemical Engineers Symposium Series*, *142*, 205-216.

Taylor, R., Krishna, R (2000). Modelling Reactive Distillation. *Chem. Engg. Sc.*, 55(22), 5183-5229

A General Quadratic Performance Approach to Binary Distillation Control

Ansgar Rehm *

* *University of Applied Sciences Osnabrück,D-49076 Osnabrück, Germany (Tel: ++49-541-9692156; e-mail: a.rehm@fh-osnabrueck.de).*

Abstract: High purity distillation control of a binary mixture in a tray column is considered in the paper at hand. The approach is based on an inferential control idea: dynamics within the column may be described as movements of concentration waves; the position of the wave front on the one hand side can be inferred from few temperature measurements, on the other hand the position implies the product concentrations. Dynamics of wave propagation is derived by simplification of a first principles model of the column. The resulting descriptor model is the basis for a recent LMI based controller design scheme that provides general quadratic performance for descriptor systems.

Keywords: Quadratic performance; descriptor system; binary distillation; inferential Control.

1. INTRODUCTION

Distillation is one of the most common separation processes in the chemical industries and it is also one of the most energy consuming ones. Therefore the control of this kind of processes has been a focus of process control for many years. Most approaches toward control of distillation columns are based on linear models which are based on identification techniques (e.g. Skogestad et al. [1988], Allgöwer and Raisch [1992]). The disadvantage of identified models is the missing physical interpretation. First principle models on the other hand are rather complex and typically not suitable for a direct model based controller computation.

In the paper at hand a reduced model for a distillation column is derived in descriptor form. The control problem is captured as a generalized quadratic performance problem. A solution to this problem is briefly reviewed (see Rehm and Allgöwer [2002] for details) and applied to the problem at hand.

The resulting controller is tested by means of a high order nonlinear model of the distillation process.

2. DESCRIPTOR MODEL

Separation of a binary mixture in a 40 tray distillation column with one feed stream is considered. A schematic representation of the process is given in the left part of Fig. 1. Exemplary the separation of two alcohols (Methanol,n-Propanol) is taken into account. The mixture is fed in the column with the feed flow rate F. Feed flow rate F and feed composition x_F (molar fraction) are determined by upstream processes.

The stationary feed flow rate and feed composition are corrupted by disturbances. The feed stream separates the column into rectifying- (upper part of the column) and stripping section (lower part of the column). Separation is achieved due to intensive heat and mass transfer between liquid flow and countercurrently rising vapor flow.

At the bottom of the column the liquid flow splits up into a liquid product stream which is removed with flow rate B from the column and a stream which is, after being heated in the reboiler, recirculated back to the column as vapor flow with flow rate V.

At the top of the column the vapor flow with the accumulated more volatile product is completely condensed in the condenser. The condensate is partly pumped back in the column with a flow rate L (reflux stream) and is partly removed as the distillate product with a flow rate D (Deshpande [1985]).

We consider the distillation column in "LV" configuration, that is: liquid flow rate L and vapor flow rate V are considered to be control inputs. Measured variables are the concentrations on trays 14 and 28.

2.1 Control Objectives

The main control objective is to stabilize the product concentrations at the top and bottom of the column at their stationary values. Additionally the deviations from the stationary values due to disturbances in the feed flow should be small.

Table 1. Notation for model variables

x_i	... liquid concentration of the more volatile component on the i^{th} tray
y_i	... vapor concentration of the more volatile component on the i^{th} tray
n_i	... liquid holdup of the i^{th} tray
$(\cdot)_{B,M,D}$	... corresponding quantities of reboiler feed tray, and condensor

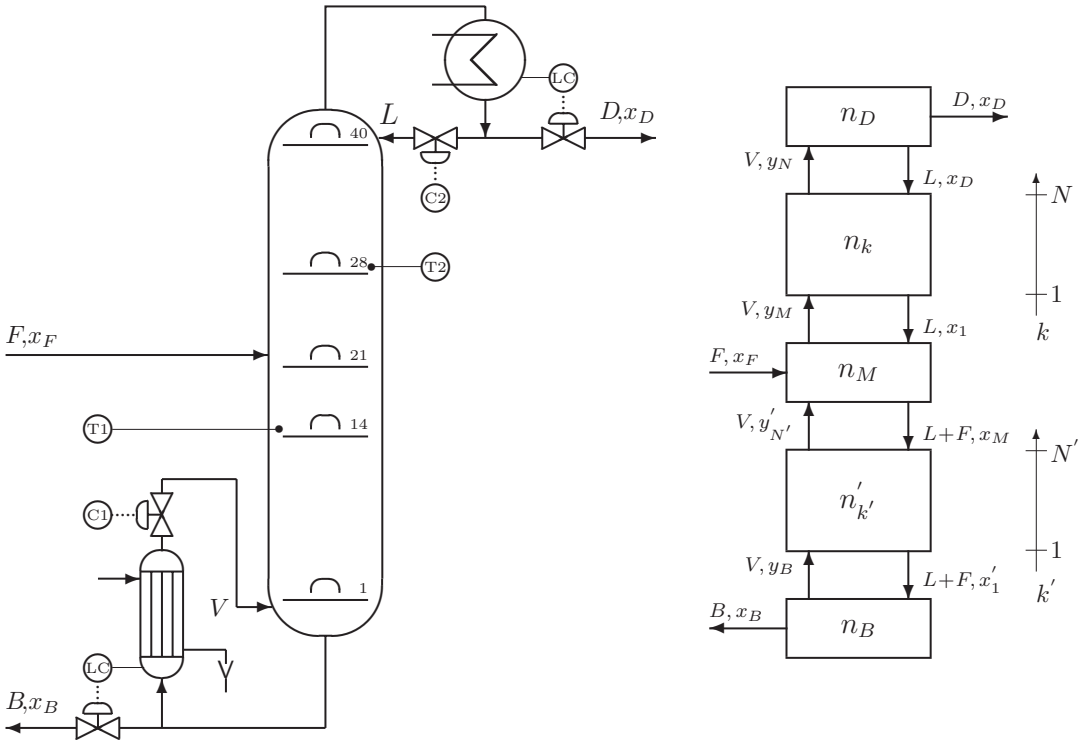


Fig. 1. Scheme of considered 40 tray distillation column (left) and subsystem structure for reduced modelling (right).

2.2 Reference Dynamics

A relatively detailed nonlinear model (CMO model without pressure losses, energy balances, and hydrodynamics (Deshpande [1985])) is used for simulation studies. The modelling equations describe the liquid concentrations of the more volatile component and are derived from the mass balance for every tray and for reboiler and condensor.

$$y_i = \frac{\alpha x_i}{1 + (\alpha - 1)x_i}, \quad \alpha = const. \tag{1}$$

The most important source of nonlinearity in the model are the equations (1) describing the vapor-liquid equilibrium (constant relative volatility α). The resulting model consists of 42 first order differential equations (40 equations from the intermediate trays plus two equations from reboiler and condensor).

2.3 Reduced Dynamics (Descriptor Model)

Starting point for the development of a reduced model in descriptor form of the distillation column is the fact (Retzbach [1986]) that qualitatively the behaviour of the column towards changes in the input values (V, L, F, x_F) can be regarded as motion and distortion of the stationary concentration profile (concentration versus tray number).

Instead of having detailed mass balances for rectifying and stripping section, the idea for a reduced model is thus to capture dynamics just by one position variable for a suitable concentration profile in every column section Due to (1) it is sufficient to consider a moving concentration profile only for the lighter component (measured in molar fractions, denoted by x in the following). Therefore the reduced model will contain two positions (s_r for the rectifying section and s_s in the stripping section) and three

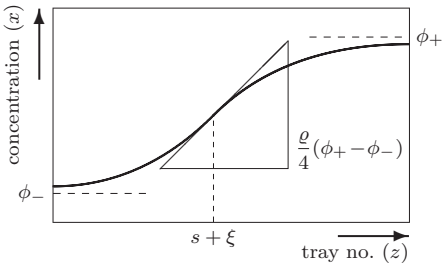


Fig. 2. Illustration of the shape parameters in function (2)

concentrations (concentration x_B in the reboiler, x_M for the feed tray, and x_D in the condenser) as state variables.

Here, only a sketch of the derivation of the reduced model in descriptor form is given, details can be found in Rehm [2004]. Furthermore we restrict ourselves to the presentation of the procedure for one column section, the deviation for the other section is completely analogous. The trays in this section are numbered by $z = 1, \ldots, N$ (see right side of Fig. 1). The concentration profile is modeled with the (continuous) function $x(z)$ (eq. (2), Fig. 2) which is well suited to describe the stationary profile in long packed columns (Kienle [1998]):

$$x(z) = \phi_- + \frac{\phi_+ - \phi_-}{1 + e^{-\varrho(z-s-\xi)}}. \tag{2}$$

With the least squares method the shape parameters ϕ_-, ϕ_+, ϱ, and ξ (see Fig. 2) are calculated such that $x(z)$ matches the stationary concentration profile ($s = 0$, i.e. s denotes the displacement relative to the stationary case) of the tray column for the discrete values $z = 1, \ldots, N$ in a least squares sense.

$$\begin{bmatrix} AY_1+Y_1^{\mathrm{T}}A^{\mathrm{T}}+ \\ B_2\hat{C}_K+(B_2\hat{C}_K)^{\mathrm{T}} & * & * & * \\ (A+B_2\hat{D}_KC_2)^{\mathrm{T}}+\hat{A}_K & X_1^{\mathrm{T}}A+\hat{B}_KC_2+ \\ & (X_1^{\mathrm{T}}A+\hat{B}_KC_2)^{\mathrm{T}} & * & * \\ B_1^{\mathrm{T}}+D_{21}^{\mathrm{T}}\hat{D}_K^{\mathrm{T}}B_2^{\mathrm{T}}+ & B_1^{\mathrm{T}}X_1+D_{21}^{\mathrm{T}}\hat{B}_K^{\mathrm{T}}+ & (D_{11}+D_{12}\hat{D}_KD_{21})^{\mathrm{T}}W_P+ \\ W_P^{\mathrm{T}}(C_1Y_1+D_{12}\hat{C}_K) & W_P^{\mathrm{T}}(C_1+D_{12}\hat{D}_KC_2) & W_P^{\mathrm{T}}(D_{11}+D_{12}\hat{D}_KD_{21})+V_P & * \\ Q_P^{\mathrm{T}}(C_1Y_1+D_{12}\hat{C}_K) & Q_P^{\mathrm{T}}(C_1+D_{12}\hat{D}_KC_2) & Q_P^{\mathrm{T}}(D_{11}+D_{12}\hat{D}_KD_{21}) & -\Sigma_P \end{bmatrix} < 0. \tag{3}$$

$$\begin{aligned} Y_1 &:= RE^{\mathrm{T}}+E^{\mathrm{T}\perp}W_Y,\ R>0, \qquad \begin{bmatrix} R & E^+ \\ E^{\mathrm{T}+} & S \end{bmatrix} > 0 \\ X_1 &:= SE+E^{\perp}W_X, \quad S>0, \end{aligned} \tag{4}$$

$$\begin{pmatrix} * & 0 & 0 & 0 & 0 \\ * & * & * & 0 & 0 \\ 0 & 0 & * & 0 & 0 \\ 0 & 0 & * & * & * \\ 0 & 0 & 0 & 0 & * \end{pmatrix} \begin{pmatrix} \frac{d\Delta x_B}{dt} \\ \frac{d\Delta s_r}{dt} \\ \frac{d\Delta x_M}{dt} \\ \frac{d\Delta s_s}{dt} \\ \frac{d\Delta x_D}{dt} \end{pmatrix} = \begin{pmatrix} * & * & * & 0 & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & * \\ 0 & 0 & * & * & * \\ 0 & 0 & * & * & * \end{pmatrix} \begin{pmatrix} \Delta x_B \\ \Delta s_r \\ \Delta x_M \\ \Delta s_s \\ \Delta x_D \end{pmatrix} + \begin{pmatrix} 0 & * \\ 0 & * \\ * & * \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x_F \\ \Delta F \end{pmatrix} + \begin{pmatrix} * & * \\ * & * \\ * & * \\ * & * \\ 0 & * \end{pmatrix} \begin{pmatrix} \Delta L \\ \Delta V \end{pmatrix}$$

$$\begin{pmatrix} \Delta x_{14} \\ \Delta x_{28} \end{pmatrix} = \begin{pmatrix} * & * & * & 0 & 0 \\ 0 & 0 & * & * & * \end{pmatrix} (\Delta x_B, \Delta s_r, \Delta x_M, \Delta s_s, \Delta x_D)^{\mathrm{T}} \tag{5}$$

However, while ϱ and ξ are kept constant, ϕ_- and ϕ_+ are used as adaptation parameters since concentration profiles not only move but also are distorted. Adaptation of these parameters is based on the requirement that (2) should also match the concentrations for the neighbouring systems when evaluated for $z = 0$ and $z = N + 1$. This adaptation rule implies that the time derivatives of x_B, x_M, and x_D influence the dynamics of wave propagation. The linearisation of the overall reduced descriptor model of the distillation column is given in (5). Here "Δ" implies deviations from the stationary value while "$*$" denotes numerical entries. A detailed derivation of the model and numerical values are given in Rehm [2004].

3. CONTROLLER COMPUTATION

3.1 Synthesis for Generalized Quadratic Performance for Descriptor Systems

The idea of generalized quadratic performance (GQP) control is to impose a general quadratic constraint of the type

$$\int_0^T \begin{bmatrix} \boldsymbol{z}(t) \\ \boldsymbol{w}(t) \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} U_P & W_P \\ W_P^{\mathrm{T}} & V_P \end{bmatrix} \begin{bmatrix} \boldsymbol{z}(t) \\ \boldsymbol{w}(t) \end{bmatrix} dt \ll 0, \tag{6}$$

on the external input/output chanel $\boldsymbol{w} \to \boldsymbol{z}$ of a generalized plant description G_{cl}. Here the notation "$\ll 0$" means that $\int_0^T Q(\boldsymbol{w}(t), \boldsymbol{z}(t))\, dt \le -\epsilon \int_0^T \boldsymbol{w}^{\mathrm{T}}(t)\boldsymbol{w}(t))\, dt$ holds for all $\boldsymbol{w}(\cdot) \in L_2$ and some fixed $\epsilon > 0$.

The rather general GQP problem contains some important control problems as a special case if the objective parameters $U_P \ge 0$, $V_P = V_P^{\mathrm{T}}$, and W_P are chosen accordingly (Scherer et al. [1997]). For example

- the H_∞ constraint $\|G_{cl}\|_\infty < \gamma$, if U_P, V_P, and W_P are specified as $U_P = \frac{1}{\gamma}I$, $V_P = -\gamma I$, $W_P = 0$;
- the strict passivity constraint $G_{cl}(j\omega)+G_{cl}(j\omega)^* > 0$ for all $\omega \in \mathbb{R} \cup \{\infty\}$, when U_P, V_P, W_P are chosen as $U_P = 0$, $V_P = 0$, $W_P = -I$;
- sector constraints of the form

$$\int_0^T (\boldsymbol{z}(t)-\alpha\boldsymbol{w}(t))^{\mathrm{T}} (\boldsymbol{z}(t)-\beta\boldsymbol{w}(t))\, dt \ll 0 \tag{7}$$

for $U_P = I$, $V_P = -\alpha\beta I$, $W_P = -\frac{1}{2}(\alpha+\beta)I$.

We consider a generalized plant description Σ in descriptor form

$$\Sigma: \quad \begin{aligned} E\dot{\boldsymbol{x}}(t) &= A\boldsymbol{x}(t) + B_1\boldsymbol{w}(t) + B_2\boldsymbol{u}(t) \\ \boldsymbol{z}(t) &= C_1\boldsymbol{x}(t) + D_{11}\boldsymbol{w}(t) + D_{12}\boldsymbol{u}(t) \\ \boldsymbol{y}(t) &= C_2\boldsymbol{x}(t) + D_{21}\boldsymbol{w}(t) \end{aligned} \tag{8}$$

with $\boldsymbol{x}(t) \in \mathbb{R}^{n_x}$, $\boldsymbol{w}(t) \in \mathbb{R}^{n_w}$, $\boldsymbol{u}(t) \in \mathbb{R}^{n_u}$, $\boldsymbol{z}(t) \in \mathbb{R}^{n_z}$, and $\boldsymbol{y}(t) \in \mathbb{R}^{n_y}$ denoting the generalized state variables, the external input variables, the control input variables, the external output variables, and the measurement variables, respectively. E and A are square constant matrices where, explicitly, E is allowed to be singular, i.e. $\mathrm{rank}(E) =: r \le n_x$. The remaining matrices are constant matrices of appropriate dimension.

The control problem is, for given matices $U_P \ge 0$, $U_P \in \mathbb{R}^{n_z \times n_z}$, $V_P = V_P^{\mathrm{T}} \in \mathbb{R}^{n_w \times n_w}$, and $W_P \in \mathbb{R}^{n_z \times n_w}$, to find a linear output feedback controller such that the undisturbed closed loop ($\boldsymbol{w} \equiv \boldsymbol{0}$) is an admissible system and such that the transfer matrix from the external input $\boldsymbol{w}$ to the external output $\boldsymbol{z}$ suffices a general quadratic performace bound (6).

The actual design problem therefore consists in the selection of matrices U_P, V_P, W_P such that the transfer matrix from $\boldsymbol{w}$ to $\boldsymbol{z}$ reflects the performance requirements (e.g. robustness, energy dissipation, ...). Since we aim at an admissible close loop, we assume in the following the corresponding necessary stabilizability/detectability properties for descriptor systems, namely stabilizability/detectability at infinity (see also Dai [1989]).

With a controller K_E,

$$K_E: \quad \begin{array}{l} E\dot{\boldsymbol{\zeta}}(t)=A_K\boldsymbol{\zeta}(t)+B_K\boldsymbol{y}(t) \\ \boldsymbol{u}(t)=C_K\boldsymbol{\zeta}(t)+D_K\boldsymbol{y}(t),\ \boldsymbol{\zeta}(t)\in\mathbb{R}^{n_x} \end{array} \tag{9}$$

parametrized by A_K, B_K, C_K, D_K the closed loop system is given by

$$\begin{aligned} E_{cl}\dot{\boldsymbol{\xi}}(t) &= A_{cl}\boldsymbol{\xi}(t)+B_{cl}\boldsymbol{w}(t) \\ \boldsymbol{z}(t) &= C_{cl}\boldsymbol{\xi}(t)+D_{cl}\boldsymbol{w}(t), \qquad \boldsymbol{\xi}(t)\in\mathbb{R}^{2n_x}, \end{aligned} \tag{10}$$

$$\begin{aligned} E_{cl} &= \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix}, & A_{cl} &= \begin{bmatrix} A+B_2D_KC_2 & B_2C_K \\ B_KC_2 & A_K \end{bmatrix}, \\ B_{cl} &= \begin{bmatrix} B_1+B_2D_KD_{21} \\ B_KD_{21} \end{bmatrix}, & C_{cl}^{\mathrm{T}} &= \begin{bmatrix} C_1^{\mathrm{T}}+C_2^{\mathrm{T}}D_K^{\mathrm{T}}D_{12}^{\mathrm{T}} \\ C_K^{\mathrm{T}}D_{12}^{\mathrm{T}} \end{bmatrix}, \\ D_{cl} &= (D_{11}+D_{12}D_KD_{21}) \end{aligned} \tag{11}$$

Then a sufficient condition for a controller K_E solving the GQP control problem for DAE systems is given by the following theorem:

Theorem 1. Consider a plant (8) and a controller (9). There exists a controller parameterization A_K, B_K, C_K, D_K such that the undisturbed (i.e. $\boldsymbol{w}\equiv 0$) closed loop system (10) is admissible with general quadratic performance if the LMIs (3), (4)[1] admit a solution $\{R, S, W_Y, W_X, \hat{A}_K, \hat{B}_K, \hat{C}_K, \hat{D}_K\}$.

Remark. The preceding theorem constitutes also a necessary condition for the existence of a controller with GQP in the cases, where the corresponding analysis result is necessary for general quadratic performance, i.e. especially in the case of the H_∞ control problem. Therefore the results of Masubuchi et al. [1997] are included in Theorem 1 as a special case.

Theorem 1 is constructive: controller computation consists of three steps:

- Solution of the LMIs (3), (4). This is possible via effective numerical tools tailored for LMI problems arising from control theoretic problem setups (e.g. Gahinet et al. [1994], El Ghaoui et al. [1995]).
- Computation of non-singular matrices X_3, Y_3 such that

$$X_1Y_1+X_2Y_3=I \tag{12}$$
$$X_3Y_1+X_4Y_3=0 \tag{13}$$

  hold together with the coupling condition $E^{\mathrm{T}}X_2 = X_3^{\mathrm{T}}E$. This is a essentially a factorization problem on the range of E which is always solvable provided (3), (4) have a solution.
- Solution of the linear equations

$$\begin{aligned} \hat{D}_K &:= D_K \\ \hat{C}_K &:= C_KY_3+D_KC_2Y_1 \\ \hat{B}_K &:= X_3^{\mathrm{T}}B_K+X_1^{\mathrm{T}}B_2D_K \\ \hat{A}_K &:= X_1^{\mathrm{T}}(A+B_2D_KC_2)Y_1+X_3^{\mathrm{T}}A_KY_3+ \\ &\quad +X_3^{\mathrm{T}}B_KC_2Y_1+X_1^{\mathrm{T}}B_2C_KY_3 \end{aligned} \tag{14}$$

  for the controller matrices D_K, C_K B_K, A_K.

[1] Here E^+ denotes any generalized inverse with the property $EE^+E = E$ and "$*$" is used in order to indicate the symmetric expansion of a block matrix.

3.2 Distillation Control Problem as S/KS Mixed Sensitivity Problem

As a special case of generalized quadratic performance, the H_∞ control problem for the distillation problem is solved. The control objectives are translated into a mixed sensitivity set-up depicted in Fig. 4. with G representing the plant (reduced model in descriptor form), K the controller, and W_1, W_2, V frequency dependent weighting matrices. Controller design by "loop shaping" requires a selection of the weighting matrices such that the solution of the H_∞ control problem

$$\left\| \begin{array}{c} W_1(I+GK)^{-1}V \\ -W_2K(I+GK)^{-1}V \end{array} \right\|_\infty \overset{!}{<} \gamma \tag{15}$$

results in a well behaved closed loop system. In this

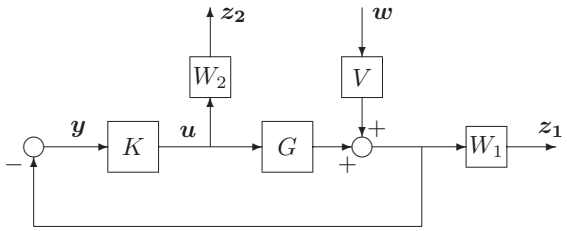


Fig. 4. Mixed sensitivity configuration

setup V can be interpreted as a filter which models the disturbance considered to be relevant for the problem at hand. With $S(s) := (I+GK)^{-1}$ being the sensitivity matrix of the closed loop the expression (15) with $\gamma = 1$ suggests to choose W_1 to be approximately the inverse of the wanted behavior for $S(s)$ and analogously W_2 to be the inverse of $K \cdot S$. General indications on selecting these weighting matrices can be found in Skogestad and Postlewaite [1996].

In case of the distillation control problem at hand an indirect approach is taken: with stabilizing the measured concentrations x_{14}, x_{28} also the stationary profiles are fixed and thus approximately also the product concentrations. In order to realize this idea the descriptor S/KS H_∞ control problem depicted in Figure 4 (with G being the descriptor model (5)) is solved by the outlined descriptor GQP synthesis procedure with specification of W, Q, and Σ as for the H_∞ set-up. The synthesis LMIs are jointly optimized with respect to γ. A final value of $\gamma = 1.01$ shows that the control objectives are approximately met.

The resulting controller has a dynamical order of 9, i.e. equal to the order of the generalized plant description. After removing the fastest two eigen-modes of the controller in order to avoid numerical problems due to stiffness, the controller is tested in simulation studies with the nonlinear CMO model of the distillation column.

4. RESULTS

In Figure 5 the stationary concentration profiles for various severe persistent disturbances for the closed loop are shown. It can be seen that the controller is able to stabilize the profile position although the disturbances result in a distortion of the stationary profile in the vicinity of the feed tray.

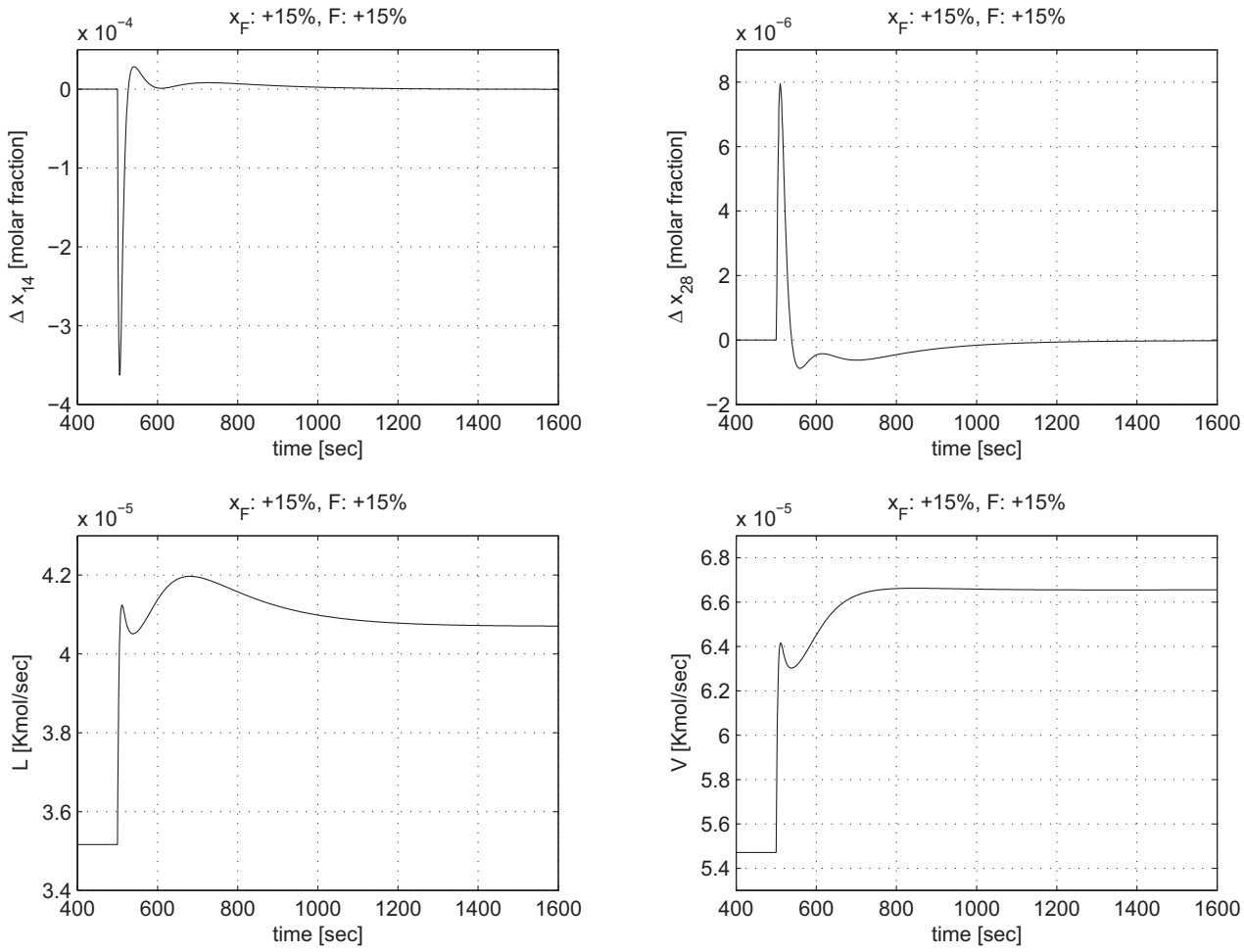


Fig. 3. Step responses for the controlled distillation column (+15% increase in feed flow rate F and +15% increase in feed concentration x_F with respect to stationary values at $t = 500sec$). Top: deviations from the steady state for the controlled variables x_{14} and x_{28}. Bottom: control variables, i.e. liquid flow rate L and vapor flow rate V.

In Figure 3 a detailed view on the control variables and the error in the controlled variables is given for a mutual step in the feed flow rate and feed composition. The plots show a fast transient behavior and small deviations. Furthermore no excessive action in the control variables is needed.

5. CONCLUSION

The generalized quadratic performance control problem for descriptor systems is solved for a reduced model of a distillation control problem. The resulting controller shows rather good results for a nonlinear reference model. The descriptor problem formulation is a direct result of reduced modeling. Furthermore, also standard approaches to build generalized plant descriptions easily fit into the descriptor system set-up. This was demonstrated by means of a S/KS - control problem formulation that directly leads to a descriptor model.

REFERENCES

F. Allgöwer and J. Raisch. Multivariable controller design for an industrial distillation column. In N.K Nichols and D.H. Owens, editors, *The Mathematics of Control Theory*. Clarendon Press, Oxford, 1992.

L. Dai. *Singular Control Systems*, volume 118 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, 1989.

P. B. Deshpande. *Distillation Dynamics and Control*. Instrument Society of America, Research Triangle Park, NC, 1985.

L. El Ghaoui, F. Delebecque, and R. Nikoukhah. *LMITOOL: A User-Friendly Interface for LMI Optimization*, 1995. Available at ftp.ensta.fr in pub/elghaoui/lmitool.

P. Gahinet, A. Nemirowskii, A.J. Laub, and M. Chilali. The LMI control toolbox. In *Proc. 33rd IEEE Conf. Decision Contr.*, pages 2038–2041, Lake Buena Vista, FL, 1994.

A. Kienle. Reduced models for multicomponent separation processes using nonlinear wave propagation theory. In *CHISA'98, 13th International Congress of Chemical and Process Engineering*, Praha, Czech Republic, 1998. Paper 109.

I. Masubuchi, Y. Kamitane, A. Ohara, and N. Suda. H_∞ control for descriptor systems: A matrix inequalities

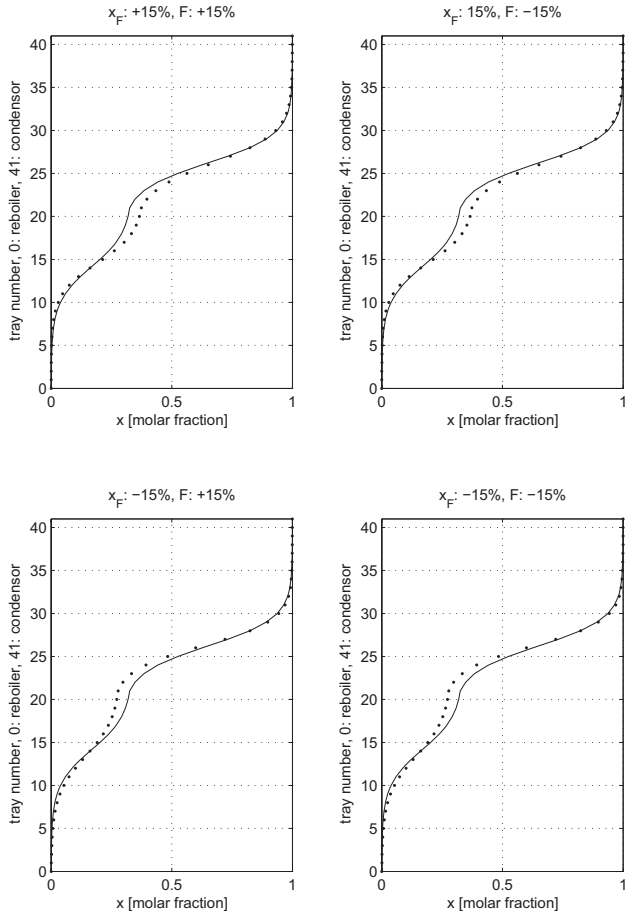


Fig. 5. Liquid concentration profiles in controlled distillation column. Solid lines: undisturbed stationary profile; dotted lines: new stationary profile for non-vanishing disturbances in feed flow rate F and feed concentration x_F (+/- 15% with respect to stationary values).

approach. *Automatica*, 33(4):669–673, 1997.

A. Rehm. *Control of Linear Descriptor Systems: A Matrix Inequality Approach*. Fortschr.-Ber. VDI Reihe 8 Nr. 1019, VDI Verlag, Düsseldorf, 2004.

A. Rehm and F. Allgöwer. General quadratic performance analysis and synthesis of differential algebraic equation (DAE) systems. *J. Proc. Contr.*, 12(4):467–474, 2002.

B. Retzbach. Control of an extractive distillation plant. In *IFAC Symposium Dynamics and Control of Chemical Reactors and Distillation Columns*, pages 225–230, Bournemouth, 1986.

C. W. Scherer, P. Gahinet, and M. Chilali. Multiobjective output - feedback control via LMI optimization. *IEEE Trans. Automat. Contr.*, AC-42(7):896–911, 1997.

S. Skogestad and I. Postlewaite. *Multivariable Feedback Control: Analysis and Design*. John Wiley and Sons, Chicester, GB, 1996.

S. Skogestad, M. Morari, and J.C. Doyle. Robust control of ill-conditioned plants: High-purity distillation. *IEEE Trans. Automat. Contr.*, AC-33(12):1092–1105, 1988.

Advances in Identification

Oral Session

Identification of low-order unstable process model from closed-loop step test

Tao Liu, Furong Gao*

*Department of Chemical and Biomolecular Engineering, Hong Kong University of Science & Technology, Kowloon, Hong Kong (Tel: +852-2358-7139; e-mail address: kefgao@ust.hk)

Abstract: Based on a closed-loop step response test, control-oriented low-order model identification algorithms are proposed for unstable processes. By using a damping factor to the closed-loop step response for realization of the Laplace transform, an algorithm for estimating the process frequency response is developed in terms of the closed-loop control structure used for identification. Correspondingly, two model identification algorithms are derived analytically for obtaining the widely used low-order process models of first-order-plus-dead-time (FOPDT) and second-order-plus-dead-time (SOPDT), respectively. Illustrative examples from the recent literature are used to demonstrate the effectiveness and merits of the proposed identification algorithms.

1. INTRODUCTION

As model-based control strategies have demonstrated apparently improved set-point tracking and load disturbance rejection for open-loop unstable processes, control-oriented identification of low-order process model, e.g., first-order-plus-dead-time (FOPDT) or second-order-plus-dead-time (SOPDT), has been increasingly explored in the process control community (Seborg, Edgar, and Mellichamp, 2003; Liu and Gao, 2008a and 2008b). For safety and economic reasons, unstable processes are usually not allowed to be operated in an open-loop manner. Closed-loop identification methods have been therefore studied in the literature. One of the mostly used identification tests is the close-loop step response test, owing to its implemental simplicity. Based on closed-loop step test in terms of the internal model control (IMC) structure, Häggblom K. E. (1996) demonstrated that closed-loop identification facilitates better representation of the process dynamic response characteristics for closed-loop operation. Using a conventional proportional (P) or proportional-integral-derivative (PID) controller for closed-loop stabilization, recent closed-loop step identification methods can be seen in the references (Paraskevopoulos, Pasgianos and Arvanitis, 2004; Sree and Chidambaram, 2006; Cheres, 2006). Using relay feedback to yield sustained oscillation within an admissible fluctuation of the process output, identification methods based on the resulting limit cycle data have been developed in the papers (Shiu, Hwang and Li, 1998; Marchetti, Scali and Lewin, 2001; Vivek and Chidambaram, 2005; Liu and Gao, 2008a). It was, however, pointed out that the conventional relay feedback structure cannot guarantee periodic oscillation for unstable processes with large time delay (Tan, Wang and Lee, 1998; Thyagarajan and Yu; 2003). Some limiting conditions to form the limit cycle from relay feedback for an unstable process have been disclosed by Liu and Gao (2008b). Besides, using the pseudo-random binary sequence (PRBS) as excitation signal to the set-point, closed-loop identification methods for application of model predictive control (MPC) have been reported in the references (Sung et al, 2001; Saffer and Doyle, 2002; Bindlish, Rawlings and Young, 2003).

In this paper, identification algorithms based on a closed-loop step test are proposed for obtaining low-order models of FOPDT and SOPDT for tuning unstable processes. By using a damping factor to the closed-loop step response for realization of the Laplace transform, an algorithm is first given to estimate the closed-loop frequency response in terms of using a conventional P, PI or PID controller for closed-loop stabilization. Accordingly, the process frequency response can be analytically derived from the closed-loop frequency response with the knowledge of the controller. Then, two identification algorithms are analytically developed for obtaining FOPDT and SOPDT models, respectively. Both the algorithms can give good accuracy if the model structure matches the process. Measurement noise tests are also performed to demonstrate identification robustness of the proposed algorithms.

2. FREQUENCY RESPONSE ESTIMATION

It is commonly known that the Fourier transform of a step response does not exist due to $\Delta y(t) \neq 0$ for $t \to \infty$, where $\Delta y(t) = y(t) - y(t_0)$ and $y(t_0)$ denotes the initial steady output. However, by substituting $s = \alpha + j\omega$ into the Laplace transform to the step response,

$$\Delta Y(s) = \int_0^\infty \Delta y(t) e^{-st} dt \tag{1}$$

we can formulate

$$\Delta Y(\alpha + j\omega) = \int_0^\infty [\Delta y(t) e^{-\alpha t}] e^{-j\omega t} dt \tag{2}$$

Note that if $\alpha > 0$, there exists $\Delta y(t) e^{-\alpha t} = 0$ for $t > t_N$, where t_N may be numerically determined using the condition of $\Delta y(t_N) e^{-\alpha t_N} \approx 0$, since $\Delta y(t)$ reaches a steady value after

the closed-loop transient response to a step change of the set-point. Therefore, by regarding α as a damping factor to the closed-loop step response for Laplace transform, we may compute $\Delta Y(\alpha + j\omega)$ from the N points of step response data as

$$\Delta Y(\alpha + j\omega) = \int_0^{t_N} [y(t)e^{-\alpha t}]e^{-j\omega t}dt \tag{3}$$

For a closed-loop step test with initial steady state, i.e., $y(t) = r(t) = c$ for $t \le t_0$, where $r(t)$ denotes the set-point value, c is a constant and t_0 is the time for step test, we may formulate the step change of the set-point by using a time shift of t_0 (i.e., letting $t_0 = 0$) as

$$\Delta r(t) = \begin{cases} 0, & t \le 0; \\ h, & t > 0. \end{cases} \tag{4}$$

where h is the magnitude of the step change. Its Laplace transform for $s = \alpha + j\omega$ with $\alpha > 0$ can be explicitly derived as

$$\Delta R(\alpha + j\omega) = \int_0^{\infty} he^{-(\alpha + j\omega)t}dt = \frac{h}{\alpha + j\omega} \tag{5}$$

Hence, the closed-loop frequency response can be derived using (3) and (5) as

$$T(\alpha + j\omega) = \frac{\alpha + j\omega}{h}\Delta Y(\alpha + j\omega), \ \alpha > 0 \tag{6}$$

Note that $T(\alpha + j\omega) \to 0$ as $\alpha \to \infty$. On the contrary, $\alpha \to 0$ will cause t_N much larger for computation of (6). A proper choice of α is therefore required for implementation. Considering that all the closed-loop transient response data to a step change of the set-point should be used to procure good estimation of the closed-loop frequency response, the following constraint is suggested to choose α,

$$\Delta y(t_{\text{set}})e^{-\alpha t_{\text{set}}} > \delta \tag{7}$$

where $\Delta y(t_{\text{set}})$ denotes the steady-state output deviation to the step change in terms of the settling time (t_{set}), and δ is a threshold of the computational precision that may be practically taken smaller than 1×10^{-6}. It follows from (7) that

$$\alpha < \frac{1}{t_{\text{set}}}\ln\frac{\Delta y(t_{\text{set}})}{\delta} \tag{8}$$

To ensure computational efficiency with respect to the complex variable, $s = \alpha + j\omega$, for frequency response estimation, the lower bound of α may be simply taken as δ, if there exists no limit on the time length of the step test.

Once α is chosen in terms of the above guideline, the time length, t_N, may be determined from a numerical constraint for computation of (3), i.e.,

$$\Delta y(t_N)e^{-\alpha t_N} < \delta \tag{9}$$

which can be solved as

$$t_N > \frac{1}{\alpha}\ln\frac{\Delta y(t_N)}{\delta} \tag{10}$$

Note that there exists the following Laplace transform with initial steady closed-loop state of $y(0) = r(0) = c$, where c is a constant,

$$L[\int_0^t \Delta y(t)dt] = \frac{\Delta Y(s)}{s} \tag{11}$$

To guarantee identification robustness against measurement noise, we may compute the frequency response by

$$T(\alpha + j\omega) = \frac{\dfrac{\Delta Y(\alpha + j\omega)}{\alpha + j\omega}}{\dfrac{\Delta R(\alpha + j\omega)}{\alpha + j\omega}} = \frac{(\alpha + j\omega)^2}{h}\int_0^{t_N}[\int_0^t \Delta y(\tau)d\tau]e^{-\alpha t}e^{-j\omega t}dt \tag{12}$$

It can be seen from (12) that, rather than use individual output data measured from the step test, a time integral for each measurement point is used to compute the outer-layer integral for obtaining the frequency response estimation. This facilitates reducing measurement errors according to the statistic averaging principle.

Denote the n-th order derivative for a complex function of $F(s)$ with respect to s as

$$F^{(n)}(s) = \frac{d^{(n)}}{ds^n}F(s), \ n \ge 1. \tag{13}$$

It follows from (3) and (6) that

$$T^{(1)}(s) = \frac{1}{h}\int_0^{\infty}(1 - st)\Delta y(t)e^{-st}dt \tag{14}$$

$$T^{(2)}(s) = \frac{1}{h}\int_0^{\infty}t(st - 2)\Delta y(t)e^{-st}dt \tag{15}$$

Hence, by letting $s = \alpha$ and choosing α as well as that for computation of (3), the single integral in (14) and (15) can be computed numerically. The corresponding time lengths of t_N can be respectively determined using the numerical constraints,

$$(1 - \alpha t_N)\Delta y(t_N)e^{-\alpha t_N} < \delta \tag{16}$$

$$t_N(\alpha t_N - 2)\Delta y(t_N)e^{-\alpha t_N} < \delta \tag{17}$$

For a conventional PID controller used in the closed-loop structure to stabilize an unstable process for step test,

$$C(s) = k_C(1 + \frac{1}{\tau_I s} + \frac{\tau_D s}{0.1\tau_D s + 1}) \tag{18}$$

where k_C denotes the controller gain, τ_I the integral constant and τ_D the derivative constant, it can be derived that

$$C^{(1)}(s) = k_C[-\frac{1}{\tau_I s^2} + \frac{\tau_D}{(0.1\tau_D s + 1)^2}] \tag{19}$$

$$C^{(2)}(s) = 2k_C[\frac{1}{\tau_I s^3} - \frac{0.1\tau_D^2}{(0.1\tau_D s + 1)^3}] \tag{20}$$

Note that the closed-loop transfer function can be derived as

$$T(s) = \frac{G(s)C(s)}{1 + G(s)C(s)} \tag{21}$$

It follows from (21) that

$$G(s) = \frac{T(s)}{C(s)[1 - T(s)]} \tag{22}$$

Its first and second derivatives can be derived accordingly as

$$G^{(1)} = \frac{T^{(1)}C + C^{(1)}T(T - 1)}{C^2(1 - T)^2} \tag{23}$$

$$G^{(2)} = \frac{CT^{(2)} + 2C^{(1)}T^{(1)}T + C^{(2)}T(T - 1)}{C^2(1 - T)^2}$$

$$-\frac{2[CT^{(1)}+C^{(1)}T(T-1)][CC^{(1)}(1-T)-C^2T^{(1)}(1-T)]}{C^3(1-T)^3} \quad (24)$$

Therefore, by substituting $s=\alpha+j\omega_k$ ($k=1,2,\ldots,M.$), where M is the number of representative frequency response points in a user specified frequency range, the process frequency response can be numerically estimated for model fitting.

3. MODEL IDENTIFICATION ALGORITHMS

Low-order unstable process models of FOPDT and SOPDT are respectively in the form of

$$G_1(s)=\frac{k_p e^{-\theta s}}{\tau_p s-1} \quad (25)$$

$$G_2(s)=\frac{k_p e^{-\theta s}}{(\tau_1 s-1)(\tau_2 s+1)} \quad (26)$$

where k_p denotes the process gain, θ the process time delay and τ_p (or τ_1 and τ_2) the process time constant(s).

By regarding $s\in\mathbb{R}$ and taking the natural logarithm for both sides of (25) in terms of $0<s<\tau_p$, we obtain

$$\ln[-G_1(s)]=\ln(k_p)-\ln(1-\tau_p s)-\theta s \quad (27)$$

Subsequently, taking the first and second derivatives for both sides of (27) with respect to s yields

$$\frac{1}{G_1(s)}\frac{d}{ds}[G_1(s)]=\frac{\tau_p}{1-\tau_p s}-\theta \quad (28)$$

$$Q_2(s)=\frac{\tau_p^2}{(1-\tau_p s)^2} \quad (29)$$

where $Q_2(s)=d[Q_1(s)]/ds$ and $Q_1(s)$ is the left side of (28).

Substituting $s=\alpha$ into (29), it can be derived that

$$\tau_p=\begin{cases}\dfrac{\sqrt{Q_2}}{\alpha\sqrt{Q_2}-1} \text{ or } \dfrac{\sqrt{Q_2}}{\alpha\sqrt{Q_2}+1}, & \text{if } Q_2>\dfrac{1}{\alpha^2};\\ \dfrac{\sqrt{Q_2}}{\alpha\sqrt{Q_2}+1}, & \text{if } Q_2\le\dfrac{1}{\alpha^2}.\end{cases} \quad (30)$$

Note that for $Q_2\ge 1/\alpha^2$, we may determine a suitable solution based on model fitting accuracy for the closed-loop step response.

It should be noted that the above parameter estimation is without loss of generality since there exists $\alpha<\tau_p$ in general, which may be verified from the guideline for choosing α as given in the earlier section.

Consequently, the other two model parameters can be derived from (28) and (25) using $s=\alpha$ as

$$\theta=-Q_1(\alpha)+\frac{\tau_p}{1-\tau_p\alpha} \quad (31)$$

$$k_p=(\tau_p\alpha-1)G_1(\alpha)e^{\alpha\theta} \quad (32)$$

Hence, the above algorithm named **Algorithm-I** for obtaining a FOPDT model for an unstable process can be summarized as:

(i) Choose $s=\alpha$ and t_N to compute $T(\alpha)$, $T^{(1)}(\alpha)$ and $T^{(2)}(\alpha)$ in terms of (6) (or (12)), (14) and (15);
(ii) Compute $C(\alpha)$, $C^{(1)}(\alpha)$ and $C^{(2)}(\alpha)$ in terms of (18), (19) and (20);
(iii) Compute $G_1(\alpha)$, $G_1^{(1)}(\alpha)$ and $G_1^{(2)}(\alpha)$ in terms of (22), (23) and (24);
(iv) Compute $Q_1(\alpha)$ and $Q_2(\alpha)$ in terms of (28) and (29);
(v) Compute the process time constant, τ_p, from (30);
(vi) Compute the process time delay, θ, from (31);
(vii) Compute the process gain, k_p, from (32).

Following a similar procedure as above, taking the natural logarithm for both sides of (26) in terms of $0<s<\tau_1$, yields

$$\ln[-G_2(s)]=\ln(k_p)-\ln(1-\tau_1 s)-\ln(\tau_2 s+1)-\theta s \quad (33)$$

Accordingly, the first and second order derivatives for both sides of (33) with respect to s can be derived respectively as

$$\frac{1}{G_2(s)}\frac{d}{ds}[G_2(s)]=\frac{\tau_1}{1-\tau_1 s}-\frac{\tau_2}{\tau_2 s+1}-\theta \quad (34)$$

$$Q_2(s)=\frac{\tau_1^2}{(1-\tau_1 s)^2}+\frac{\tau_2^2}{(\tau_2 s+1)^2} \quad (35)$$

where $Q_2(s)=d[Q_1(s)]/ds$ and $Q_1(s)$ is the left side of (34).

Substituting $s=\alpha$ into (35) yields

$$\begin{aligned}Q_2(\alpha)=&[2\alpha^2-\alpha^4Q_2(\alpha)]\tau_1^2\tau_2^2+[2\alpha-2\alpha^3Q_2(\alpha)](\tau_1^2\tau_2-\tau_1\tau_2^2)\\&+4\alpha^2Q_2(\alpha)\tau_1\tau_2+[1-\alpha^2Q_2(\alpha)](\tau_1^2+\tau_2^2)+2\alpha Q_2(\alpha)(\tau_1-\tau_2)\end{aligned} \quad (36)$$

To solve τ_1 and τ_2 from (36), we reformulate (36) in the LS form of

$$\psi(\alpha)=\phi(\alpha)^T\gamma \quad (37)$$

where

$$\begin{cases}\psi(\alpha)=Q_2(\alpha),\\ \phi(\alpha)=[2\alpha^2-\alpha^4Q_2(\alpha),\ 2\alpha-2\alpha^3Q_2(\alpha),\ -\alpha^2Q_2(\alpha),\ 1,\ 2\alpha Q_2(\alpha)]^T,\\ \gamma=[\tau_1^2\tau_2^2,\ \tau_1^2\tau_2-\tau_1\tau_2^2,\ \tau_1^2+\tau_2^2-4\tau_1\tau_2,\ \tau_1^2+\tau_2^2,\ \tau_1-\tau_2]^T.\end{cases} \quad (38)$$

By choosing 5 different values of α in terms of the guideline given in (8) and denoting $\Psi=[\psi(\alpha_1),\ \psi(\alpha_2),\ldots,\ \psi(\alpha_5)]^T$ and $\Phi=[\phi(\alpha_1),\ \phi(\alpha_2),\ldots,\ \phi(\alpha_5)]^T$, an LS solution can be derived from the linear regression,

$$\gamma=(\Phi^T\Phi)^{-1}\Phi^T\Psi \quad (39)$$

It is obvious that all the columns of Φ are linearly independent with each other, such that Φ is guaranteed non-singular for computation of (39). Accordingly, there exists a unique solution of γ for parameter estimation.

Then, the model parameters can be retrieved from γ as

$$\begin{cases}\tau_1=\dfrac{\gamma(5)}{2}+\dfrac{1}{2}\sqrt{\gamma^2(5)+4\dfrac{\gamma(2)}{\gamma(5)}}\\ \tau_2=\tau_1-\gamma(5)\end{cases} \quad (40)$$

Note that there exist three redundant fitting conditions in the parameter estimation of γ, which can be surely satisfied if the model structure matches the process to be identified. To procure fitting accuracy for a high-order process, we may use $\gamma(1)$, $\gamma(3)$ and $\gamma(4)$ together with $\gamma(2)$ and $\gamma(5)$ to derive

an LS fitting solution for parameter estimation in terms of using the natural logarithm, i.e.,

$$\begin{bmatrix} 2 & 2 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \ln \tau_1 \\ \ln \tau_2 \end{bmatrix} = \begin{bmatrix} \ln \gamma(1) \\ \ln[\dfrac{\gamma(4)-\gamma(3)}{4}] \\ \ln[\dfrac{\gamma(5)}{2}+\dfrac{1}{2}\sqrt{\gamma^2(5)+4\dfrac{\gamma(2)}{\gamma(5)}}] \\ \ln[\dfrac{\gamma(5)}{2}+\dfrac{1}{2}\sqrt{\gamma^2(5)+4\dfrac{\gamma(2)}{\gamma(5)}}-\gamma(5)] \end{bmatrix} \tag{41}$$

Consequently, the other two model parameters can be derived from (34) and (26) as

$$\theta = -Q_1(\alpha) + \frac{\tau_1}{1-\tau_1\alpha} - \frac{\tau_2}{\tau_2\alpha+1} \tag{42}$$

$$k_p = (\tau_1\alpha - 1)(\tau_2\alpha + 1)G_2(\alpha)e^{\alpha\theta} \tag{43}$$

Hence, the above algorithm named **Algorithm-II** for obtaining an SOPDT model for an unstable process can be summarized as:

(i) Choose $s = \alpha$ and t_N to compute $T(\alpha)$, $T^{(1)}(\alpha)$ and $T^{(2)}(\alpha)$ in terms of (6) (or (12)), (14) and (15);
(ii) Compute $C(\alpha)$, $C^{(1)}(\alpha)$ and $C^{(2)}(\alpha)$ in terms of (18), (19) and (20);
(iii) Compute $G_2(\alpha)$, $G_2^{(1)}(\alpha)$ and $G_2^{(2)}(\alpha)$ in terms of (22), (23) and (24);
(iv) Compute $Q_1(\alpha)$ and $Q_2(\alpha)$ in terms of (34) and (35);
(v) Compute the time constants, τ_1 and τ_2, from (40) (or (41));
(vi) Compute the process time delay, θ, from (42);
(vii) Compute the process gain, k_p, from (43).

4. ILLUSTRATION

Example 1. Consider the FOPDT unstable process studied in the recent literature (Padhy and Majhi, 2006),

$$G(s) = \frac{1}{s-1}e^{-0.8s}$$

Based on relay feedback test with two P controllers, Padhy and Majhi (2006) derived a FOPDT model, $G_m = 1.0e^{-0.8033s}/(1.0007s-1)$. For illustration, the unity feedback control structure with a proportional controller of $k_c = 1.2$, which is equivalent to that of Padhy and Majhi (2006), is used for closed-loop step test. By adding a step change with a magnitude of $h = 0.05$ to the set-point, the closed-loop step response is shown in Fig.1. According to the guidelines given in (8) and (10), $\alpha = 0.1$ and $t_N = 150$ (s) are chosen to use the proposed Algorithm-I, resulting in a FOPDT model listed in Table 1, which indicates high accuracy. The fitting error is given in terms of the closed-loop transient response in the time interval [0, 50]s.

To demonstrate identification robustness against measurement noise, assume that a random noise of $N(0, \sigma_N^2 = 0.035\%)$, causing the noise-to-signal ratio (NSR) to 5%, is added to the output measurement which is then used for feedback control. By performing 100 Monte-Carlo tests in terms of varying the 'seed' of the noise generator from 1 to 100, the identified results are listed in Table 1, where the model parameters are respectively the mean of 100 Monte-Carlo tests, and the values in the adjacent parentheses are the sample standard deviation. The results for the noise levels of NSR=10% and 20% are also listed in Table 1 to show the achievable identification accuracy and robustness.

Example 2. Consider the SOPDT unstable process studied by Cheres (2006) and Sree and Chidambaram (2006),

$$G(s) = \frac{1}{(2s-1)(0.5s+1)}e^{-0.5s}$$

Based on a closed-loop step test in terms of a PID controller ($k_c = 2.71$, $\tau_I = 4.43$ and $\tau_D = 0.319$) and a unity step change to the set-point, Cheres (2006) derived only a referential FOPDT model for controller tuning, and so was done in Sree and Chidambaram (2006). By performing the same closed-loop step test, the proposed Algorithm-II based on the choice of $\alpha = 0.1, 0.15, 0.2, 0.25, 0.3$ and $t_N = 300$ (s) gives a SOPDT model listed in Table 1, again demonstrating good accuracy. The fitting error is given in terms of the closed-loop transient response in the time interval [0, 30]s.

To demonstrate identification robustness against measurement noise, 100 Monte Carlo tests are performed in terms of NSR=5%, 10% and 20%, respectively. The identified results are listed in Table 1 for comparison, which indicate again that good identification accuracy and robustness is therefore obtained.

5. CONCLUSIONS

Low-order model identification methods have been increasingly appealed for improving control system design to operate unstable processes. By applying a damping factor to the closed-loop step response for realization of the Laplace transform, a frequency response estimation algorithm has been proposed for model fitting. Based on the process frequency response estimated from the closed-loop step response with the knowledge of the controller, two model identification algorithms have been analytically developed for obtaining the widely used low-order process models of FOPDT and SOPDT for practical applications. Two illustrative examples from the recent literature have been performed to demonstrate the achievable accuracy of the proposed algorithms. The results under Monte Carlo noise tests have also demonstrated good identification robustness of the proposed algorithms.

ACKNOWLEDGEMENT

This work is supported in part by the Hong Kong Research Grants Council under Project No. 613107.

REFERENCES

Åström, K.J. and T. Hägglund (2005). *Advanced PID Control*. Research Triangle Park, NC, ISA Society of America.

Bindlish, R., J.B. Rawlings and R.E. Young (2003). Parameter estimation for industrial polymerization processes. *AIChE J.*, **49**, 2071-2078.

Cheres, E. (2006). Parameter estimation of an unstable system with a PID controller in a closed loop configuration. *Journal of the Franklin Institute*, **343**, 204-209.

Häggblom, K.E. (1996). Combined internal model and inferential control of a distillation column via closed-loop identification. *Journal of Process Control*, **6**, 223-232.

Liu, T. and F. Gao (2008a). Alternative identification algorithms for obtaining a first-order stable/unstable process model from a single relay feedback test. *Ind. Eng. Chem. Res.*, **47**, 1140-1149.

Liu, T. and F. Gao (2008b). Identification of integrating and unstable processes from relay feedback. *Computers and Chemical Engineering*, **32**, 3038-3056.

Marchetti, G., C. Scali and D.R. Lewin (2001). Identification and control of open-loop unstable processes by relay methods. *Automatica*, **37**, 2049-2055.

Paraskevopoulos, P.N., G.D. Pasgianos and K.G. Arvanitis (2004). New tuning and identification methods for unstable first order plus dead-time. *IEEE Trans. Control Systems Technology*, **12**, 455-464.

Padhy, P.K. and S. Majhi (2006). Relay based PI-PD design for stable and unstable FOPDT processes. *Computers and Chemical Engineering*, **30**, 790-796.

Saffer, D.R. and F.J. Doyle (2002). Closed-loop identification with MPC for an industrial scale CD-control problem. *IET Control Theory & Appl.*, **149**, 448-456.

Seborg, D.E., T.F. Edgar and D.A. Mellichamp (2003). *Process Dynamics and Control, 2nd ed.*, John Wiley & Sons, New Jersey.

Shiu, S.J., S.H. Hwang and M.L. Li (1998). Automatic tuning of systems with one or two unstable poles. *Chemical Engineering Communication*, **167**, 51-72.

Sree, R.P. and M. Chidambaram (2006). Improved closed loop identification of transfer function model for unstable systems. *Journal of the Franklin Institute*, **343**, 152-160.

Sung, S.W., S.Y. Lee, H.J. Kwak and I.-B. Lee (2001). Continuous-time subspace system identification method. *Ind. Eng. Chem. Res.*, **40**, 2886-2896.

Tan, K. K., Q.G. Wang and T.H. Lee (1998). Finite spectrum assignment control of unstable time delay processes with relay tuning. *Ind. Eng. Chem. Res.*, **37**, 1351-1357.

Thyagarajan, T. and C.C. Yu (2003). Improved autotuning using the shape factor from feedback. *Ind. Eng. Chem. Res.*, **42**, 4425-4440.

Vivek, S. and M. Chidambaram (2005). An improved relay auto tuning of PID controllers for unstable FOPTD systems. *Computers and Chemical Engineering*, **29**, 2060-2068.

Yu, C.C. (2006). *Autotuning of PID Controllers: A Relay Feedback Approach, 2nd ed*. Springer-Verlag, London.

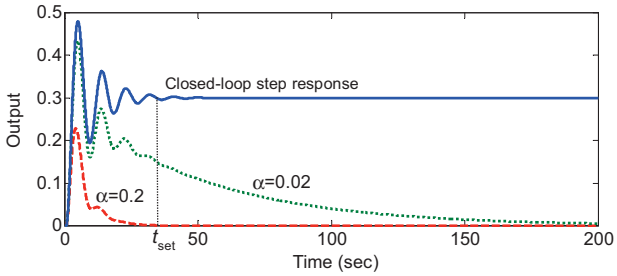


Fig. 1. Illustration of choosing α for example 1

Table 1. Closed-loop step response identification under different measurement noise levels

NSR	Example 1	*err*	Example 2	*err*
0	$\frac{1.0000e^{-0.8008s}}{1.0006s-1}$	8.46×10^{-8}	$\frac{0.9999e^{-0.4996s}}{(2.0000s-1)(0.5000s+1)}$	1.47×10^{-8}
5%	$\frac{1.0001(\pm0.0011)e^{-0.8019(\pm0.045)s}}{1.0015(\pm0.033)s-1}$	4.15×10^{-7}	$\frac{0.9999(\pm0.0025)e^{-0.4998(\pm0.0082)s}}{[2.0002(\pm0.0037)s-1][0.5004(\pm0.0098)s-1]}$	1.72×10^{-7}
10%	$\frac{1.0003(\pm0.0021)e^{-0.8018(\pm0.089)s}}{1.0013(\pm0.066)s-1}$	5.08×10^{-7}	$\frac{0.9999(\pm0.0054)e^{-0.4992(\pm0.023)s}}{[2.0005(\pm0.0087)s-1][0.5009(\pm0.034)s-1]}$	7.61×10^{-7}
20%	$\frac{1.0006(\pm0.0041)e^{-0.7971(\pm0.183)s}}{0.9973(\pm0.136)s-1}$	2.86×10^{-6}	$\frac{0.9995(\pm0.0097)e^{-0.4985(\pm0.026)s}}{[2.0008(\pm0.017)s-1][0.5011(\pm0.026)s-1]}$	2.79×10^{-6}

Multivariable System Identification for Integral Controllability – Computational Issues

Mark L. Darby, Michael Nikolaou

Chemical & Biomolecular Engineering Department, University of Houston, Houston, TX 77204-4004 USA
(darbymark@sbcglobal.net, nikolaou@uh.edu)

Abstract: A process model satisfies the integral controllability (IC) condition if the model can be used in a model-based controller that can be arbitrarily detuned without jeopardizing closed-loop stability. For decoupling multivariable control this requirement is equivalent to the inequality $\mathrm{Re}\left[\lambda\left(\mathbf{G}\hat{\mathbf{G}}^{-1}\right)\right]>0$ for the actual and estimated process steady-state gain matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$. This necessitates experiments for identification of $\hat{\mathbf{G}}$ that satisfies the IC inequality. In this work we explore, via computer simulations, computational issues related to the design of such experiments for an FCC process. The proposed approach is based on a general mathematical optimization framework we presented in prior work.

Keywords: Identification, Integral controllability, multivariable systems

1. INTRODUCTION

A process model satisfies the integral controllability (IC) condition if the model can be used in a model-based controller that can be arbitrarily detuned without jeopardizing closed-loop stability. For decoupling multivariable control this requirement is equivalent to the inequality

$$\mathrm{Re}\left[\lambda\left(\mathbf{G}\hat{\mathbf{G}}^{-1}\right)\right]>0 \tag{1}$$

for the actual and estimated process steady-state gain matrices $\mathbf{G}$ and $\hat{\mathbf{G}}$ (Garcia and Morari 1985). The problem is acute for ill-conditioned processes. This necessitates experiments for identification of a model $\hat{\mathbf{G}}$ that satisfies the IC inequality. "The main weakness of the eigenvalue conditions [eqn. (1)] is that they consist of a coupling between the plant model and the true plant which is highly cumbersome for use in robust control analysis and design." (Featherstone and Braatz 1998b). A number of attempts have been made to address this weakness. Featherstone and Braatz (1998a) showed that for processes with constant rotation matrices in the singular-value decomposition (svd) of their transfer matrix the problem reduces to D-optimal design of experiments. Using insightful geometric reasoning to ensure IC for general linear 2×2 systems, Koung and MacGregor (1993) introduced experiment designs in terms of rotated PRBS input vectors, with power of each component of the rotated input vector reciprocally proportional to the corresponding singular value of $\hat{\mathbf{G}}$. Koung & MacGregor (1994) heuristically extended these design rules to $n\times n$ multivariable systems. The same rules were also used by Bruwer & MacGregor (2006) for the design of identification experiments subject to input and output bounds in the time domain. Darby and Nikolaou (2008) showed that the design rules proposed by Koung and MacGregor (1993; 1994) accept the same deep theoretical justification for both 2×2 and $n\times n$ systems in a number of cases. However, Darby and Nikolaou (2008) also showed that these design rules are not optimal for a number of typical cases, such as when outputs and/or inputs are constrained or when input rather than output variance alone must be maintained at a minimum. Furthermore, the same authors provided rigorous design rules for optimal inputs in a number of such cases. These design rules from solution of corresponding optimization problems. The purpose of this article is to explore the nature of the optimal input designs produced by the mathematical framework introduced by Darby and Nikolaou (2008) when applied to a realistic system, such as a 5×5 fluid catalytic cracking (FCC) unit.

2. BACKGROUND: EXPERIMENT DESIGN FOR IC

Consider a stable, linear, time-invariant, multivariable system with steady-state input-output relationship

$$\mathbf{y}=\mathbf{G}\mathbf{m} \tag{2}$$

where $\mathbf{y},\mathbf{m}\in\mathbb{R}^n$, $\mathbf{G}$ and $\hat{\mathbf{G}}\triangleq[\hat{\mathbf{g}}_1,...,\hat{\mathbf{g}}_n]^T\in\mathbb{R}^{n\times n}$. Because the IC condition, eqn. (1), involves the real process $\mathbf{G}$ and identified model $\hat{\mathbf{G}}$, it cannot directly guide input design for an $n\times n$ system. The following results (Darby and Nikolaou 2008) avoid that difficulty and can be used directly to design experiments pursuing IC.

Theorem 1. Experiment design for IC. Let the model uncertainty matrix $\mathbf{D} = \mathbf{G} - \hat{\mathbf{G}} \in \mathbb{R}^{n\times n}$ *belong to the ellipsoidal uncertainty set*

$$D \triangleq \left\{ [\mathbf{d}_1 ... \mathbf{d}_n]^T \in \mathbb{R}^{n\times n} \middle| \mathbf{d}_k^T \mathbf{M}^T \mathbf{M} \mathbf{d}_k \le c^2, 1 \le k \le n \right\}. \quad (3)$$

Then, an experiment design guarantees IC if the resulting information matrix $\mathbf{M}^T\mathbf{M}$ *and identified model* $\hat{\mathbf{G}}$ *satisfy the inequality*

$$\sum_{k=1}^{n} a_k \sqrt{\hat{\mathbf{v}}_k^T \left(\mathbf{M}^T\mathbf{M}\right)^{-1} \hat{\mathbf{v}}_k} < 1. \quad (4)$$

where

$$a_k \triangleq c \left\| \hat{\mathbf{u}}_k \right\|_1 / \hat{\sigma}_k, \quad k = 1,...,n \quad (5)$$

and $\hat{\sigma}_1 \ge \cdots \ge \hat{\sigma}_n$, $\hat{\mathbf{u}}_k$, $\hat{\mathbf{v}}_k$ *are defined through the svd*

$$\hat{\mathbf{G}} = \hat{\mathbf{U}}\hat{\Sigma}\hat{\mathbf{V}}^T = \sum_{k=1}^{n} \hat{\sigma}_k \hat{\mathbf{u}}_k \hat{\mathbf{v}}_k^T. \quad (6)$$

Eqn. (4) clearly suggests that IC can be satisfied if the information matrix $\mathbf{M}^T\mathbf{M}$ is "large enough". Given bounds on the input vector $\mathbf{m}$, a large enough $\mathbf{M}^T\mathbf{M}$ can be achieved if (a) the identification experiment is run long enough, or (b) $\mathbf{m}$ is shaped appropriately. While the first alternative is straightforward, it is far less desirable than the second one. Therefore, the essence of experiment design for IC is how to *shape process inputs that satisfy eqn. (4) subject to relevant constraints*. Darby and Nikolaou (2008) showed that numerical or analytical solutions can be developed for a number of cases. While for some of these cases the resulting designs are similar to designs that have appeared in literature, for others the resulting designs are entirely different.

2.1. Analytical solutions

Suppose that the quantity $\sum_{k=1}^{n} a_k [\hat{\mathbf{v}}_k^T (\mathbf{M}^T\mathbf{M})^{-1} \hat{\mathbf{v}}_k]^{1/2}$ in eqn. (4) is to be minimized with respect to the zero-mean random input $\mathbf{m}$, subject to the total weighted variance inequality

$$x \operatorname{var}(\mathbf{y}) + (1-x) \operatorname{var}(\mathbf{m}) \le W^2 \quad (7)$$

where $0 \le x \le 1$ and $0 < a_1 < ... < a_n$. Then it can be shown that the optimal input vector $\mathbf{m}$ is

$$\mathbf{m} = \hat{\mathbf{V}}\boldsymbol{\xi}, \quad (8)$$

where $\boldsymbol{\xi}$ is a zero-mean multivariable PRBS with

$$E\left[\xi_k^2\right] = \left(\frac{a_k}{b_k^2}\right)^{2/3} \frac{W^2}{\left[\sum_{j=1}^{n} \left(a_j b_j\right)^{2/3}\right]^2}, \quad k = 1,...,n \quad (9)$$

$$b_k^2 \triangleq x\sigma_k^2 + 1 - x, \quad k = 1,...,n. \quad (10)$$

and $b_1 \ge ... \ge b_n > 0$. Reversing the role of the above objectives and constraints, the minimum of the cost function $x \operatorname{var}(\mathbf{y}) + (1-x) \operatorname{var}(\mathbf{m})$ subject to eqn. (4) can be shown to be attained at an optimal $\mathbf{m}$ satisfying eqn. (8) with

$$E\left[\xi_k^2\right] > \left(\frac{a_k}{b_k^2}\right)^{2/3} \left[\sum_{j=1}^{n} \left(a_j b_j\right)^{2/3}\right]^2, \quad k = 1,...,n. \quad (11)$$

Note that for $x = 1$ (all cost on output variance, as is desirable in early stages of an identification experiment) both eqns. (9) and (11) result in the well known design rule

$$r_{kj} \triangleq \sqrt{\frac{E\left[\xi_k^2\right]}{E\left[\xi_j^2\right]}} = s_{kj} \frac{\hat{\sigma}_j}{\hat{\sigma}_k} \approx \frac{\hat{\sigma}_j}{\hat{\sigma}_k}. \quad (12)$$

where $s_{kj} \triangleq (\| \hat{\mathbf{u}}_k \|_1 / \| \hat{\mathbf{u}}_j \|_1)^{1/3} \approx 1$ for most cases of practical interest, and $1/n^{1/6} \le s_{kj} \le n^{1/6}$ when $\hat{\mathbf{u}}_k$, $\hat{\mathbf{u}}_j$ are any orthonormal vectors in $\mathbb{R}^n$. However, for $x = 0$ (all cost on input variance) we get the new input design

$$r_{kj} \triangleq \sqrt{\frac{E\left[\xi_k^2\right]}{E\left[\xi_j^2\right]}} = s_{kj} \left(\frac{\hat{\sigma}_j}{\hat{\sigma}_k}\right)^{1/3} \approx \left(\frac{\hat{\sigma}_j}{\hat{\sigma}_k}\right)^{1/3}. \quad (13)$$

The above design would keep inputs small to avoid inadvertent loss of IC by failure to excite the process by inputs along directions corresponding to small singular values.

Finally, a D-optimal design subject to eqns. (4) and (7) can be shown to be attained, if feasible, at an optimal $\mathbf{m}$ as in eqn. (8) with

$$E\left[\xi_k^2\right] = W^2 / \left(n b_k^2\right), \quad k = 1,...,n, \quad (14)$$

if eqn. (4) is satisfied by the above ξ_k, or each $E\left[\xi_k^2\right]$ equal to the unique positive solution of the equation

$$\frac{\rho}{t-1}\left(\frac{a_k}{\sqrt{(t-1)E\left[\xi_k^2\right]}} - \frac{1}{n}\right) = b_k^2 E\left[\xi_k^2\right] - \frac{W^2}{n} \quad (15)$$

for a value of $\rho > 0$ such that eqn. (4) is satisfied. It has been shown that eqn. (15) guarantees that $E\left[\xi_{k+1}^2\right] > E\left[\xi_k^2\right]$ and that D-optimality is compatible with IC by Cauchy's inequality (Darby and Nikolaou 2008). Note that as $t \to \infty$ eqns. (14) and (15) coincide asymptotically.

2.2. Numerical solutions

The preceding section 2.1 summarized analytical solutions for simple cases, offering insight into the nature of corresponding solutions. However, in many practical situations individual constraints on m_i and y_i may be present, such as

$$\sum_{\tau=1}^{t} m_{i,\tau}{}^2 = [\mathbf{M}^T\mathbf{M}]_{ii} \le (t-1) M_i^2 \quad (16)$$

$$\sum_{\tau=1}^{t} y_{i,\tau}{}^2 = [\mathbf{Y}^T\mathbf{Y}]_{ii} = [\hat{\mathbf{G}}\mathbf{M}^T\mathbf{M}\hat{\mathbf{G}}^T]_{ii} \le (t-1)Y_i^2 \quad (17)$$

corresponding to bounds on the variance of individual inputs m_k or outputs y_k. In such cases a numerical solution is required. To obtain a numerical solution, assume a zero-mean input vector $\mathbf{m}$, approximate the information matrix as $\mathbf{M}^T\mathbf{M} \approx (t-1)\mathbf{C}_m$, parametrize the input covariance matrix $\mathbf{C}_m$ in terms of the triangular matrix $\mathbf{Q}$ through the Cholesky factorization $\mathbf{C}_m = \mathbf{Q}\mathbf{Q}^T$, and substitute $\mathbf{M}^T\mathbf{M}$ into eqn. (4), to get

$$\frac{c}{\sqrt{t-1}} \underbrace{\sum_{i=1}^{n} \frac{\|\hat{\mathbf{u}}_k\|_1}{\hat{\sigma}_k} \sqrt{\hat{\mathbf{v}}_k^T \left(\mathbf{Q}\mathbf{Q}^T\right)^{-1} \hat{\mathbf{v}}_k}}_{\beta} \triangleq \frac{c}{\sqrt{t-1}} \beta < 1. \quad (18)$$

(Other parametrizations of a symmetric matrix in terms of corresponding basis matrices could be used. This is subject of ongoing investigation.) We can then design experiments for IC using eqn. (18) as a constraint or by minimizing β with respect to $\mathbf{Q}$ subject to input and output constraints such as in eqns. (16) and (17). Then, the corresponding optimal input $\mathbf{m}$ is

$$\mathbf{m} = \mathbf{Q}_{\text{opt}}\mathbf{z} \quad (19)$$

where $\mathbf{z}$ is a zero-mean PRBS with $\text{cov}(\mathbf{z}) = \mathbf{I}$. Even though β is not convex, resulting design problems are not prohibitively large for realistic systems, as demonstrated in section 4. It should be stressed that minimizing β may result to neither uncorrelated rotated inputs $\boldsymbol{\xi}$, nor magnitudes of rotated input components reciprocally proportional to corresponding singular values of the steady-state gain matrix, eqn. (12). In fact, the advantage of the above numerical formulation is that no such underlying assumptions on the nature of optimal inputs are necessary. Rather, the numerical optimization determines the nature of optimal input designs.

3. SUMMARY OF PROPOSED APPROACH

a. Establish constraints commensurate with time t available for identification experiments.

b. Obtain preliminary estimates of $\hat{\mathbf{G}}$ and c^2.

c. Compute the svd of $\hat{\mathbf{G}}$ to get $\hat{\mathbf{U}}$, $\hat{\boldsymbol{\Sigma}}$, $\hat{\mathbf{V}}^T$, eqn. (6).

Case I

d. Compute $\lambda_{k,\text{opt}}$ via eqns. (9) or similar

e. Design $\mathbf{m} = \hat{\mathbf{V}}\boldsymbol{\xi}$ with $\boldsymbol{\xi}$ zero-mean PRBS and $\text{cov}(\boldsymbol{\xi}) = \text{diag}(\lambda_{\ell_1,\text{opt}}^2, ..., \lambda_{\ell_n,\text{opt}}^2)/(t-1)$

Case II

d. Compute $\mathbf{Q}_{\text{opt}} = \arg\min \beta$ subject to constraints.

e. Design $\mathbf{m} = \mathbf{Q}_{\text{opt}}\mathbf{z}$ with $\mathbf{z}$ zero-mean PRBS and $\text{cov}(\mathbf{z}) = \mathbf{I}$ (eqn. (19)).

f. Implement $\mathbf{m}$ and collect data, to update $\hat{\mathbf{G}}$ and c^2.

g. If $\hat{\mathbf{G}}$ is adequate, stop. Otherwise go to step c.

4. CASE STUDY

A steady-state gain matrix is obtained from a linear empirical dynamic model of an industrial reactor-regenerator from a FCC unit, identified from plant testing (Harmse 2007). Note that the specific inputs and outputs are not indicated. Scaling is performed according to the inverse of the typical operating ranges of the inputs and outputs. The resulting gain matrix is

$$\mathbf{G} = \begin{bmatrix} 0.3866 & 0.0 & 0.1192 & 0.0 & 0.0630 \\ 0.0 & -0.6935 & 1.5463 & -0.1311 & -0.2462 \\ 0.0 & 0.0 & 0.5225 & -0.1298 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.1058 & 0.0 \\ 0.0 & -0.5803 & -0.3669 & -0.2057 & -0.4435 \end{bmatrix} \quad (20)$$

We tested the designs shown in Table 1. The constraints shown in Table 2 were considered.

Table 1. Summary of experimental designs tested

Design	Objective	Constraints
ICmin	$\min_{\mathbf{Q}} \beta$	$\mathbf{Q}$ triangular
KM	(Koung and MacGregor 1994)	
PRBSmax	$\min_{\mathbf{C}_m} -\log(\det \mathbf{C}_m)$	$\mathbf{C}_m \triangleq \text{diag}(v_i) \succ \mathbf{0}$

Table 2. Constraints considered for experiment design

Design case	Bounds	
A	$\text{var}(m_i) \le \infty$	$\text{var}(y_1) \le 0.50$
	$i = 1,...,5$	$\text{var}(y_2) \le 0.47$
		$\text{var}(y_3) \le 0.44$
		$\text{var}(y_4) \le 0.41$
		$\text{var}(y_5) \le 0.38$
B	$\text{var}(m_i) \le 0.5$	$\text{var}(y_i) \le 0.5$
	$i = 1,...,5$	

In all simulations, parameter estimation is initiated at time step 5 and is performed at each subsequent time step. The IC condition, eqn. (1), is calculated based on the true gain matrix and the inverse of the gain estimate at each time step. For simulations of case A, independent Gaussian noise of zero-mean and unit variance is added to all outputs.

Realizations of the inputs and outputs for each design are shown in Figure 1. We see that due to the high noise levels, the actual output variances are significantly higher than the optimal output variances (which are based on the model without noise). Note that the relatively low signal-to-noise ratio allows us to observe the evolution of gains over a longer period of time.

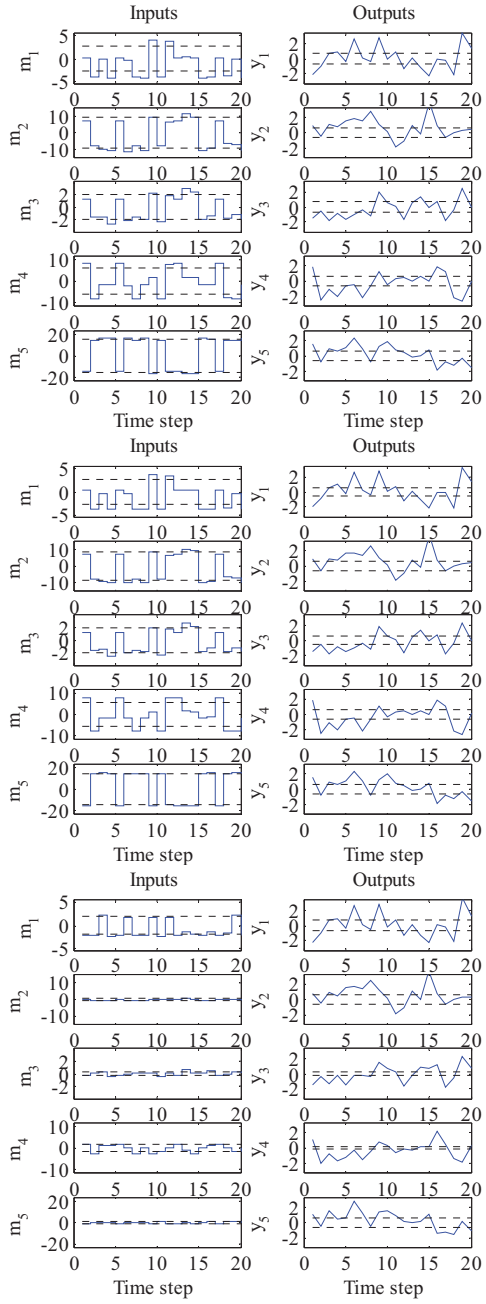


Figure 1 – Example realization of reactor regenerator for the three designs of case A (Table 1). The dotted lines represent $\pm(\mathrm{var}(m_i)^{\mathrm{opt}})^{1/2}$ and $\pm(\mathrm{var}(y_i)^{\mathrm{opt}})^{1/2}$ values.

Trends of the gain errors and an indicator of the IC condition are shown in Figure 2 For this realization, we see that IC is achieved first by ICmin at time step 7, followed by KM at time step 8, and finally PRBSopt at time step 13. The evolution of the gain errors is similar for ICmin and KM, whereas the PRBSopt results show higher gain errors and slower error reduction over time, consistent with the lower value of det(cov(**m**)) for PRBSopt.

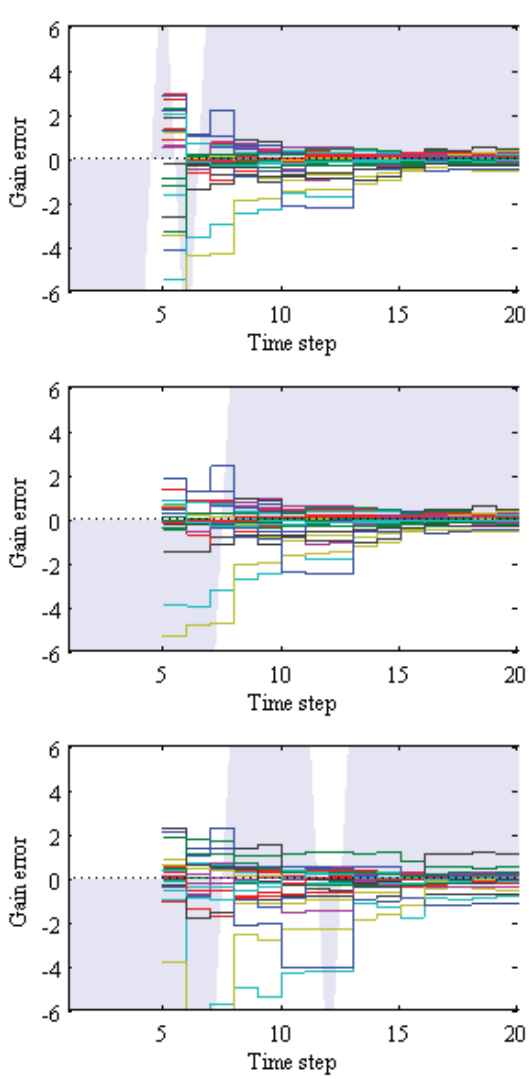


Figure 2 – Gain errors and IC condition for example realization of design case A (Table 1). Satisfaction or violation of the IC condition corresponds to shading above or below 0, respectively.

For the simulations of case B, independent, Gaussian noise of zero mean and variance (0.15^2) is added to all outputs. Realizations of the inputs and outputs for each design are shown in Figure 3. Note that the smaller signals for the KM design, compared to the other two designs, are due to the

fact that inputs must satisfy all inequality constraints, as well as equality constraints on rotated input ratios. The latter are clearly not optimal for case B.

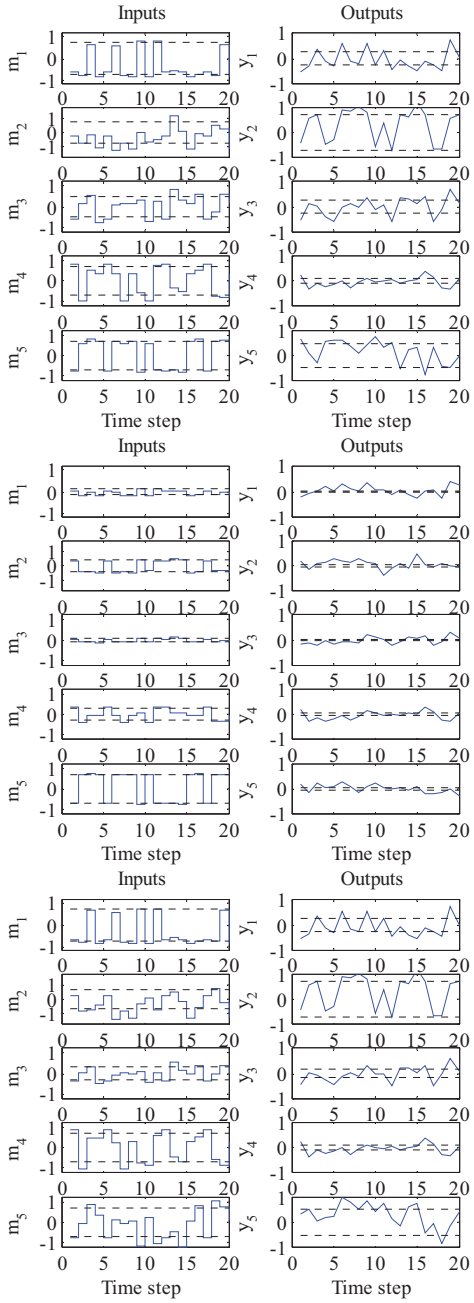


Figure 3 – Example realization of reactor regenerator case B (Table 1). The dotted lines represent $\pm\sqrt{\mathrm{var}(m_i)^{\mathrm{opt}}}$ and $\pm\sqrt{\mathrm{var}(y_i)^{\mathrm{opt}}}$ values.

Trends of the gain errors and an indicator of the IC condition are shown in Figure 4. We see that the parameters estimates for the KM design are significantly inferior to the ICmin and PRBSopt designs, reflecting the much smaller value of det(cov(**m**)) for KM. Further, while IC is achieved by ICmin at time step 6 and by PRBSopt at time step 5, the KM design does not satisfy the IC condition by time step 20 (it is actually achieved at time step 22 – not shown – more then 3 times longer than required for either ICmin or PRBSopt).

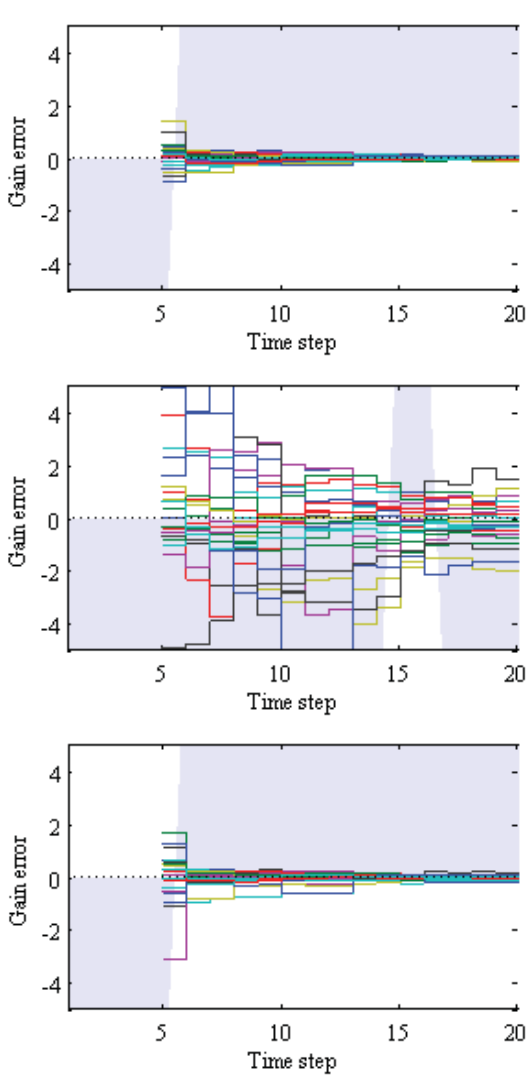


Figure 4 – Gain errors and IC condition for example realization of design case B (Table 1). Satisfaction or violation of the IC condition corresponds to shading above or below 0, respectively.

To develop these designs we used the Matlab routine fmincon with multiple starting points, to reduce the possibility of missing a global optimum for non-convex optimization problems. Run time for all simulations was, on the average, of the order of 0.25 seconds when using the Yalmip interface, and of the order of 0.1 seconds without it.

5. CONCLUSIONS

The purpose of this paper was to explore numerical aspects of a numerical optimization approach proposed in prior work for the design of experiments targeting IC. Given that analytical solutions for this approach are available only for some cases, it is natural to ask how well numerical optimization can work, given that the proposed problems are non-convex. In this work, we develop what appear to be useful designs for a 5×5 multivariable FCC unit. Convergence time appeared not to be an issue. These results suggest that the proposed approach can work reasonably well for problems of that size. Clearly, other optimization methods (either deterministic or probabilistic) can be explored.

Acknowledgments

Partial support from the National Science Foundation (Grant No. CTS-0227232) is gratefully acknowledged.

6. REFERENCES

Bruwer, M. J. and J. F. MacGregor (2006). "Robust multi-variable identification: Optimal experimental design with constraints." Journal Of Process Control **16**(6): 581-600.

Darby, M. L. and M. Nikolaou (2008). "Multivariable system identification for integral controllability." Automatica: submitted.

Featherstone, A. P. and R. D. Braatz (1998a). "Input Design for Large-Scale Sheet and Film Processes." Industrial & Engineering Chemistry Research **37**: 449-454.

Featherstone, A. P. and R. D. Braatz (1998b). "Integrated robust identification and control of large-scale processes." Industrial & Engineering Chemistry Research **37**(1): 97-106.

Garcia, C. E. and M. Morari (1985). "Internal model control: 2. design procedure for multivariate systems." Ind. Eng. Chem. Process Des. Dev. **24**: 472-484.

Harmse, M. (2007). "Personal communication."

Koung, C. W. and J. F. MacGregor (1993). "Design Of Identification Experiments For Robust-Control - A Geometric Approach For Bivariate Processes." Industrial & Engineering Chemistry Research **32**(8): 1658-1666.

Koung, C. W. and J. F. MacGregor (1994). "Identification For Robust Multivariable Control - The Design Of Experiments." Automatica **30**(10): 1541-1554.

Internal Excitation Approaches for Closed-loop Identification of Processes Controlled by MPC

Oscar A. Z. Sotomayor[1], Darci Odloak[2]

[1]*Department of Electrical Engineering*
Center for Exact Sciences and Technology, Federal University of Sergipe (UFS), BRAZIL
E-mail: oscars@ufs.br
[2]*Department of Chemical Engineering*
Polytechnic School of the University of São Paulo (EPUSP), BRAZIL
E-mail: odloak@usp.br

Abstract: This work is concerned with the model re-identification of processes controlled by MPC systems. The MPC system considered here has a two-layer structure, where in the upper layer a steady-state optimization algorithm determines a set of optimal targets for the process inputs and passes this set to the MPC controller that determines the best way to drive the process to such targets. This is the case of several commercial MPC packages applied in industry. In this paper, it is proposed two internal excitation approaches aiming to obtain closed-loop data sufficiently rich for process identifiability. Here, the term internal is used to indicate that the excitation test signal is applied within the MPC control structure. In the first excitation approach, the test signal is introduced as a weighting factor in the objective function of the target calculation layer. In the second approach, the test signal is injected as a dither signal in the objective function of the dynamic controller layer. These two approaches are compared to the usual method where the excitation signal is added to the controller output. The application of the methodologies is illustrated through numerical simulations carried out on a depropanizer column of the oil industry. The results show the effectiveness of the proposed approaches and their good potential to be applied in practice.

1. INTRODUCTION

Model identification has become a bottleneck of MPC technology. It is the most expensive, difficult and time-consuming step of the MPC project. Although industrial processes present nonlinear dynamic characteristics, typically, empirical linear and time-invariant (LTI) models based on process input-output data obtained in open-loop operation are used in MPC implementation (Qin & Badgwell, 2003). While this approach is only acceptable at operating conditions around the operating point where the model was obtained, the control system with this model works satisfactorily in most applications.

However, after some operation time (2-3 years), MPC is seldom performing as when it was commissioned. The main cause of this problem is related to the model deterioration resulting from changes in the dynamics of the plant or persistent unmeasured disturbances that force the plant to a different operating point (Conner & Seborg, 2005). Changes in the dynamics of the plant may result from fatigue conditions, fouling, debottlenecking, etc, or changes in the operating conditions or product specifications. In general, the above listed problems intensify with time and tend to accentuate the plant/model mismatch, leading to poor output prediction and, therefore, degradation of the control system performance. In order to keep the performance of the MPC at an acceptable level, it is essential to carry out the MPC re-commissioning in a periodic basis, which means to re-identify the process model and, if necessary, to retune the MPC considering the new model (Gugaliya et al., 2005). However, due to production goals and safety aspects, model re-identification means, in most cases, to develop a new model based on plant data obtained in closed-loop conditions.

Closed-loop identification is a research subject with growing interest in the last decade (Van den Hof, 1998; Forssell & Ljung, 1999; Hjalmarsson, 2005). Important aspects on model identification have been studied and several identification strategies have been proposed, which can be categorized as variants of the following three approaches (Forssell & Ljung, 1999): direct, indirect and joint input-output methods. Both indirect and joint input-output methods require prior knowledge of the controller or assume that it has a certain LTI structure. Obviously, these methods are not suitable for MPC applications, because MPC presents nonlinear and time-variant features, especially when operating under constraints. For this sort of control strategy, the direct method is the recommended choice for closed-loop identification. See for example Rivera & Flores (1999).

In closed-loop identification, the use of routine operating data would be an ideal goal. But, the inherent reduction in the excitation resulting from the presence of the controller may result in a poor signal-to-noise ratio. In this case, and in order to achieve the necessary and sufficient conditions for process identifiability, an external persistently exciting (PE) test signal is required. External excitation is a dither signal that may be introduced on the controlled variable set-point and/or on the manipulated variable (added to the controller output).

However, adding such a signal is often undesirable or too expensive, and there is no guarantee that the process constraints and product specifications will be attended during the execution of the excitation procedure. On the other hand, an insufficient excitation may compromise the identification requirements.

The main goal of this paper is to compare internal excitation approaches that exploit the two-layer structure of MPC packages. Motivation for this work is due primarily to commercial needs and as an attempt to overcome the significant gap between practical applications and theory in closed-loop identification with MPC. In the proposed methodologies an external PE test signal is applied within the MPC control structure: (a) the test signal is introduced in the objective function of the target calculation layer, and (b) the test signal is injected in the objective function of the dynamic MPC control layer. These approaches does not modify the optimization code and they allows the adequate excitation of the process coupled with the continuous operation of the system as the process constraints and product specification can be satisfied during the test. Results from the proposed excitation methodologies are also compared with the one provided by a conventional external excitation procedure in the closed-loop identification of an industrial depropanizer column.

2. THE TWO-LAYER MPC STRUCTURE

In modern processing plants, MPC control systems are usually implemented in a two-layer scheme (Ying & Joseph, 1999; Qin & Badgwell, 2003; Nikandrov & Swartz, 2009). The two-layer MPC considered here is shown in Figure 1. The upper layer usually corresponds to a simplified steady-state target optimization and the lower layer stands for a constrained dynamic optimization in which the outputs are controlled in specified zones or ranges instead of fixed references. It is at the dynamic layer where the main control objectives (setpoint tracking, disturbance rejection) are pursued. All commercial MPC packages offer the option of zone control. In the target calculation layer, one searches for the optimum steady-state values to the system (input targets), by usually solving a linear or quadratic objective function subject to bound constraints in the inputs and outputs. The outputs at the optimal steady-state are computed through a static model, consistent with the dynamic MPC model, and the available steady-state output prediction computed at the previous time step in the MPC algorithm. The optimal input targets are sent to the dynamic layer, where the control cost is extended with a term that penalizes the distance between the present value of the input and the optimal target. Both layers are executed with the same sampling period.

In this paper, the target calculation layer solves a LP (linear programming) problem where the objective may be to maximize production by forcing one or more inputs to their bounds, while keeping the outputs inside the bounds:

$$\min_{\Delta u_s, \delta_y} W_1^T \Delta u_s + \left| W_2^T \delta_y \right| \tag{1}$$

subject to:

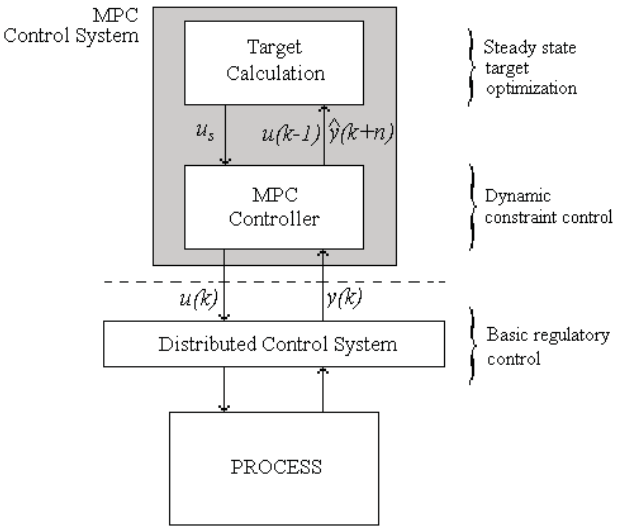


Fig. 1. Two-layer MPC structure

$$\begin{aligned} &\Delta u_s = u_s - u(k-1) \\ &y_s = G_0 \Delta u_s + \hat{y}(k+n) \\ &u_{\min} \le u_s \le u_{\max} \\ &y_{\min} \le y_s + \delta_y \le y_{\max} \end{aligned} \tag{2}$$

where $u(k-1)$ is the last implemented control action, k is the present time, u_s is the vector of steady-state targets for the manipulated inputs, y_s is the vector of predicted output at steady-state, $\hat{y}(k+n)$ is the prediction of the controlled output at time instant $k+n$ (n is the model horizon or settling time of the process in open-loop) computed at time k, δ_y is the vector of slack variables for the controlled outputs, G_0 is the steady state gain matrix model, W_1 and W_2 are weight vector of appropriate dimensions, $u_{\min}$ and $u_{\max}$ are the bounds of the manipulated inputs, $y_{\min}$ and $y_{\max}$ are the bounds of the controlled outputs.

As a result of the solution to the problem formulated in (1) and (2), it is obtained the input target u_s that is passed to the dynamic layer, which typically solves the following QP (quadratic programming) problem:

$$\begin{aligned} \min_{\Delta u} \sum_{i=1}^{p} \left\| Q\left(\hat{y}(k+i) - y_{sp}\right) \right\|_2^2 + \sum_{j=1}^{m} \left\| R\Delta u(k+j-1) \right\|_2^2 + \\ \sum_{j=1}^{m} \left\| R_u \left(u(k+j-1) - u_s\right) \right\|_2^2 \end{aligned} \tag{3}$$

subject to:

$$\begin{aligned} &-\Delta u_{\max} \le \Delta u(k+j-1) \le \Delta u_{\max}, \quad j = 1, \cdots, m \\ &u_{\min} \le u(k+j-1) \le u_{\max}, \qquad j = 1, \cdots, m \end{aligned} \tag{4}$$

with $u(k+j-1) = u(k-1) + \sum_{i=1}^{j} \Delta u(k+i-1)$

where $\hat{y}(k+i)$ is the output prediction at time $k+i$, y_{sp} is the setpoint to the system output, $\Delta u = \left[\Delta u(k)^T \quad \Delta u(k+1)^T \quad \cdots \quad \Delta u(k+m-1)^T \right]^T$ is the vector of control moves, $\Delta u_{\max}$ is the upper limit to the control moves, m is the control horizon, p is the prediction horizon, and Q, R and R_u are diagonal weighting matrices of appropriate dimensions. Note that only the first element of the computed input sequence Δu is implemented in the plant, i.e. $u(k) = \Delta u(k) + u(k-1)$.

In this controller, the zone control strategy is implemented as follows: if the prediction of a given output is inside its reference zone or range, the error in this output is considered to be equal to zero and the output is not included in the controller optimization problem. When the output prediction lies outside the corresponding reference range, depending on whether the prediction is above or below the max or min values of this range, one of these bounds is adopted as the output reference. In general, the zone control strategy is used as an attempt to release some degrees of freedom to allow the inputs to approach their optimal targets (constraints pushing) and to smooth out the system response. For more details see Sotomayor et al. (2009).

The two-layer MPC algorithm as described above is similar to the structure of several MPC packages widely applied to control the refining and petrochemical processes. For instance, this MPC algorithm, with slight modifications, is supported by the advanced control package SICON[@], which is the standard process control software in the oil refineries of PETROBRAS in Brazil.

3. INTERNAL EXCITATION APPROACHES

To solve the problem of lack of excitation during normal operation of MPC systems, some authors have proposed a new class of excitation methods for MPC that can be considered as internal excitation methods. Genceli & Nikolaou (1996) use the MPCI framework (model predictive control and identification) where the PE characteristic of the inputs is imposed as a constraint in the optimization problem related to the MPC. The drawback of this method is that the additional constraint is non-convex, resulting in a non-convex optimization problem. Since solving non-convex problems is significantly more involved than solving convex problems, the additional non-convex constraints are undesirable and the method cannot be directly applied to existing commercial MPC packages, without extensive modifications of the controller code, which limits its practical application.

On the other hand, a reasonable consideration of the layered MPC is that when the model is biased, the target calculation layer will change the input target to the dynamic layer quite often, and so, the input target could be viewed as a possible test signal. However, to assume that these natural moves on the input targets will be PE is a questionable matter. Here, taking advantage of the layered structure of MPC, and in order to guarantee the necessary excitation of the input targets, an external PE test signal, namely a binary signal of magnitude ± 1, is applied within the MPC control structure according to the following approaches:

Method 1. Introducing the test signal into the LP layer

Given that the weight vector W_1 in the objective function of the target calculation layer is usually available to be set on-line by the user of the MPC package, the external test signal can be introduced in the MPC system as a variable that multiplies W_1. Thus, Equation (1) is re-written as follows:

$$\min_{\Delta u_s, \delta_y} \left(W_{exc} \otimes W_1 \right)^T \Delta u_s + \left| W_2^T \delta_y \right| \tag{5}$$

subject to (2)

where W_{exc} is a vector whose elements are the components of the binary test signal and operator $\otimes$ denotes the Schur (or element-by-element) product. Then, if the product $W_{exc,i} W_{1,i}$ is positive (negative), the solution to the LP problem will tend to reduce (increase) input u_i until it reaches its lower (upper) bound or the output predictions lies outside the control zones. Note that if W_{exc} is set equal to a vector of ones, then the excitation procedure will end and the original objective function (1) is recovered. This method is described with details in Sotomayor et al. (2009).

Method 2. Introducing the test signal into the dynamic layer

In this case, the external test signal, conveniently scaled, is injected as a dither signal in the input target u_s that enters the MPC layer. This is similar to writing the objective function (3) as follows:

$$\begin{aligned} \min_{\Delta u} & \sum_{i=1}^{p} \left\| Q\left(\hat{y}(k+i) - y_{sp} \right) \right\|_2^2 + \sum_{j=1}^{m} \left\| R \Delta u(k+j-1) \right\|_2^2 + \\ & \sum_{j=1}^{m} \left\| R_u \left(u(k+j-1) - (u_s + u_{s,dith}) \right) \right\|_2^2 \end{aligned} \tag{6}$$

with $u_{s,dith} = \lambda W_{exc}$, subject to (4), where W_{exc} is the excitation vector as defined in Section 3.1 and λ is a scaling factor. Then, the achievement of excitation of the system will largely depend on the value of tuning parameter R_u, which will define if the MPC layer will implement the input target faster than the main process dynamics. Observe that if λ is set equal to zero, the excitation procedure is ended and the original cost function (3) is obtained.

Particularly, Method 1 can be easily implemented in existing MPC packages with a structure similar to the one detailed in Section 2, as the excitation signal is introduced through a tuning parameter of the controller. In both methods the MPC problem is still solved through a QP, and the problem is reduced to design the binary sequence for W_{exc} such that the on-line solution of the problems (5)-(2) and (6)-(4) produces persistent excited inputs, which is the primary requirement for the process identifiability (Ljung, 1999).

As it will be shown in the application section, with the approaches proposed here, the inputs can be adequately excited and if the outputs are controlled within zones, the feedback effect on the test data may be minimized. Also, the approaches attend the process safety requirements and the product specifications can be satisfied adequately.

In the next section it is illustrated the application of the proposed excitation procedures to the closed-loop identification of a depropanizer column. The identification procedure follows same steps as the usual identification methodology applied to industrial processes (Ljung, 1999): design of the test signal and generation of dataset, model structure selection, computation of the model parameters and model validation.

4. APPLICATION: DEPROPANIZER COLUMN

Figure 2 presents the process considered in this work. It is an industrial depropanizer column of the FCC unit at the PETROBRAS Refinery of Cubatão (RPBC), Brazil.

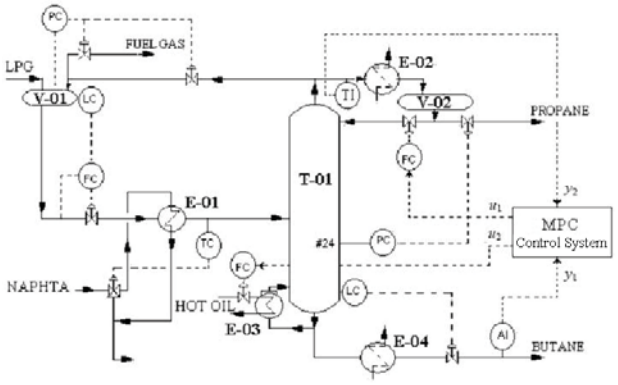


Fig. 2. Schematic diagram of the depropanizer column

In the depropanizer column, the C3 stream (propane and propene) is separated from a C4 stream (butane and butene). The operation of this process is controlled by a commercial MPC system as detailed in Section 2. Basically, it is a 2x2 control system, where the output variables y_1 and y_2 are the molar concentration (%) of C3 in the bottom stream and the temperature (°C) at the first stage of the top section of the column, respectively. The input variables u_1 and u_2 are the reflux flowrate to the top of the column (m^3/d) and the flowrate of hot oil to the reboiler (m^3/d), respectively. Transfer function models of order 2 corresponding to points FD and 1 from Porfirio et al. (2003) are used to simulate the "true" process and to represent the nominal process model (as it is used by the MPC system), respectively. MPC tuning parameters are here omitted but they can be found in Sotomayor et al. (2009).

PE test signal and generation of the dataset

Based on the guidelines provided by Zhu (2001) and a priori knowledge of the process (already existing model in the MPC), two independent GBN (generalized binary noise) (Tulleken, 1990) signals of magnitude ± 1 are designed. These test signals are applied directly to the LP layer of the MPC system if Method 1 is used or they are firstly scaled using $\lambda_1 = 0.065 u_{s,1}$ and $\lambda_2 = 0.07 u_{s,2}$, respectively, and applied to the MPC dynamic layer if excitation Method 2 is employed. In addition to the these excitation approaches, the system is also perturbed using a conventional external excitation as in MacArthur &. Zhan (2007). For this purpose, the MPC controller first calculates the normal movement for each controller output. New projected outputs are computed by superimposing a dither signal on the moves proposed by the controller. The projected outputs are then compared to the controller's high and low limits (constraints). Projected moves are then modified to ensure that all constraints are honored. In the present case, the dither signals are the GBN test signals scaled to ± 0.0034 and ± 0.0043, respectively. In all the cases, the duration of the excitation test is 4500 min. The data were collected with a sampling time of 1 min, resulting 4500 samples of input-output data.

For better identification results, the dataset is normalized, de-trended and filtered. Next, the dataset is divided into two subsets, where the first one containing 3000 data points is used to identify the model while the second one containing the remaining points is used to cross-validate the model. The PE characteristic of the inputs for the three cases is tested for order $\eta = 4$, which means that 2nd-order transfer functions can be satisfactory identified (Söderström & Stoica, 1989).

Model structure selection

In the present paper, it is considered that the model structure is defined by a continuous-time multi-input and single-output (MISO) output-error (OE) transfer function model, with the stochastic model parameterized as unitary:

$$\hat{y}_j(t,\rho_j) = \sum_{i=1}^{n_u} \hat{G}_{j,i}(s,\rho_{j,i}) u_i(t) + \varepsilon_j(t), \quad j = 1,\cdots,n_y \quad (7)$$

where $\hat{G}_{j,i}(s,\rho_{j,i})$ is the (j,i)th transfer function defined as:

$$\hat{G}_{j,i}(s,\rho_{j,i}) = \frac{\hat{B}_{j,i}(s)}{\hat{A}_{j,i}(s)} = \frac{\sum_{k=0}^{m_{j,i}} \hat{b}_{j,i,k} s^k}{\sum_{k=0}^{n_{j,i}} \hat{a}_{j,i,k} s^k} e^{-s\hat{\theta}_{j,i}}$$

with $\hat{a}_{j,i,n_{j,i}} = 1$, $n_{j,i} \geq m_{j,i}$, where $u(t)$ is the input vector, $\hat{y}(t,\rho)$ is the model output, $\hat{\theta}_{j,i}$ is the estimated time-delay between the i_{th} input and the j_{th} output, $\varepsilon(t)$ is the residual or total model error (bias plus variance), n_u and n_y are the number of inputs and outputs, respectively, and $\rho_{j,i} = \left[\hat{b}_{j,i,m_{j,i}} \cdots \hat{b}_{j,i,0}\ \hat{a}_{j,i,(n_{j,i}-1)} \cdots \hat{a}_{j,i,0}\ \hat{\theta}_{j,i}\right]^T \in \mathbb{R}^{q_{j,i}}$, with $q_{j,i} = n_{j,i} + m_{j,i} + 2$, where $n_{j,i}$ and $m_{j,i}$ denotes the denominator and numerator orders of $\hat{G}_{j,i}(s,\rho_{j,i})$, respectively.

Computation of the model parameters

The goal is to build a model as defined in eq. (7) based on closed-loop sampled data, focusing on the parameters of each transfer function $\hat{G}_{j,i}(s,\rho_{j,i})$ rather than on the model error. Thus, the task of the identification procedure is to compute the vector of model parameters:

$$\rho_j = \left[\rho_{j,1}^T \cdots \rho_{j,mu}^T \right]^T \in \mathbb{R}^{q_j \times 1}, \quad q_j = \sum_{i=1}^{n_u} q_{j,i} \tag{8}$$

Here it is used the CONTSID toolbox (Garnier et al., 2008) to find the vector ρ_j for each closed-loop sampled dataset from the depropanizer column, assuming $n_u = n_y = 2$, $n_{j,i} = 2$ and $m_{j,i} = 1$. The identification is carried out off-line considering the values of the parameters of the existing (old) process model as the initial solution to the identification problem.

Three new models are obtained and they are evaluated based on the following performance criteria:

$$FIT = 100 \times \left(1 - \frac{norm(y_j - \hat{y}_j)}{norm(y_j - mean(y_j))} \right)$$

$$R_T^2 = 1 - \frac{\mathrm{var}\left(y_j - \hat{y}_j\right)}{\mathrm{var}\left(y_j\right)}$$

where y is the true system output and $\hat{y}$ the model output. Coefficient FIT indicates the percentage of the output variation that can be associated to the model, while coefficient R_T^2 measures how well the model output explains the behavior of the system output, and this parameter will be close to 1 in low noise conditions.

Model validation

Figure 3 shows the step response comparison between the old model and the new models obtained with the three excitation methods considered here. Observing the responses of the old model used in the controller and the new models obtained with the re-identification procedure, one may conclude that the re-identification of the process model is quite justified not only because of the difference between the gains of the old and new models, but also because of the different dynamics. Also, observing the responses of the re-identified models, one can confirm that the model obtained with excitation method 3 is, in general, inferior to the models obtained with the two other excitation methods, showing a significant bias on the gain of $\hat{G}_{2,1}$. Moreover, from a practical point of view, one may conclude that the internal excitation methods 1 and 2 can be considered equivalent in terms of the model that is obtained, particularly if the step response of the process is to be used as in several MPCs. This result is in concordance with the performance indicators obtained from the cross-validation procedure of the three new models (not presented here).

5. CONCLUSIONS

Three closed-loop excitation methods for systems being controlled by MPCs with a two layer structure were studied here. These excitation methods allow the closed-loop model re-identification that should be used for periodic MPC monitoring and maintenance and for the design of an explicit adaptive MPC. The first two methods are based on the introduction of a persistently exciting signal within the MPC structure. The third method corresponds to the traditional approach of adding the excitation signal to the controller output. The three methods showed equivalence in terms of producing a data set that is adequate for model identification. The main difference between the excitation methods lies in the implementation of the approach in practice. Method 1 introduces the excitation signal in the coefficients of the objective function of the target calculation layer which are usually available as tuning parameters of the controller. So, there is no need of any modification in the controller code and the method does not require any particular attention of the operator while the process excitation is performed. Thus, this method seems to be the most adequate in practical terms. Method 2 adds the excitation signal to the input target of the dynamic layer of the controller. Besides, some new tuning parameters this method requires a slight modification in the controller code and, consequently, can only be implemented if the source code is available. The third or conventional method, that adds the excitation signal to the output of the controller, requires more care in the design stage and more attention of the operator because the control action really injected in the process will not satisfy the process constraints.

REFERENCES

Conner, J.S., D.E. Seborg (2005). Assessing the need for process re-identification. *Industrial & Engineering Chemistry Research*, **44**(8): 2767-2775.

Forssell, U., L. Ljung (1999). Closed-loop identification revisited. *Automatica*, **35**(7): 1215-1241.

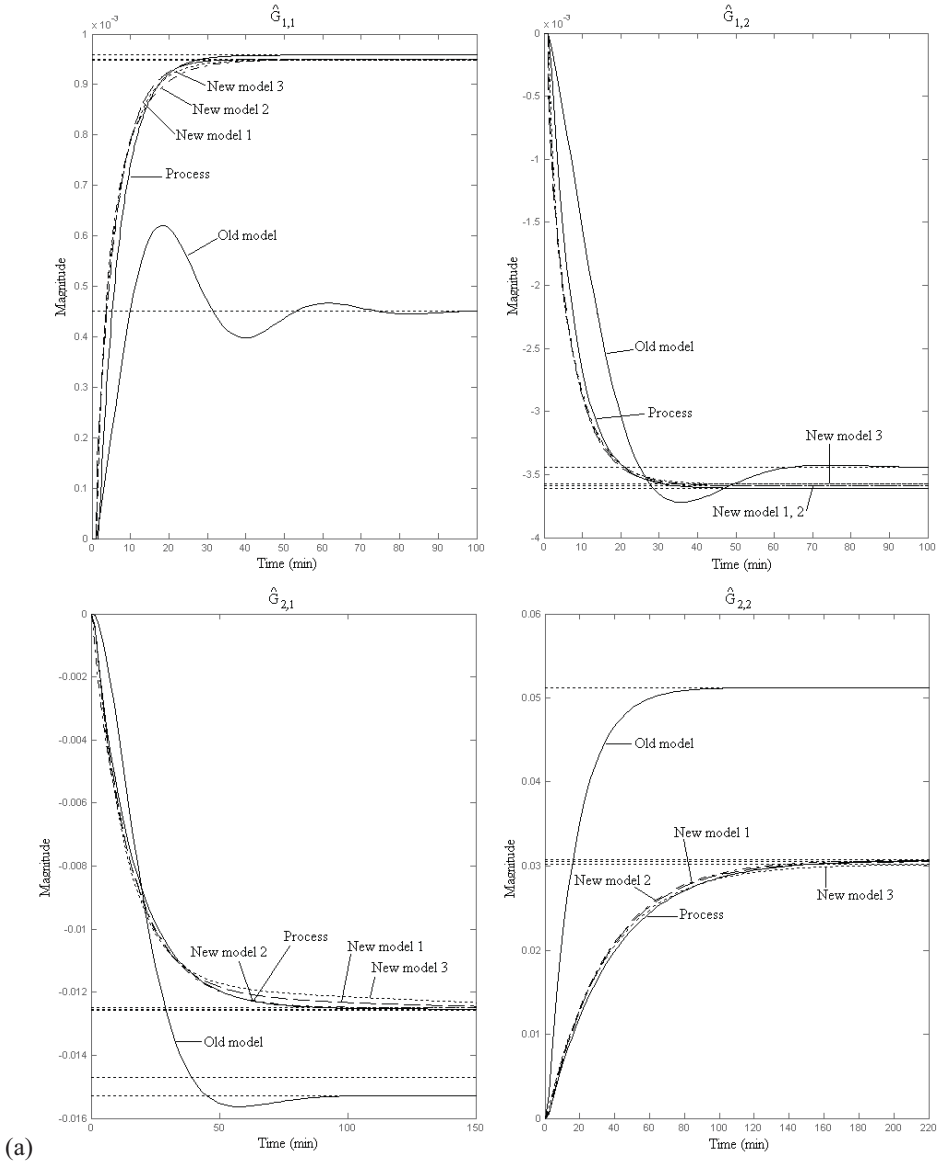
Garnier, H., M. Gilson, T. Bastogne, M. Mensler (2008). The CONTSID toolbox: A software support for continuous-time data-based modelling. In. H. Garnier & L. Wang (Eds.), *Identification of continuous-time models from sampled data*. Berlin: Springer.

Genceli, H., M. Nikolaou (1996). New approach to constrained predictive control with simultaneous model identification. *AIChE Journal*, **42**(10): 2857-2868.

Gugaliya, J.K., S.C. Patwardhan, R.D. Gudi, H. Raghavan, R. Yelchuru (2005). Model maintenance for industrial process control. In: *2005 AIChE Annual Meeting*, Cincinnati – OH.

Hjalmarsson, H. (2005). From experiment design to closed loop control. *Automatica*, **41**(3): 393-438.

Ljung, L. (1999). *System identification – Theory for the user* (2nd Ed.), Upper Saddle River, NJ: Prentice-Hall.



(a)

Fig. 3. Step response for the depropanizer column

MacArthur, J., C. Zhan (2007). A practical global multi-stage method for fully automated closed-loop identification of industrial processes. *Journal of Process Control*, **17**(10): 770-786.

Nikandrov, A., C.L.E. Swartz (2009). Sensitivity analysis of LP-MPC cascade control systems. *Journal of Process Contro,* **19**(1): 16-24.

Porfirio, C.R., E.A. Neto, D. Odloak (2003). Multi-model predictive control of an industrial C3/C4 splitter. *Control Engineering Practice*, **11**(7): 765-779.

Qin, S.J., T.A. Badgwell (2003). A survey of industrial model predictive control technology. *Control Engineering Practice*, **11**(7): 733-764.

Rivera, D.E., M.E. Flores (1999). "Plant-friendly" closed-loop identification using model predictive control. In: *1999 AIChE Annual Meeting*, Dallas – TX.

Söderström, T., P. Stoica (1989). *System identification*. Hemel Hempstead, UK: Prentice-Hall.

Sotomayor, O.A.Z, D. Odloak, L.F.L. Moro (2009). Closed-loop model re-identification of processes under MPC with zone control. *Control Engineering Practice*, **17**(5): 551-563.

Tulleken, H.J.A.F. (1990). Generalized binary noise test-signal concept for improved identification-experiment design. *Automatica*, **26**(1): 37-49.

Van den Hof, P.M.J. (1998). Closed loop issues in system identification. *Annual Reviews in Control*, **22**(1): 173-186.

Ying, C.-M., B. Joseph (1999). Performance and stability analysis of LP-MPC and QP-MPC cascade control systems. *AIChE Journal*, **45**(7): 1521-1534.

Zhu, Y. (2001). *Multivariable system identification for process control*. Oxford, UK: Elsevier Science Ltda.

Identification of Low Order Models for Large Scale Systems ⋆

Satyajit Wattamwar * **Siep Weiland** * **Ton Backx** *

* *All authors are with Dept. Of Electrical Engg. Technical University of Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands*
s.wattamwar@tue.nl

Abstract: In this paper we propose a novel procedure for obtaining a low order non-linear model of a large scale multi-phase, non-linear, reactive fluid flow systems. Our approach is based on the combinations of the methods of Proper Orthogonal Decomposition (*POD*), and non-linear System Identification (*SID*) techniques. The problem of non-linear model reduction is formulated as parameter estimation problem. In the first step POD is used to separate the spatial and temporal patterns and in the second step a model structure and it's parameters of linear and of non-linear polynomial type are identified to approximate the temporal patterns obtained by the POD in the first step. The proposed model structure treats POD modal coefficients as states rather than outputs of the identified model. The state space matrices which happens to be the parameters of a black-box to be identified, comes linearly in parameter estimation process. For the same reason, Ordinary Least Square (*OLS*) method is used to estimate the model parameters. The simplicity and reliability of the proposed method gives computationally very efficient linear and non-linear low order models for extremely large scale processes. The method is of generic nature. The efficiency of proposed approach is illustrated on a very large scale benchmark problem depicting Industrial Glass Manufacturing Process (*IGMP*). The results show good performance of the proposed method.

1. INTRODUCTION

Industrial processes involving fluid flows are usually modeled by Navier-Stoke's equations which are solved by some kind of spatial discretization. Due to this modeling approach they are referred to as Distributed Parameter Systems (DPS). Spatial discretization of DPS is done by means of Finite Volume or Finite Element methods and Galerkin or Petro-Galerkin projection techniques and they are simulated in a Computational Fluid Dynamic (CFD) software environment. Although such discretizations approximate the dynamic process behavior reasonably well, it leads to very large order process model. It takes huge computational efforts (time, CPU requirement) to simulate such models and therefore such process models can not be used for online plant optimization and control purposes. Model Order Reduction (MOR) is therefore an important step before proceeding to control design, see e.g. Shvartsman and Kevrekidis [1998]. The method of Proper Orthogonal Decomposition (POD) or Principle Component Analysis (PCA) is widely used for deriving lower dimensional models from the First Principle Model (FPM). The POD method searches for dominant patterns in the given process and defines an optimal, data-dependent basis, that is subsequently used as a projection space to infer reduced order models through Galerkin type of projections, see Astrid [2004] and the references therein. POD methods are empirical (data dependent) in nature and therefore these methods are susceptible to changes in process inputs and process parameters. The reduced model obtained by POD techniques with Galerkin projections are usually very dense and one loses the original sparse model structure. Such a dense model does not always give computational advantage over original full scale CFD model. This motivates one to look for other possible approaches which can give computationally efficient, reliable models which can be used for the online control and optimization purpose. Other motivations for the method proposed in this paper are that in many commercial CFD packages sometimes it becomes impossible to get access to the Partial Differential Equations (PDE) used to implement full scale model and even with access to the PDEs the reduced order modeling efforts can also be very expensive and laborious. In such situations one needs to explore the other possible ways to get a low order model by some identification method. One of such methods is explained in Wattamwar et al. [2008], which uses POD and system identification tools like N4SID algorithms, as explained in e.g. Overschee and Moor [1996], Favoreel et al. [2000]. But the method proposed there results in linear models which are not sufficient for approximating the non-linearities of large scale applications like IGMP. Moreover in the method proposed there, the states of the linear reduced order model have no physical meaning. These problems have motivated us to investigate another model reduction strategy which can approximate process non-linearity. The identification based approach proposed in this paper can be very useful, because it allows to use the available large-scale first principle based detailed non-linear process model in the form of commercial package, not just for the purpose of computationally extremely efficient dynamic

⋆ This work was supported by the European Union within the Marie-Curie Training Network PROMATCH under the grant number MRTN-CT-2004-512441.

process analysis but also for the purpose of design of the process controller and optimization. Therefore the method proposed here helps in minimizing the dependence on the expensive testing of the plant required for the controller design.

This paper is organized as follows. The overall methodology involving necessary tools from system theory like POD, a black-box type of system identification for linear and non-linear polynomial system is explained in section 2. The application/motivation is IGMP and is explained in the section 3. Some results of the proposed method on the motivation problem are presented in the section 4 which is followed by future work and references.

2. THEORY BACKGROUND AND METHODOLOGY

One of the most promising and successful techniques for an efficient reduction of large-scale nonlinear systems in fluid dynamics is the method of Proper Orthogonal Decompositions (POD) also known as the Karhunen-Loève method Holmes et al. [1996]. The method is based on the observation that flow characteristics reveal coherent structures or *patterns* in many processes in fluid dynamics. This has led to the idea that the solutions of model equations may be approximated by considering a small number of dominant coherent structures (called *modes* or *basis*) that are inferred in an *empirical* manner from measurements or simulated data. Given an ensemble of K measurements $\mathbf{T}^k(\cdot), k = 1, \ldots, K$ with each measurement defined on some spatial domain Ω, the POD method amounts to assuming that each observation $\mathbf{T}^k$ belongs to a Hilbert space $\mathcal{H}$ of functions defined on Ω. With the inner product defined on $\mathcal{H}$, it then makes sense to call a collection $\{\varphi_j\}_{j=1}^{\infty}$ an *orthonormal basis* of $\mathcal{H}$ if any element, say $\mathbf{T} \in \mathcal{H}$, admits a representation

$$\mathbf{T}(z) = \sum_{j=1}^{\infty} a_j \varphi_j(z), \quad z \in \Omega \tag{1}$$

Here, the a_j's are referred to as the *modal coefficients*(MC) and the φ_j's are the *modes* or *basis* of the expansion. The truncated expansion

$$\mathbf{T}_n(z) = \sum_{j=1}^{n} a_j \varphi_j(z), \quad z \in \Omega \tag{2}$$

causes an approximation error $\|\mathbf{T} - \mathbf{T}_n\|$ in the norm of the Hilbert space. We will call $\{\varphi_j\}_{j=1}^{\infty}$ a *POD basis* of $\mathcal{H}$ whenever it is an orthonormal basis of $\mathcal{H}$ for which the *total approximation error* in some norm over the complete ensemble is

$$\sum_{k=1}^{K} \|\mathbf{T}^k - \mathbf{T}_n^k\| \tag{3}$$

is minimal for all truncation levels n. This is an *empirical basis* in the sense that every POD basis depends on the data ensemble. Using variational calculus, the solution to this optimization problem amounts to finding the normalized eigenfunctions $\varphi_j \in \mathcal{H}$ of a positive semi-definite operator $R : \mathcal{H} \to \mathcal{H}$ that is defined as

$$\langle \psi_1, R\psi_2 \rangle := \frac{1}{K} \sum_{k=1}^{K} \langle \psi_1, \mathbf{T}^k \rangle \cdot \langle \psi_2, \mathbf{T}^k \rangle \tag{4}$$

with $\psi_1, \psi_2 \in \mathcal{H}$. R is well defined in this manner and corresponds to a positive semi-definite matrix whenever $\mathcal{H}$ is finite dimensional. In that case, a POD basis is obtained from the normalized eigenvectors of R, see e.g. Astrid [2004].

The POD modal coefficients 'a_j' are then obtained by the projection of the ensemble on the span of dominant POD modal coefficients as given by:

$$a_j(k) = \langle \varphi_j(z), \mathbf{T}_n(k, z) \rangle \tag{5}$$

Subsequently, a Galerkin projection is used to obtain the reduced order model as follows. Suppose that the system is governed by a PDE of the form

$$\frac{\partial T_n}{\partial t} = \mathcal{A}(T_n) + \mathcal{B}(u) + \mathcal{F}(T_n, u, d) \tag{6}$$

and let $\mathcal{H}_n$ denote an n dimensional subspace of $\mathcal{H}$ and let $P_n : \mathcal{H} \to \mathcal{H}_n$ and $I_n : \mathcal{H}_n \to \mathcal{H}$ denote the canonical projection and canonical injection maps or operators respectively. The injection map reconstruct the full scale model from reduced space. The reduced model is then given by

$$P_n \frac{\partial T_n}{\partial t} = P_n \mathcal{A}(T_n) + P_n \mathcal{B}(u) + P_n \mathcal{F}(T_n, u, d) \tag{7}$$

where observation $T_n(\cdot, k) = \mathbf{T}_n(k) \in \mathcal{H}_n = P_n \mathcal{H} \quad \forall k$, $\mathcal{A}$ is the spatial operator for convection and diffusion, and is of linear nature $\mathcal{B}$ defines input matrix and $\mathcal{F}$ is nonlinear source term. In the specific case of a POD basis, the finite dimensional subspace $P_n = \text{span}\{\varphi_j\}$, $j = 1, \ldots, n$ where the φ_j's denote POD basis functions. In that case eq. (6) becomes an *ordinary differential equation* in the coefficients $a_j(k)$ in the expansion of T_n as eq. (8) and eq. (9)

$$\frac{\partial \langle P_n, \mathbf{T}_n \rangle}{\partial t} = \mathcal{A}\langle P_n, \mathbf{T}_n \rangle + \mathcal{B}\langle P_n, u \rangle + P_n \mathcal{F}(T_n, u, d) \tag{8}$$

or equivalently,

$$\frac{da_n}{dt} = \mathcal{A}_n a_n + \mathcal{B}_n u + P_n \mathcal{F}(P_n^{-1} a_n, u, d) \tag{9}$$

Eq. (9) is reduced order model (ROM) and the POD modal coefficients 'a_j' are the states of the ROM. Therefore the POD MC can also be viewed as *dominant temporal patterns/dynamics* along which system evolves. The optimization problem to obtain POD basis as mentioned above in eq. (4) equivalently can also be solved for the ensemble $\mathbf{T}_n$ as a 'Singular Value Decomposition' *SVD* which then gives POD basis function (*spatial patterns*) in the form of left singular vectors and POD modal coefficients (*temporal patterns* as singular values multiplied by the right singular vectors. From the property of SVD these patterns are arranged as per their importance, i.e. the first POD basis corresponds to the direction of maximum energy. Usually a tolerance criterion based on amount of energy captured in the reduced model is used to decide the order or the reduced model, i.e. the span of POD basis as defined above in $\mathcal{H}_n$. The criterion is usually called projection energy and is given as below:

$$P_{tol} = \frac{\sum_{k=1}^{r} \lambda_k}{\sum_{k=1}^{n} \lambda_k} \tag{10}$$

where λ_k is the 'k^{th}' eigenvalue of the correlation operator as defined in eq. (4), 'r' is order of ROM and 'n' is order of finite dimensional full scale model. The first two terms of

eq. (9) on RHS are linear and the third non-linear term do not appear for the systems defined by linear PDEs. For the system governed by linear PDEs the differential equation eq. (9) can be transformed in equivalent discrete time form as:

$$a_n(k+1) = A_d a_n(k) + B_d u(k) \tag{11}$$

At this point one can observe that given the ensemble $\mathbf{T}_n$ one can obtain POD basis and corresponding MC, and from this knowledge of MC and system inputs 'u' the system parameters 'A_d' and 'B_d' can be easily estimated by ordinary least square (OLS) estimation techniques. If one now think of the possible approach to identify the system parameters when the governing equations are non-linear like the one in eq. (9), one needs then some approximation for non-linear terms. There are many possible ways to approximate the non-linearities like black-box, neural net, fuzzy logic, grey box, e.g. see Romijn et al. [2008] and many other input-output based fit of Weiner-Hammerstein type. It is also well known that Taylor series expansion of a nonlinear function can be a good approximation of a non-linear function. The use of Taylor Series is not considered in usual input-output identification methods due to the lack of state information. But as explained earlier in the case of model reduction, the states of ROM are accessible and therefore one can make use of Taylor series to approximate the non-linear terms. If one is interested in approximating the original full scale non-linear model then one need to include the Jacobian terms of the Taylor series in ROM. But if the approximation by linear system is not sufficient enough then one must consider the Hessian and other higher terms from the Taylor series. Note that the inclusion of the Hessian terms results into polynomial form of the identified ROM. Replacing the non-linear part by a polynomial system for multi-variable system is cumbersome due to the involvement of the tensor algebra (*Hessian computation*). For this reason we will briefly explain what does a Taylor series expansion for a scalar valued function means and then we will explain it for the vector valued function, and its implementation for the computation purpose. Another interesting feature of he polynomial systems is that they are promising candidates and have structure better suited for analytical analysis and for extension of the notions from linear system theory, e.g. see Ebenbauer et al. [2005]. For a scalar valued function,

$$\dot{x} = f(x), \text{ where } f : \mathcal{R} \to \mathcal{R} \ \& f(x^*) = 0 \tag{12}$$

Taylor series expansion in x as a nominal variable and $\tilde{x}$ as a deviation variable, $\tilde{x} = x - x^*$

$$\dot{\tilde{x}} = f(x^*) + f^{'}(x^*)\tilde{x} + (1/2!) f^{''}(x^*)\tilde{x}^2 + ... \tag{13}$$

where, $f^{'}(x) = \mathcal{J}(x) : \mathcal{R} \to \mathcal{R}$, system jacobian operator
$f^{''}(x) = H(x) : \mathcal{R} \to \mathcal{R}$, system Hessian operator.
For a vector valued function $f : \mathcal{R}^n \to \mathcal{R}^n$,

the first derivative is defined as a map: $f^{'} : \mathcal{R}^n \to \mathcal{L}(\mathcal{R}^n, \mathcal{R}^n)$, and when the first derivative is evaluated at $x^* \in \mathcal{R}^n$ then $f^{'}(x^*) \in \mathcal{L}(\mathcal{R}^n, \mathcal{R}^n)$, i.e. $f^{'}(x^*)$ is a linear operator, and when it acts on the 'n' dimensional vector 'x' then its image is $\in \mathcal{R}^n$, i.e. $f'(x^*)(x) \in \mathcal{R}^n$. This lets us to understand first derivative as a map, $f' : \mathcal{R}^n * \mathcal{R}^n \to \mathcal{R}^n$. As $f'(x^*)$ is constant term (fixed operator), we better write it as $[f'(x^*)](x) \in \mathcal{R}^n$.

We usually refer the above operator as system Jacobian matrix as, $[f'(x^*)] := \mathcal{J}(x^*)$.

The operator defined in the last expression can be written in terms of partial derivatives as,

$$[f'(x^*)](x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x^*) \dots \frac{\partial f_1}{\partial x_n}(x^*) \\ . \\ \frac{\partial f_n}{\partial x_1}(x^*) \dots \frac{\partial f_n}{\partial x_n}(x^*) \end{bmatrix} \begin{bmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{bmatrix} \tag{14}$$

equivalently,

$$[f^{'}(x^*)](x) = \begin{bmatrix} \sum_{k=1}^{n} \frac{\partial f_1(x^*)}{\partial x_k} x_k \\ . \\ \sum_{k=1}^{n} \frac{\partial f_n(x^*)}{\partial x_k} x_k \end{bmatrix} \tag{15}$$

The same procedure is repeated for computing the second derivative of the function,

$f'' : \mathcal{R}^n * \mathcal{R}^n * \mathcal{R}^n \to \mathcal{R}^n$, i.e.
$f'' : \mathcal{R}^n \to \mathcal{L}(\mathcal{R}^n, \mathcal{L}(\mathcal{R}^n, \mathcal{R}^n))$, i.e.
$f''(x^*) \in \mathcal{L}(\mathcal{R}^n, \mathcal{L}(\mathcal{R}^n, \mathcal{R}^n))$, i.e.
$f''(x^*)(x) \in \mathcal{L}(\mathcal{R}^n, \mathcal{R}^n)$, i.e.
$f''(x^*)(x)(x) \in \mathcal{R}^n$, i.e. $[f''(x^*)](x, x) \in \mathcal{R}^n$
$[f''(x^*)] := H(x^*)$, system Hessian operator.

It is clear from the above discussions that the Hessian operator is a tensor with argument from two domains while its codomain remains the same that of the Jacobian operator. The linearity of Hessian operator allows us to compute it like the Jacobian operator as in (15), but now with one more argument as:

$$[f^{''}(x^*)](x, x) = \begin{bmatrix} \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f_1(x^*)}{\partial x_k \partial x_j} x_k x_j \\ . \\ \sum_{k=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f_n(x^*)}{\partial x_k \partial x_j} x_k x_j \end{bmatrix} \tag{16}$$

the above expression can be written as:

$$[f^{''}(x^*)](x, x) = A_1(x \otimes x) \tag{17}$$

where, $(x \otimes x)$ is the Kroneckar product.

The complete simplification procedure mentioned above is aimed to express, $f'' : \mathcal{R}^n \to \mathcal{L}(\mathcal{R}^n, \mathcal{L}(\mathcal{R}^n, \mathcal{R}^n))$ as, $f'' : \mathcal{R}^n \to \mathcal{L}(\mathcal{R}^{n^2}, \mathcal{R}^n)$. This is possible due to the notion of the linearity of the tensor operator.

From the discussion above, a nonlinear equation of the form $\dot{x} = f(x, u)$ can be expanded in Taylor series as in (13) which can be approximated by a polynomial of the form,

$$\begin{aligned} \dot{x} =& Ax(t) + Bu(t) + A_1(x(t) \otimes x(t)) \\ &+ B_1(u(t) \otimes u(t)) + Q(x(t) \otimes u(t)) \end{aligned} \tag{18}$$

Where, A_1, B_1, Q are equivalent Hessian operators and $x \in \mathcal{R}^n$, $u \in \mathcal{R}^l$, $A \in \mathcal{R}^{n*n}$, $B \in \mathcal{R}^{n*l}$, $A_1 \in \mathcal{R}^{(n*n)*n}$, $B_1 \in \mathcal{R}^{(l*l)*n}$, $Q \in \mathcal{R}^{(l*n)*n}$ and $\otimes$ is the Kronecker products.

These methodological developments are based on CFD software as plant model, so for the moment we are not considering the output equations here.

Equivalent discrete form of Eq. (18) can be written as:

$$x(k+1) = A_d x(k) + B_d u(k) + A_{1d}(x(k) \otimes x(k)) + B_{1d}(u(k) \otimes u(t)) + Q_d(x(k) \otimes u(k)) \quad (19)$$

As we are considering the discrete identification problem here, for the convenience in remaining part of the paper we have dropped the superscript 'd' from eq. (19).

Please note that the polynomial equation (19) is non-linear in states and inputs but it is linear in all the system parameters (equivalent Jacobian and Hessian terms). This is a big advantage. Because if the states and inputs are known then by fixing the above polynomial model structure we can estimate the system parameters by Least Square parameter Estimation (LSE) techniques.

Coming back to the problem of the reduced model identification, the states in the (18) can be seen as POD modal coefficients (MC) and then linear and non-linear part in (9) can be written as (18).

Another interesting feature of the proposed framework is that for a large scale parameter varying systems, given the knowledge of the variation of the time varying parameter, similar approach as proposed above can be used. But the uncertain parameter should then be treated like process inputs and therefore the corresponding process snapshots due to the parameter excitation need to be included while computing the POD basis functions and MCs. As per the knowledge of the author, this approach of model reduction for very large scale process under process parameter uncertainty is never studied in past.

Once the MC and POD basis are obtained from the full scale CFD model as mentioned earlier, then by using the tensors decomposition as in eq.(16) for eq.(18), the problem of polynomial model parameter estimation is an ordinary least square estimation(*OLS*) problem and if we define,

$$\xi_k := col\,(x(k), u(k), (x(k) \otimes x(k)), (x(k) \otimes u(k)), (u(k) \otimes u(k))) \quad (20)$$

then from (19), $x_{k+1} \simeq \Theta\,\xi_k$ Where, $\Theta = [A\,B\,A_1\,B_1\,Q]$ and define the parameter estimation error at each time instance as

$$e_{k+1} = x_{k+1} - \Theta\,\xi_k \quad (21)$$

similarly the estimation error that is minimized by LSE method over the complete simulation horizon 'N' is

$$E := [x_1 \dots x_N] - \Theta[\xi_0 \dots \xi_{N-1}] \quad (22)$$

equivalently, $E := X - \Theta\Xi$
where, N is the number of samples and ,
$X \in \mathcal{R}^{n*(N-1)}$, $\Xi \in \mathcal{R}^{(n+l+n*n+l*l+n*l)*(N-1)}$ and $\Theta \in \mathcal{R}^{n*(n+l+n*n+l*l+n*l)}$

The least square solution will be

$$\Theta = X\,\Xi^T\,(\Xi\,\Xi^T)^{-1} \quad (23)$$

Please make a note here that the system parameter vector Θ is rank deficient due to the involved Kronecker product. Nevertheless, there are some simple ways to estimate the parameters for rank deficient problem as well. We simply used *Matlab* routines for our case.

The complete CFD spatio-temporal information can be reconstructed by projecting back the solution of reduced model (19) on the span of dominant POD basis P_n. The reconstructed CFD state space will be:

$$\tilde{\mathcal{I}}_n(k) = I_n\,a_n(k) = P_n^{-1}\,a_n(k), \text{or equivalently} \quad (24)$$

$$\tilde{\mathcal{I}}_n(k) = \sum_{j=1}^{r} \phi_j^{-1}\,a_j(k) \quad (25)$$

As this study is based on software simulations, the outputs can be chosen as per the user choice. In our study we have decided them close to the real life situation. The constructed output equations can be approximated as:

$$\tilde{y}(k) = C\,\tilde{\mathcal{I}}_n(k) \quad (26)$$

Note that the original Navier-Stokes equation (*non-linear PDEs*) modeled in CFD software are continuous in time and in the approach presented above we have proposed to approximate them by using discrete time linear or polynomial type non-linear equations.
The error involved here will be the sum of projection error and the statistical fit in the identification step to the few selected POD modal coefficients corresponding to the maximum energy content as per eq. (10).

One of the serious drawbacks of this approach is that the OLS estimation method for as described earlier can easily lead to an unstable system, although the original system could be a stable one. Notion of stability is discussed here as divergence of simulation results. We think one of the possible explanation could be the small data set, another could be that the POD MC obtained from SVD are right singular vectors and they are orthonormal vectors. These Orthonormal basis functions (MC) are considered as signals while they are being fitted by using a polynomial model. The orthogonality of vectors is equivalent to the property of uncorrelatedness of signals. Or, orthogonality of MC in terms of the inner product is

$$< a_i,\, a_j > \begin{cases} = 1,\ i = j \\ = 0,\ i \neq j \end{cases} \quad (27)$$

To overcome this drawback of spurious instability one might like to try some other parameter estimation method or to impose the stability in the proposed polynomial model by using some regularization trick. But usually regularization leads to bad performance of the identified model. Moreover regularization if not carried out smartly can lead to completely different dynamics of the identified model. Typically in subspace state space linear model identification techniques, regularization is imposed in the form of forcing the eigenvalues of the identified model to lie in the unit circle, e.g. see Gestel et al. [2000].
In this paper we have not solved the stability issue as the research in polynomial systems is still relatively new and imposing the stability in identification procedure will need considerable amount of further efforts.

3. MOTIVATION: GLASS MANUFACTURING

This section describes the motivating example of Industrial Glass Manufacturing Process, *IGMP*. IGMP is usually carried out in large furnaces which are very well designed in order to have a desired laminar flow pattern of the glass. A 2D view of a typical furnace is shown in figure 1. The flow pattern of glass determines the residence time of the glass in the melting furnace which in turn determines the quality of the glass produced. The process is an example of very large scale integrated systems. Most of the process variables like temperature, velocity, pressure, viscosity are

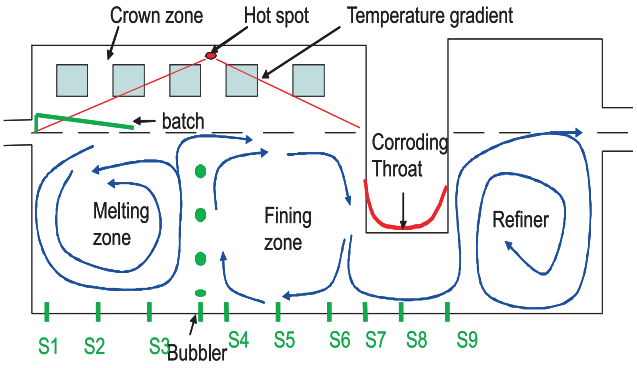


Fig. 1. Glass Manufacturing Furnace

interacting with each other. Due to this interacting nature the control of the furnace has to be done carefully. Usually pull rate (production rate), heat input and pressure valve positions are some of the control variables. Whereas variables of interest are temperature distribution and velocity profiles in the furnace. The product quality is determined by these two factors. The temperature maintained inside the furnace varies between $1400 - 1650\ ^0$C. The glass raw material enters from the left side (inlet) in the form of a batch blanket, it is melted by applying heat from the top. After circulating through the glass furnace for many hours glass passes through the throat and then leaves via the outlet. Based on the process operation there are roughly three regimes - glass melting, fining to remove high concentration of dissolved gases from the molten glass and refining to remove all remaining undissolved gases from the glass. The IGMP shows large variation in the time constants, from minutes to days. The transport of physical quantities in IGMP can be approximated with reasonable accuracy by modeling it by a set of nonlinear Navier-Stokes equations, see Huisman [2005]. There are many different types of glass furnaces and many different ways to manufacture glass depending on the type of glass required. Most of the glass manufacturing process dynamics are series combination of Continuous Stirrer Tank Reactor (CSTR) and Plug Flow Reactor (PFR). Some more details about mathematical modeling of glass can be found in Huisman [2005], Patankar [1980], Post [1988].

Due to very high process temperature and due to the viscous nature of glass, the glass furnace is a hostile environment for sensor systems. Sensors are largely limited to temperature measurements in the bottom refractory of the melting furnace. As 3 dimensional glass furnace model easily consist of $10^4 - 10^6$ finite elements, simulating its steady and/or dynamic behavior takes days for a normal configured PC and therefore it becomes very difficult to generate and process sufficient data that can be used to develop a model reduction method. For this reason we are using an approximate $2D$ glass furnace which mimics the vertical cross section along the length of $3D$ glass furnace and has only 2 grids cells in width direction.

Currently, apart from modeling the process non-linearity in the reduced order model, we are also trying to model the very slow geometric changes that take place in real $3D$ furnace in the form of throat or dam wall corrosion, see figure 1. This corrosion results into back-flow of molten glass from the refining zone to the fining zone. Such back-flow behavior causes undesired changes in the temperature distribution in the furnace which ultimately leads to economic losses. In this paper we are not addressing the corrosion problem but interested readers can refer Wattamwar et al. [2008] and a Linear Parameter Varying (LPV) system approximation in Wattamwar et al. [2009].

4. RESULTS AND DISCUSSION

In this paper a $2D$ benchmark CFD model of the original process is considered. The full scale CFD model has 3000 cells. It has many variables like temperature, velocity, concentration, pressure, etc. in each grid cell. Although most of the variables are interconnected, for the study here we have considered only temperature as variable of the interest. Therefore the order of the full scale model is 3000. From the method explained in the section 2 we have obtained a fourth order linear and non-linear polynomial model. The choice of fourth order approximation is decided based on the stability issue of polynomial model. Approximation order larger than four leads to an unstable ROM. For the linear model as well, there is not much improvement in the parameter fit above fourth order. This means that for the linear reduced model larger than fourth order there is no way to improve its performance merely by increasing its order, and there is need for non-linear reduced order model. The four POD modal coefficients corresponding to the order approximately capture 80% of the total projection energy. Usually it is desired to capture approximately 99% of the energy of the full model. But due to the stability limitation we can not satisfy this requirement.

The input considered for the identification purpose is pull-rate(feed) in terms of tons/day, which varies 5% around the nominal value in the form of +/- steps superimposed by PRBS signal. This is done to excite the slow and fast dynamics. The simulation horizon is 120 hrs and sampling time is 16 mins, therefore we have 450 snapshots. Like most of the POD related methods, this identification process is very sensitive to the type of input excitation signal. For such complex process it is also very important to know what non-linearity the identification input signal excites. If one excites soft non-linearities for such a complex process then one can expect to get a better and stable polynomial model which would fit more POD MC.

Figure 2 shows the identification result for both linear and polynomial models as proposed in this paper. Figure 3 shows zoomed version of the faster dynamics from the figure 2. Plot shows the result for four outputs which are temperature at the bottom of the four main zones of the glass, viz. Melting, Fining, Throat and Refining section. The sensors are assumed to be placed at the bottom of the tank. This is close to the real life situation. The readers can refer to the figure 1 for sensor locations. $S1$ to $S9$ are the sensors in the figure. Plot shows that the both the models approximates the overall trend very well, but the linear model fails to capture the PRBS signal dynamics precisely compared to the polynomial model.

Figure 4 shows the performance of the two models for the validation signal, which is a step input on the raw-material feed rate. Plot shows that both models follows the trend very well, but both models do not match the time constant and the final gain exactly. This is due to the two reasons. First, this is a distributed system and the

excitation signal used for the identification was designed based on the average time constant of the whole glass tank and it was not designed based on only the four location shown in the figure. Reason for the mismatch of the final gain is that these ROM could not capture 99% of the projection energy of the full scale model. One can expect smaller offset if the approximation order of the reduced model is higher. Unfortunately, as explained earlier in 2, approximation order can not be increased more than 4^{th} for the polynomial form of ROM. Nevertheless, for the size and involved complexity in GMP, even the current results seems to be very interesting.

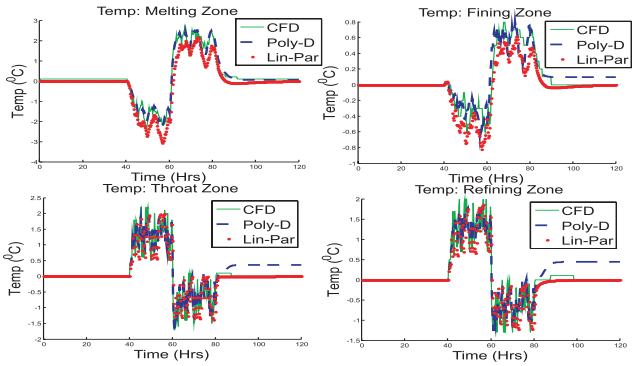


Fig. 2. Model Identification

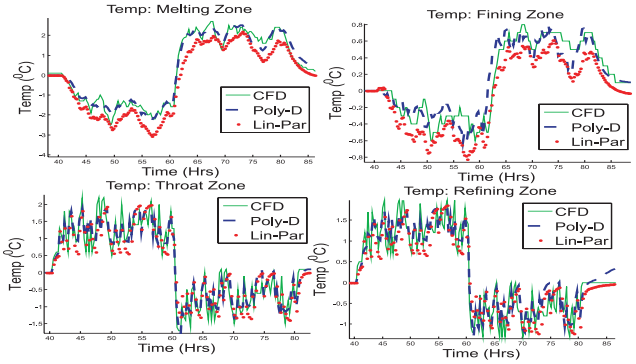


Fig. 3. Model Id: Zoom

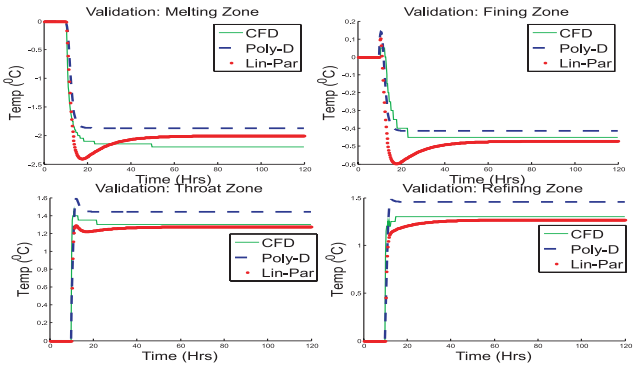


Fig. 4. Model Validation

5. CONCLUSION AND FUTURE RESEARCH

In this paper we have proposed a new model reduction method and its application on large scale industrial application. The proposed method is promising and suited especially for the very large scale processes where complexity reduction by using merely physical insight is not possible. The proposed method is also well formulated in technical aspects and with further improvements in imposing the stability in the identification of polynomial system could make this method of great potential.

We want to explore following topics in near future which has never/rarely been explored in literature like:
1. To investigate the possibility of imposing the stability in the identification process for the polynomial systems.
2. It is also possible to identify multiple linear/polynomial ROM at different working points by the method explained in this paper and construct a non-linear LPV ROM like the one described in Wattamwar et al. [2009].
3. Observer and controller design for polynomial ROM.

REFERENCES

P. Astrid. *Reduction of Process Simulation Models- A proper Orthogonal Decomposition Approach*. PhD report, Technical University of Eindhoven, 2004.

C. Ebenbauer, R. Renz, and F. Allgower. Polynomial feedback and observer design using nonquadratic lyapunov function. *Conference on Decision and Control*, 2005.

W. Favoreel, B. D. Moor, and P. Van Overschee. Subspace state space system identication for industrial processes. *Journal of Process Control*, 10:149–155, 2000.

T. V. Gestel, J. A. K. Suykens, P. V. Dooren, and B. De Moor. Imposing stability in subspace identification by regularization. *Proceedings of IEEE-CDC*, 2000.

P. Holmes, J. L. Lumley, and G. Berkooz. *Turbulence, cohorent structures, dynamical systems and symmetry*. Cambridge University Press: In Cambridge monograph on mechanics, 1996.

L. Huisman. *Control of Glass Melting Processes based on Reduced CFD Models*. PhD report, Technical University of Eindhoven, 2005.

P. Van Overschee and B. De Moor. *Identication for Linear Systems: Theory, Implementation, Applications*. Kluwer Academic, Dordrecht, The Netherlands, 1996.

S.V. Patankar. *Numerical Heat Transfer and Fluid Flow*. Hemisphere, 1980.

L. Post. *Modeling of Flow and Combustion in a Glass Melting Furnace*. PhD report, Technical University of Delft, 1988.

R. Romijn, L. Ozkan, S. Weiland, W. Marquardt, and Ludlage J. A grey-box modeling approach for the reduction of nonlinear systems. *Journal of Process Control*, 18(9):906–914, 2008.

S. Y. Shvartsman and I. G. Kevrekidis. Nonlinear model reduction for control of distributed parameter systems: A computer assisted study. *AIChE Journal*, 44(7):1579–1595, 1998.

S. Wattamwar, S. Weiland, and T. Backx. Identification of low dimensional models for slow geometric parameter variation in an industril glass manufacturing process. *Proceedings of IEEE MSC (CCA)*, 2008.

S. Wattamwar, S. Weiland, and T. Backx. Identification of low dimensional parameter varying models for large scale systems. *sent to the Proceedings of IFAC SYSID*, 2009.

Identification of Nonlinear State-Space Models: The Case of Unknown Model Structure

R.Bhushan Gopaluni *

* *Department of Chemical and Biological Engineering, University of British Columbia, Vancouver, BC, Canada V6T 1Z4, Email: gopaluni@chml.ubc.ca.*

Abstract: This article presents an algorithm for identification of nonlinear state-space models when the "true" model structure of a process is unknown. In order to estimate the parameters in a state-space model, one needs to know the model structure and have an estimate of states. An approximation of the model structure is obtained using radial basis functions centered around a *maximum a posteriori* estimate of the state trajectory. A particle filter approximation of smoothed states is then used in conjunction with expectation maximization algorithm for estimating the parameters. The proposed approach is illustrated through an example.

Keywords: Nonlinear Systems, Maximum Likelihood Parameter Estimation, Expectation Maximization, Particle Filters.

1. INTRODUCTION

Nonlinear models are commonly used to describe the behavior of many chemical processes. Process variables, typically, can be divided into latent variables (that are not measured) and measured variables. A combination of latent and measured variables can be elegantly used to represent the dynamic behavior of a nonlinear process in the following form,

$$\begin{aligned} x_{t+1} &= f(x_t, u_t, \theta) + w_t \\ y_t &= g(x_t, u_t, \theta) + v_t \end{aligned} \tag{1}$$

where $x_t \in \mathbf{R}^n$ is the n-dimensional state vector, $u_t \in \mathbf{R}^s$ is the s-dimensional input vector, $y_t \in \mathbf{R}^m$ is the m-dimensional output or measurement vector, and w_t, v_t are independent and identically distributed Gaussian noise sequences of appropriate dimension and variances Q and R respectively, $\theta \in \mathbf{R}^p$ is a p-dimensional parameter vector and $f(.)$, $g(.)$ are some nonlinear functions that describe the dynamics of the process. The nonlinear functions $f(.)$ and $g(.)$ are typically obtained using physical laws such as energy and mass balance expressions for the process. However, often, due to the complexity of chemical processes, it is difficult to develop accurate and reliable nonlinear functions. This article provides an algorithm for approximation and estimation of $f(.)$ and $g(.)$ using a combination of radial basis functions and Expectation Maximization (EM) algorithm Shumway and Stoffer (2000).

The complexity of the parameter estimation problem considered in this article arises due to unknown nonlinearities, and presence of unmeasured latent variables. If the latent variables are measured, then the model parameters can be estimated using a straightforward nonlinear least squares method Ljung (1999). If the process dynamic functions are linear, then any sub-space approach can be used Van Overschee and Moor (1996). On the other hand, if the process dynamic functions are nonlinear and latent variables are not measured, then approximate maximum likelihood approaches such as the one based on local linearization in Goodwin and Agüero (2005) and the one based on particle filter approximation in Gopaluni (2008) can be used.

The algorithm presented in this article extends the one in author's previous work on parameter estimation for known model structure Gopaluni (2008). The central idea is to find the parameter vector, θ, that maximizes the likelihood function of the observations, y_t. Due to the presence of latent variables, x_t, it is difficult to develop this likelihood function. On the other hand, due to Markov property of latent variables, it is rather straightforward to develop a joint likelihood function of the the latent and measured variables. Hence, expectation maximization, a maximum likelihood approach, that iteratively maximizes the likelihood of the observations by maximizing the joint likelihood function in each iteration, is used. EM algorithm is implemented by iteratively finding the expected value of the joint likelihood function in the first step and maximizing it in the second step Dempster et al. (1977).

This approach using EM algorithm for parameter estimation poses two problems. A structure of the process model (or in other words, the functions $f(.)$ and $g(.)$), and the distribution of noise sequence is needed to develop the joint likelihood function required in EM algorithm. Moreover, since the process is nonlinear, the distribution of latent variables, x_t, and measurements is not Gaussian even if Gaussian noise is assumed. As a result, the expected value of the joint likelihood function required in EM algorithm can not be analytically calculated. In this article, an approach that uses radial basis functions to approximate the process dynamics and particle filters to approximate the expected value of the joint likelihood function is presented.

The rest of this paper is divided into following sections: The problem is mathematically defined in section 2, a summary of expectation maximization algorithm is presented in section 3, the proposed algorithm is presented in section 4, and an example is presented in section 5.

2. PROBLEM DEFINITION AND NOTATION

As explained in the previous section, it is assumed that the process dynamics are unknown and therefore an approximation of the dynamics is needed to apply EM algorithm. It is well-known that any function can be approximated to an arbitrary degree of accuracy using basis functions such as radial basis functions. Hence, the model in (1) is approximated using radial basis functions as follows:

$$x_{t+1} = \sum_{i=1}^{I_x} h_i \rho_i(x_t, u_t, c_i, \Sigma_x) + Ax_t + Bu_t + w_t$$

$$y_t = \sum_{i=1}^{I_y} g_i \gamma_i(x_t, u_t, d_i, \Sigma_y) + Cx_t + Du_t + v_t$$

where $\rho_i(.,.)$ and $\gamma_i(.,.)$ are the radial basis functions centered around c_i and d_i with variance Σ_x and Σ_y respectively, A, B, C, D are appropriate matrices that are used to capture any linear dynamics in the model. w_t, and v_t are identically and independently distributed Gaussian noise sequences with zero mean and covariances Q and R respectively. h_i and g_i are constant vectors of appropriate dimensions. I_x and I_y are the number of basis functions used in the state and observation equations. Theoretically, even linear dynamics in the process can be approximated if sufficiently large number of radial basis functions are used. In order to reduce the total number of parameters, and capture linear dynamics, linear terms involving the matrices A, B, C, and D are added. In this article, radial Guassian basis functions of the following form are used:

$$\rho_i(x_t, u_t, c_i, \Sigma_x) = e^{\left[-(\bar{x}_t - c_i)^T \Sigma_x^{-1} (\bar{x}_t - c_i)\right]}$$

$$\gamma_i(x_t, u_t, d_i, \Sigma_y) = e^{\left[-(\bar{x}_t - d_i)^T \Sigma_y^{-1} (\bar{x}_t - d_i)\right]}$$

where $\bar{x}_t$ is the concatenated vector of states and inputs. The input-output data from the nonlinear model in (1) are denoted by $\{y_{1:T}, u_{1:T}\}$, where $y_{1:T}$ are the observations from time, $t = 1$, to $t = T$, and $u_{1:T}$ are corresponding inputs during that time period. The parameter vector θ includes all the constant parameters in the above model that describe the process behavior, and is defined as $\theta = (\theta_l, \theta_{nl})$, where θ_l consists of all "linear" parameters, h_i, g_i, Q and R, and θ_{nl} consists of all "nonlinear" parameters, c_i, d_i, Σ_x, Σ_y.

The expectation maximization algorithm plays a central role in the method developed in this article, and hence a summary of EM algorithm is presented below. It is an elegant optimization algorithm that constructs a complete likelihood function including the latent states and observations, and maximizes the likelihood function of observed data through iterations. A brief description of the EM algorithm is presented in this section to facilitate the development of the proposed algorithm in later sections.

For the state-space model described in this article, let $p(y_{1:T}|\theta)$ denote the likelihood function of the observed data. The maximum likelihood estimate of the parameter vector is obtained by maximizing this observed data likelihood function. For certain classes of state-space models, such as linear systems, it is possible to derive an explicit expression for this joint density. However, for the model considered in this paper, it is difficult to develop such an expression due to the presence of latent states. Instead, using the Markov property of the states it is straightforward to develop an expression for the complete (including states and observations) likelihood function, $p(x_{1:T}, y_{1:T}|\theta)$. In light of this feature of the Markovian states, the joint probability density function of the states and observations is iteratively maximized to obtain a maximizing θ for $p(y_{1:T}|\theta)$.

This maximization approach is called EM algorithm and can be summarized in four steps:

(1) Choose an initial guess of the parameter vector, say θ_0.
(2) Estimate the states given the parameter vector and the observations and evaluate

$$Q(\theta_i, \theta) = \int \log[p(x_{1:T}, y_{1:T}|y_{1:T}, \theta)] p(x_{1:T}|y_{1:T}, \theta_i) dx_{1:T} \quad (2)$$

where $p(x_{1:T}|y_{1:T})$ is the joint conditional density function of the states given the observations, and θ_i is an estimate of the parameter vector from a previous iteration.
(3) Maximize $Q(\theta_i, \theta)$ with respect to θ. Call the maximizing value θ_{i+1}.
(4) Repeat the above two steps until the change in parameter vector is within a specified tolerance level.

The second step in the above algorithm is called E-step and the third step is called M-step. The likelihood function, $p(y_{1:T}|\theta)$, increases monotonically through these iterations. Due to the nonlinear nature of the dynamics it is not possible to analytically evaluate the Q-function in (2). In the next section, an approximation of the Q-function and an approach to maximize it are presented.

3. MAIN ALGORITHM

3.1 Approximation of Q function

A number of approximations of EM algorithm, involving different approximations of Q function have been proposed in the literature. They either involve approximating the nonlinearities Goodwin and Agüero (2005) or approximating expected value of the joint likelihood function using particle filters and other simulation based approaches Schön et al. (2006); Gopaluni (2008). Methods involving approximation of nonlinearities will fail if the nonlinearities are prominent and on the other hand, methods involving approximation of expected value are usually computationally intensive. In this article, the Q function is approximated using a combination of particle filters and smoothers. The complete details of this approximation and its extension to handle missing data are presented in the author's work in Gopaluni (2008). For continuity a summary of this approach is presented in this section.

The Q function, using Markov property of states, can be expanded to

$$\begin{aligned}
&Q(\theta_i,\theta)\\
&=\int \log[p(x_1|y_{1:T},\theta)]p(x_1|y_{1:T},\theta_i)dx_1\\
&\quad+\sum_{t=2}^{T}\int \log[p(x_t|x_{t-1},\theta)]p(x_{t-1:t}|y_{1:T},\theta_i)dx_{t-1:t}\\
&\quad+\sum_{t=1}^{T}\int \log[p(y_t|x_t,\theta)]p(x_t|y_{1:T},\theta_i)dx_t. \qquad (3)
\end{aligned}$$

From the above expression, it is easy to notice that in order to obtain an approximation of the Q function, the following density functions are needed:

(1) $p(x_1|y_{1:T},\theta_i)$,
(2) $p(x_{t-1:t}|y_{1:T},\theta_i)$, and
(3) $p(x_t|y_{1:T},\theta_i)$.

It is possible to obtain the following particle approximations of the above density functions (please see Gopaluni (2008) for details)

$$\begin{aligned}
p(x_1|y_{1:T},\theta_i)&=\sum_{i=1}^{N} w_{1|1}^{(i)}\delta(x_1-x_1^{(i)})\\
p(x_t|y_{1:T},\theta_i)&=\sum_{i=1}^{N} w_{t|T}^{(i)}\delta(x_t-x_t^{(i)})\\
p(x_{t-1},x_t|y_{1:T},\theta_i)&=\sum_{i=1}^{N} w_{t-1,t}^{(i)}\delta(x_{t-1}-x_{t-1}^{(i)})\delta(x_t-x_t^{(i)})
\end{aligned}$$

where $w_{1|1}^{(i)}$, $w_{t|T}^{(i)}$ and $w_{t-1,t}^{(i)}$ are appropriate weights calculated using Bayes rule and importance sampling Klaas et al. (2006), $\delta(.)$ is the Kronecker delta function, and $x_t^{(i)}$ are N samples of states obtained through simulations. Using the above approximations of the density functions in the Q function, one can write the following expression,

$$\begin{aligned}
Q(\theta_i,\theta)\approx&\sum_{i=1}^{N} w_{1|1}^{(i)}\log[p(x_1^{(i)}|y_{1:T},\theta)]+\sum_{t=2}^{T}\sum_{i=1}^{N} w_{t-1,t}^{(i)}\\
&\log[p(x_t^{(i)}|x_{t-1}^{(i)},y_{1:T},\theta)]+\sum_{t=1}^{T}\sum_{i=1}^{N} w_{t|T}^{(i)}\log[p(y_t|x_t^{(i)},\theta)]
\end{aligned} \qquad (4)$$

In the above approximation, since the noise sequences are assumed to be Gaussian, the density functions, $p(x_t^{(i)}|x_{t-1}^{(i)},y_{1:T},\theta)$, and $p(y_t|x_t^{(i)},\theta)$ can be written in terms of Gaussian density functions and hence,

$$\begin{aligned}
&\log[p(x_t^{(i)}|x_{t-1}^{(i)},y_{1:T},\theta)]=\\
&-\frac{1}{2}\log(2\pi)-\frac{1}{2}\log(\det(Q))-\frac{1}{2}(x_t^{(i)}-\hat{x}_{t-1}^{(i)})^T Q^{-1}\\
&(x_t^{(i)}-\hat{x}_{t-1}^{(i)})
\end{aligned}$$

$$\begin{aligned}
&\log[p(y_t|x_t^{(i)},\theta)]=\\
&-\frac{1}{2}\log(2\pi)-\frac{1}{2}\log(\det(R))-\frac{1}{2}(y_t^{(i)}-\hat{\hat{x}}_{t-1}^{(i)})^T R^{-1}\\
&(y_t^{(i)}-\hat{\hat{x}}_{t-1}^{(i)})
\end{aligned}$$

where $\hat{x}_t^{(i)}=\sum_{i=1}^{I_x} h_i\rho_i(x_t^{(i)},u_t,c_i,\Sigma_x)+Ax_t^{(i)}+Bu_t$, and $\hat{\hat{x}}_t^{(i)}=\sum_{i=1}^{I_y} g_i\gamma_i(x_t^{(i)},u_t,d_i,\Sigma_y)+Cx_t^{(i)}+Du_t$.

3.2 Maximization of Q function

The maximization of Q function is performed in two steps using separable least squares. It is easy to notice that the parameters in θ_l enter the model linearly, while those in θ_{nl} enter the model nonlinearly. Hence, a two step procedure (called separable least squares) where the linear parameters are estimated in the first step using linear least squares, and the nonlinear parameters are estimated in the second step through nonlinear least squares. The procedure is explained below.

Step 1 Starting with an initial guess for the nonlinear parameter vector, θ_{nl}, the Q function is maximized with respect to θ_l. This maximization can be achieved through linear least squares. Before providing the maximizing value of the linear parameter vector, define the following matrices,

$$\begin{aligned}
\Omega_x&=[h_1\ h_2\ \cdots h_{I_x}\ A\ B]\\
s_t&=[I_1\rho_1(x_t,u_t,c_1,\Sigma_x)\ I_1\rho_2(x_t,u_t,c_2,\Sigma_x)\cdots\\
&\quad I_1\rho_{I_x}(x_t,u_t,c_{I_x},\Sigma_x)\ x_t\ u_t]
\end{aligned}$$

where I_1 is a vector of ones of appropriate dimensions. Noting that the Q function is quadratic in Ω_x, through straightforward calculations, it can be shown that

$$\Omega_x=\left[\sum_{t=1}^{T}<x_t s_t^T>_{xx}\right]\left[\sum_{t=1}^{T}<s_t s_t^T>_{xx}\right]^{-1} \qquad (5)$$

where $<.>_{xx}$ is used to denote integration with respect to the density function $p(x_{t-1:t}|y_{1:T},\theta)$. This integration can be approximated using the particle approximation of $p(x_{t-1:t}|y_{1:T},\theta)$. The state co-variance can be shown to be

$$Q=\frac{1}{T}\sum_{t=1}^{T}\left\langle(x_{t+1}-\Omega_x s_t)x_{t+1}^T\right\rangle_{xx}.$$

Similarly, defining the matrices,

$$\begin{aligned}
\Omega_y&=\left[g_1\ g_2\ \cdots g_{I_y}\ C\ D\right]\\
r_t&=[I_1\gamma_1(x_t,u_t,d_1,\Sigma_y)\ I_1\gamma_2(x_t,u_t,d_2,\Sigma_y)\cdots\\
&\quad I_1\gamma_{I_y}(x_t,u_t,d_{I_x},\Sigma_y)\ x_t\ u_t]
\end{aligned}$$

and noticing that the Q function is quadratic in Ω_y, it can be shown that,

$$\Omega_y=\left[\sum_{t=1}^{T}<y_t r_t^T>_x\right]\left[\sum_{t=1}^{T}<r_t r_t^T>_{yx}\right]^{-1} \qquad (6)$$

where $<.>_x$ denotes integration with respect to the density function $p(x_t|y_{1:T},\theta)$ and $<.>_{yx}$ denotes integration with respect to the density function $p(y_t|x_t,\theta)$. The measurement noise co-variance can be shown to be,

$$R = \frac{1}{T}\sum_{t=1}^{T} \left\langle (y_t - \Omega_y r_t) y_t^T \right\rangle_x .$$

The parameters in the matrices Ω_x, Ω_y, Q, and R constitute the linear parameter vector, θ_l.

Step 2 : In step one, it is assumed that the centers and widths of the radial basis functions are known. However, in practice, it is difficult estimate them. In this step an approach to estimate centers and radii is presented. The idea is to obtain a *maximum a posteriori* (MAP) estimate of the state trajectory and fix centers and radii that provide the best possible predictions of MAP state estimate and the observations. The MAP estimate of the state is obtained using a modified Viterbi algorithm as described in Godsill et al. (2001). For the sake of completeness, Viterbi algorithm is described below [1] .

Viterbi Algorithm

1. *Initialization*: For $1 \le i \le N$, $\delta_1(i) = \log(p(x_1^{(i)}) + \log(p(y_1|x_1^{(i)}))$.
2. *Recursion*: For $2 \le t \le T$, and $1 \le j \le N$,
$$\delta_t(j) = \log(p(y_t|x_t^{(j)})) + \max_i \left[\delta_{t-1}(i) + \log(p(x_t^{(j)}|x_{t-1}^{(i)}))\right]$$
$$\psi_t(j) = \arg\max_i \left[\delta_{t-1}(i) + \log(p(x_t^{(j)}|x_{t-1}^{(i)}))\right]$$
3. *Termination*: $i_T = \arg\max_i \delta_T(i)$ and $x_{MAP}(T) = x_T^{(i_T)}$.
4. *Backtracking*: For $t = T-1, T-2, \cdots, 1$, $i_t = \psi_{t+1}(i_{t+1})$, and $x_{MAP}(t) = x_t(i_t)$.

An estimate of θ_{nl} is now obtained from the data $\{x_{MAP}(1:T), y_{1:T}, u_{1:T}\}$ by fixing θ_l to its estimated value from step 1 and using nonlinear least-squares. Step 1 and Step 2 are iterated until changes in θ_l and θ_{nl} between iterations are less than a specified tolerance level.

3.3 Proposed Algorithm

The complete proposed identification algorithm is summarized below:

0. **Initialization**: Initialize the parameter vector to θ_0.
1. **Expectation**: Approximate the expected value of the complete log-likelihood function (E-step) using particle filters.
2 **Maximum a Posteriori Estimate**: Obtain a maximum *a posteriori* estimate of the state trajectory using Viterbi algorithm. Using this MAP estimate of the state trajectory, fix the centers and variances of the radial basis functions. In other words, estimate $(\theta_{nl})_{i+1}$, where i denotes the number of EM algorithm iterations performed so far.
3. **Maximization**: Maximize the Q function with respect to θ_l and call the maximizing parameter, $(\theta_l)_{i+1}$. Then set $\theta_{i+1} = [(\theta_l)_{i+1} \ \ (\theta_{nl})_{i+1}]$.
4. **Iterate**: Repeat steps 1, 2, and 3 until the change in parameter vector is within a specified tolerance level.

[1] for notational clarity, the parameter dependence is not shown in the density functions below

4. ILLUSTRATIVE EXAMPLES

The proposed approach is tried on data collected from a real continuous stirred tank reactor. The governing

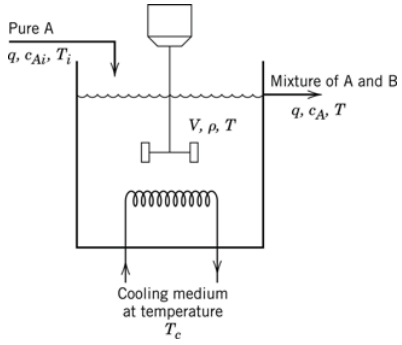


Fig. 1. Continuous Stirred Tank Reactor - Picture taken from Seborg et al. (2004).

equations of this popular CSTR, shown in figure 1, are given below (Morningred et al. (1992); Chen (2004))

$$\frac{dC_A}{dt} = \frac{q}{V}(C_{Ai} - C_A) - k_0 C_A e^{-E_A/T}$$
$$\frac{dT}{dt} = \frac{q}{V}(T_i - T) - \frac{\Delta H}{\rho C_p} k_0 C_A e^{-E_A/T} - \frac{\rho_c C_{pc}}{\rho C_p V} q_c (1 - e^{-\frac{hA}{q_c \rho_c C_{pc}}})(T - T_c)$$

where C_A is the concentration of the reactant in the reactor, T is the temperature in the reactor, q is the flow rate, V is the volume of the reactor, C_{Ai} and T_i are inflow concentration and temperature, $k_0 C_A e^{-E_A/T}$ is the reaction rate, ΔH is the reaction heat, ρ and ρ_c are the densities of the reactant and the cooling fluid respectively, C_p and C_{pc} are the corresponding specific heats, h and A are the effective heat transfer coefficient and area respectively, T_c and q_c are the temperature and flow rate of the cooling fluid. Finite difference discretization of the above continuous time differential equations results in the following model,

$$f(x_t, u_t, \theta) = x_{t-1} + \Delta t \begin{bmatrix} \frac{q}{V}(C_{Ai} - x_{t-1}(1)) - \theta_1 x_{t-1}(1) e^{-E_A/x_{t-1}(2)} \\ \frac{q}{V}(T_i - x_{t-1}(2)) - \theta_2 x_{t-1}(1) e^{-E_A/x_{t-1}(2)} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\rho_c C_{pc}}{\rho C_p V} u_{t-1} \left[1 - e^{-\theta_3 A/(u_{t-1}\rho_c C_{pc})}\right](T_c - x_{t-1}(2)) \end{bmatrix}$$

where the state vector is $x_t = [x_t(1) \quad x_t(2)] = [C_A(t) \quad T(t)]$, $\theta_1 = k_0$, $\theta_2 = (k_0 \Delta H)/(\rho C_p)$, $\theta_3 = hA$, $u_t = q_c$, $g(x_t, u_t, \theta) = x_t$ and Δt is the discretization sample time. C_{Ai} and q_c are input variables. Real data [2] collected from this reactor is shown in figures 2 and 3. Our goal is to fit a state-space model to this data assuming that the energy and mass balance expressions provided above are unknown.

[2] Please note that this data is available at http://homes.esat.kuleuven.be/ smc/daisy/

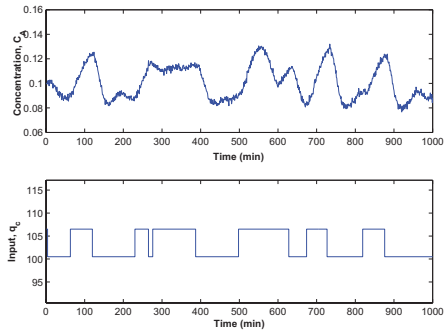


Fig. 2. The concentration, C_A, measurements.

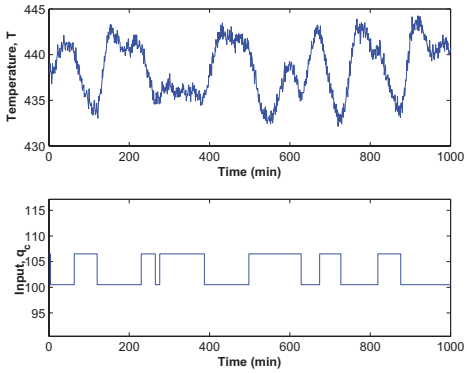


Fig. 3. The temperature, T, measurements.

The proposed algorithm is applied on this data, with a single radial basis function to describe the nonlinearities in the state and observation equations *i.e.,* with $I_x = 1$, $I_y = 1$. The accuracy of the model can definitely be increased by increasing the number of radial basis functions used. The predictions of concentration from this model for different prediction horizons are shown in figure 4. The %-fit, at these prediction horizons, calculated with the estimated model is comparable to that of input-output Hammerstein-Weiner (HW) models built using Matlab system identification toolbox. However, it should be noted that while there is no realistic and fair way to compare the complexities of HW and state space models, an attempt is made to compare the "best" trial and error based HW model with the state-space model estimated using the proposed approach.

The main advantage of the proposed method, over other nonlinear input-output identification methods, is in its ability to handle missing data - both in states and observations. In this article, missing observations are not considered. However, as shown in Gopaluni (2008), it is possible to derive particle approximations of density functions, required to approximate the Q function, even if there are some missing observations. Hence, extension of this approach to handle missing observations is rather straightforward.

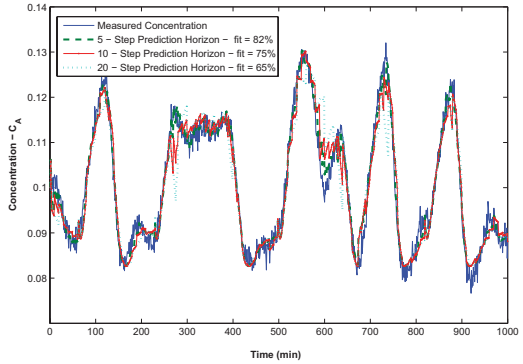


Fig. 4. True and predicted concentration profiles.

5. CONCLUSIONS

An approach to identify stochastic nonlinear systems using a combination of expectation maximization algorithm and particle filters is presented. In the proposed approach it is assumed that the model structure is unknown, and is approximated using radial basis functions. The expectation step in the algorithm is approximated using particle approximations of certain density functions. The maximization step is performed by separable least squares, where linear parameters are estimated using linear least squares, and the nonlinear parameters are estimated by using nonlinear least squares on a sequence of maximum *a posteriori* states and observations. The developed algorithm is applied to a real continuous stirred tank reactor. The proposed approach is easily extendable to handle missing observations.

REFERENCES

Chen, W. (2004). *Bayesian Estimation by Sequential Monte Carlo Sampling For Nonlinear Dynamical Systems*. Ph.d. thesis, Ohio State University.

Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B*, 39, 1–38.

Godsill, S., Doucet, A., and West, M. (2001). Maximum a posteriori sequence estimation using Monte Carlo simulations. *Ann. Inst. Stat. Math*, 53(1), 82–96.

Goodwin, G. and Agüero, J. (2005). Approximate EM algorithms for paramter and state estimation in nonlinear stochastic models. *Proceedings of IEEE Conference on Decision and Control*, 368–373.

Gopaluni, R.B. (2008). Identification of nonlinear processes with known structure under missing observations. In *Proceedings of IFAC World Congress*. Seoul, Korea.

Klaas, M., Briers, M., Nando, D., and Doucet, A. (2006). Fast particle smoothing: If I had a million particles. *International Conference on Machine Learning*.

Ljung, L. (1999). *System Identification: Theory for the user*. Prentice Hall.

Morningred, J., Paden, B.E., Seborg, D.E., and Mellichamp, D.A. (1992). An adaptive nonlinear predictive controller. *Chemical Engineering Science*, 47, 755–762.

Schön, T., Wills, A., and Ninness, B. (2006). Maximum likelihood nonlinear system estimation. In *Proceedings*

of IFAC Symposium on System Identification, 1003–1008.

Seborg, D., Edgar, T., and Mellichamp, D. (2004). *Process Dynamics and Control*. John Wiley, New Jersey.

Shumway, R. and Stoffer, D. (2000). *Time Series Analysis and Its Applications*. Springer.

Van Overschee, P. and Moor, B.D. (1996). *Subspace identification for linear systems*. Kluwer.

Subspace closed loop identification using the integration of MOESP and N4SID methods

Santos Miranda, Claudio Garcia

Polytechnic School of the University of São Paulo.
e-mail: santos.borjas@poli.usp.br, clgarcia@lac.usp.br

Abstract: Linear identification of time invariant systems operating in closed loop is of special interest for a large number of engineering applications. There are different techniques and methods to carry out this type of identification. For example, modifying the N4SID method, one can derive a closed loop subspace identification method. The same can be done using the MOESP method. Based on them, the MON4SID method is introduced, which estimates the extended observability matrix and the state sequence directly from a LQ decomposition, using a combination of the techniques contained in both, MOESP and N4SID. This new method uses an algorithm to identify state space model of a plant in a closed loop system, in the same way as in MOESP method. The advantage of the proposed algorithm is that it does not require any knowledge about the controller, whereas such information is essential for other methods (e.g. N4SID). The disadvantage of this algorithm is that it needs a great amount of data to obtain better estimates. A simulated process to show the performance of this algorithm is presented.

Keywords: Subspace identification; closed loop identification; state space models.

1. INTRODUCTION

Great part of the literature referring to system identification deals with how to find polynomial models as Prediction Error Method (PEM). In case of complex systems, there is a parameterization problem in the PEM model, so the state space model appears as an alternative to PEM, such as Multivariable Output-Error State sPace (MOESP) (Verhaegen, 1994), Canonical Variate Analysis (CVA) (Larimore, 1990) and Numerical algorithms for Subspace State Space System Identification (N4SID) (Van Overschee; De Moor, 1996). Statistical properties such as consistency and efficiency of these algorithms were studied by (Bauer, 2003; Bauer; Ljung, 2002; Chiuso; Picci, 2005). One of the main assumptions of theses methods is that the process and the measurement noises are independent of the plant input. This assumption is violated when the system is working in closed loop. The closed loop identification is of special interest for a large number of engineering applications (Ljung, 1999), since closed loop experiments are necessary if the open loop plant is unstable, or the feedback is an inherent mechanism of the system (Forssell; Ljung, 1999; Van den Hof, 1997). Several closed loop identification methods have been suggested in the last years and can be broadly classified into three main groups: direct, indirect and joint input output identification methods (Forssell; Ljung, 1999). The results of any of the N4SID, MOESP and CVA methods cited above are asymptotically biased when closed loop identification is applied. To solve this problem, the MOESP method (Verhaegen, 1993) proposed a closed loop subspace identification method through the identification of an overall open loop state space. Based on it, the plant and the controller models are estimated. To do so, it is necessary to know the order of the controller. In the N4SID case (Van Overschee; De Moor, 1997) it is necessary to know a limited number of impulse response samples of the controller and, via direct identification, the plant model is estimated. There are other possible solutions to the closed loop identification problem; the reader can consult (Huang *et al*., 2005; Katayama, 2005; Katayama *et al.*, 2002; Katayama *et al.*, 2005; Ljung; McKelvey, 1996; Qin *et al.*, 2005).

Combining the MOESP and N4SID methods, we obtain the MON4SID algorithm, which estimates the extended observability matrix in the same way it occurs in the MOESP method, the state sequence is computed through the oblique projection, as it is done in the N4SID method. From this sequence, the past and future states are obtained, and finally a consistent estimate of the system matrices is obtained, applying the least squares method. In this paper, it is proposed an algorithm to identify the state space model of a plant in a closed loop system, in the same way as it was proposed in the MOESP method, that first computes a global model from which is extracted the plant model. **T**his method does not need any knowledge about the controller.

1.1 Open Loop Subspace identification

Consider a time discrete linear time invariant dynamic system described by the state space models in the innovation form:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Ke_k \\ y_k &= Cx_k + Du_k + e_k \end{aligned} \tag{1}$$

where $u_k \in R^m$ and $y_k \in R^l$ denote the input and output signals, respectively and $x_k \in R^n$ is the vector of states. $e_k \in R^l$ is zero-mean Gaussian white noise and independent of past input and output data. *A*, *B*, *C*, *D* and *K* are matrices with appropriate dimensions.

1.2 Open Loop Subspace identification problem

The subspace identification problem is: given $u = [u_1,..,u_{nd}]$ and $y = [y_1,..,y_{nd}]$ a set of input-output measurements, determine the order *n* of the unknown system, the system matrices (*A*, *B*, *C*, *D*) up to within a similarity transformation and Kalman filter gain *K* (Van Overschee; De Moor, 1996).

1.3 Subspace matrix equation

Making successive iterations in equation (1), one can derive the following matrix equation:

$$Y_f = \Gamma_i X_f + H_i^d U_f + H_i^s E_f \tag{2}$$

where subscript f stands for the "future" and p for the "past". For the definition of the matrices H_i^d and H_i^s given in (2), see (Van Overschee; De Moor, 1996). The past and future input block-Hankel matrices are defined as:

$$\begin{pmatrix} U_p \\ U_f \end{pmatrix} = \left(\begin{array}{cccc} u_0 & u_1 & \dots & u_{j-1} \\ u_1 & u_2 & \dots & u_j \\ \dots & \dots & \dots & \dots \\ u_{i-1} & u_i & \dots & u_{i+j-2} \\ \hline u_i & u_{i+1} & \dots & u_{i+j-1} \\ u_{i+1} & u_{i+2} & \dots & u_{i+j} \\ \dots & \dots & \dots & \dots \\ u_{2i-1} & u_{2i} & \dots & u_{2i+j-2} \end{array}\right) \tag{3}$$

where $U_p, U_f \in R^{mixN}$. The output and noise innovation block-Hankel matrices $Y_p, Y_f \in R^{lixN}$ and $E_p, E_f \in R^{mixN}$, respectively, are defined in a similar way to (3).

The states are defined as:

$$X_p = X_0 = [x_0,..,x_{j-1}] \tag{4}$$

$$X_f = X_i = [x_i,..,x_{i+j-1}] \tag{5}$$

The extended observability matrix Γ_i is given by:

$$\Gamma_i = \begin{pmatrix} C \\ CA \\ \dots \\ CA^{i-1} \end{pmatrix} \tag{6}$$

The orthogonal projection of the row space of A_x into the row space of B_x is:

$$A_x / B_x = A_x B_x^T (B_x B_x^T)^* B_x \tag{7}$$

where $(\bullet)^*$ denotes the Moore-Penrose pseudo-inverse of the matrix $(\bullet)$.

The projection of the row space of A_x into the orthogonal complement of the row space of B_x is:

$$A_x / B_x^\perp = A_x - A_x / B_x \tag{8}$$

The oblique projection of the row space of *G* along the row space *H* into the row space of *J* is:

$$G /_H J = (G / H^\perp) \cdot (J / H^\perp)^* \cdot J \tag{9}$$

Properties of the orthogonal and oblique projections:

$$A_x / A_x^\perp = 0 \tag{10}$$

$$A_x /_{A_x} C_x = 0 \tag{11}$$

For a proof, see (Van Overschee; De Moor, 1996).

2. PROPOSED IDENTIFICATION METHOD

2.1 MON4SID identification method

In this subsection, the MON4SID method is presented. To solve the problem in section 1.2, it is used the POMOESP method to calculate the extended observability matrix Γ_i and the N4SID method is employed to calculate the matrices *A*, *B*, *C*, *D* through the least squares method. Therefore, it is necessary to eliminate the last two terms in the right side of equation (2). That is done in two steps: first, eliminating the term $H_i^d U_f$ in (2), performing an orthogonal projection of equation (2) into the row space of $U_f^\perp$, which yields:

$$Y_f / U_f^\perp = \Gamma_i X_f / U_f^\perp + H_i^d U_f / U_f^\perp + H_i^s E_f / U_f^\perp \tag{12}$$

And by the property (10), equation (12) can be simplified to:

$$Y_f / U_f^\perp = \Gamma_i X_f / U_f^\perp + H_i^s E_f / U_f^\perp \tag{13}$$

Second, to eliminate the noises in (13), an instrumental variable $Z = (U_p^T Y_p^T)^T$ is defined. Multiplication of (13) by *Z* yields:

$$Y_f / U_f^\perp Z = \Gamma_i X_f / U_f^\perp Z + H_i^s E_f / U_f^\perp Z \tag{14}$$

As it is assumed that the noise is uncorrelated with input and output past data (Verhaegen; Dewilde, 1992), which means that $E_f / U_f^\perp Z = 0$. Therefore, (14) is written as:

$$Y_f / U_f^\perp Z = \Gamma_i \hat{X}_f \tag{15}$$

In equation (15), $X_f / U_f^\perp Z = \hat{X}_f$ is the estimate of the Kalman filter state. Equation (15) indicates that the column space of Γ_i can be calculated by the SVD decomposition of $Y_f / U_f^\perp Z$. For further details, see (Verhaegen; Dewilde, 1992). Γ_i, given in (15), can be derived from a simple LQ factorization of a matrix constructed from the block-Hankel matrices U_f, U_p and Y_f, Y_p, in the form:

$$\begin{pmatrix} U_f \\ Z_p \\ Y_f \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \end{pmatrix} \tag{16}$$

and the orthogonal projection in the left side of (15) can be computed by matrix L_{32} (Verhaegen; Dewilde, 1992). The SVD of L_{32} can be given as:

$$L_{32} = (U_1\ U_2)\begin{pmatrix} S_n & 0 \\ 0 & S_2 \end{pmatrix}\begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = USV^T \tag{17}$$

The order *n* of the system is equal to the number of non-zero singular values in S. The column space of U_1 approximates that of Γ_i in a consistent way (Verhaegen; Dewilde, 1992), that is:

$$\Gamma_i = U_1 \tag{18}$$

The system (1) can be written as:

$$\begin{pmatrix} \widetilde{X}_{i+1} \\ Y_{i|i} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} \widetilde{X}_i \\ U_{i|i} \end{pmatrix} + \begin{pmatrix} r_i \\ r_2 \end{pmatrix} \tag{19}$$

In equation (19), suppose (ideally) that $\widetilde{X}_{i+1}$ and $\widetilde{X}_i$ are given, then the system matrices (*A*, *B*, *C*, *D*) could be computed through the least squares method. Therefore, the problem now is to find the state sequences.

$\Theta_i = Y_f /_{U_f} Z_p$ is the oblique projection (Van Overschee; De Moor, 1996), which is achieved by performing an oblique projection of equation (2), along the row space U_f onto the row space of Z_p, that is:

$$Y_f /_{U_f} Z_p = \Gamma_i X_f /_{U_f} Z_p + H_i^d U_f /_{U_f} Z_p + \\ + H_i^s E_f /_{U_f} Z_p \tag{20}$$

It is easy to see that the last two terms of equation (20) are zero, by the property of the oblique projection, equation (11); and by the assumption that the noise is uncorrelated with input and output past data (Van Overschee; De Moor, 1996). Thus, equation (20) can be simplified to:

$$Y_f /_{U_f} Z_p = \Gamma_i \widetilde{X}_i \tag{21}$$

where $\widetilde{X}_i = X_f /_{U_f} Z_p$. Then equation (21) is written as:

$$\Theta_i = \Gamma_i \widetilde{X}_i \tag{22}$$

The oblique projection Θ_i given in equation (22) can be computed from (16) by:

$$\Theta_i = Y_f / U_f W_p = L_{32}(L_{22})^{-1}(L_{21}\ L_{22})\begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}. \tag{23}$$

An estimate of the state sequence X is given by:

$$X = (\Gamma_i)^* L_{32}(L_{22})^{-1} Z_p. \tag{24}$$

The following matrices are defined: $\widetilde{X}_i = X(:,1:N-1)$, $\widetilde{X}_{i+1} = X(:,2:N)$. Thus, the system matrices can be estimated from equation (19). To estimate *K* see (Van Overschee; De Moor, 1996) or (Verhaegen; Dewilde, 1992).

2.2 Closed loop identification method

Figure 1 shows a typical standard closed loop system, where *P* and *C* denote respectively the plant and the controller, r_k is the exogenous input, u_k the input control, y_k the plant output, w_k the plant disturbance and v_k the plant noise.

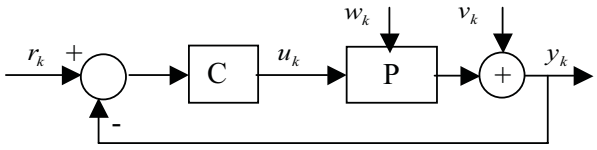


Figure1. Standard closed loop system.

P is given by equation (1) and *C* can be described by the following state equation:

$$\begin{aligned} s_{k+1} &= A_c s_k + B_c(r_k - y_k) \\ u_k &= C_c x_k + D_c(r_k - y_k) \end{aligned} \tag{25}$$

where A_c, B_c, C_c and D_c are matrices with appropriate dimensions.

2.3 Closed loop subspace identification problem

Given (r_k, u_k, y_k), a set of input output measurements finite data, of a well posed problem (Katayama, 2005), one considers the problem of identifying the deterministic part of the plant, that is, one determines the order *n* of the unknown system, the system matrices (*A*, *B*, *C*, *D*) up to within a similarity transformation.

2.3 Identification by joint input output approach

The objective of this paper is to obtain a state space model of the deterministic part of the plant *P*, based on finite measurement data (r_k, u_k, y_k). The present problem is practically the same as it was exposed in Verhaegen (1993), but the approach is quite different, as it is not necessary to know any information about the controller.

Using equations (1) and (25), it is possible to obtain a global state space model (Verhaegen, 1993):

$$\begin{aligned} \Psi_{k+1} &= \widetilde{A}\Psi_k + \widetilde{B}r_k + \Xi_k \\ \Phi_k &= \widetilde{C}\Psi_k + \widetilde{D}r_k + \Lambda_k \end{aligned} \tag{26}$$

where $\Psi_k = [x_k^T\ s_k^T]^T$, $\Phi_k = [u_k^T\ y_k^T]^T$. Ξ_k, Λ_k are noises and $\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}$ are matrices with appropriate dimensions.

The method MON4SID is applied to find an estimate of the matrices $\widetilde{A}, \widetilde{B}, \widetilde{C}, \widetilde{D}$, and based on them, to estimate Φ_k. Once Φ_k is known, it is easy to compute the matrices of the plant. To do so, the method POMOESP (Verhaegen, 1994) is used.

3. SIMULATION

In this section, we provide a simulation example to evaluate the performance of the MON4SID algorithm and to compare it with other existing identification algorithms PEM, N4SIDC and ARXS. N4SIDC here denotes the algorithm of Van Overschee and De Moor (1997) and ARXS the algorithm of Ljung and MacKelvey (1996). This example was used by (Huang *et al.*, 2004; Katayama, 2005; Overschee; De Moor, 1997 and Verhaegen, 1993). It is important to stress that the algorithm N4SIDC has three versions (Overschee; De Moor, 1997): two of them are unbiased and one is biased. The version implemented in this paper is the biased one, based on

states. In this version, it is used two different initial conditions, one for past states and the other for future states, what causes the bias. For further details see (Overschee; De Moor, 1997).

The plant is a discrete time model of a laboratory plant setup of two circular plates rotated by an electrical servo motor with flexible shafts. For further details, see (Hakvoort, 1990). The model of the plant is given by equation (1), where:

$$A = \begin{bmatrix} 4.4 & 1 & 0 & 0 & 0 \\ -8.09 & 0 & 1 & 0 & 0 \\ 7.83 & 0 & 0 & 1 & 0 \\ -4 & 0 & 0 & 0 & 1 \\ 0.86 & 0 & 0 & 0 & 0 \end{bmatrix}, \; B = 10^{-3} \begin{bmatrix} 0.98 \\ 12.99 \\ 18.59 \\ 3.3 \\ -0.02 \end{bmatrix}, \; C^T = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$K = \begin{bmatrix} 2.3 \\ -6.64 \\ 7.515 \\ -4.0146 \\ 0.86336 \end{bmatrix} \text{ and } D = 0$$

e_k is the white noise, which generates the disturbance on the plant, with standard deviation equal to 0 (for the case of deterministic system), 0.001 (to denote a system of little noise) and 0.01 (to denote a system of high noise). The controller has a state space description as in the equation (25), where:

$$A_c = \begin{bmatrix} 2.65 & -3.11 & 1.75 & -0.39 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \; B_c = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$C_C^T = \begin{bmatrix} -0.4135 \\ 0.8629 \\ -0.7625 \\ 0.2521 \end{bmatrix} \text{ and } D_c = 0.61$$

PRBS was used as an exogenous input signal, that is, persistently exciting of any finite order. There were collected 3000 samples and the number of block rows $i = 20$.

The simulation results for a closed loop deterministic identification without noise is shown in figure 2, where the order of the plant is n = 5. From figure 2, one can observe that all the algorithms had a good performance, apart from the algorithm N4SIDC, which had am improvement using n =7, as it is shown in figure 4. Figure 3 shows the poles of the original open loop plant and the estimated systems, where • denotes the original poles of the plant. One can see that all the algorithms had a good performance in relation to the estimation of the poles, which are on the unit circle.

To see an advantage of the proposed algorithm, a white noise is added to the plant, first with measurement noise variance 0.001 and then with 0.01. The comparison results are shown in figures 5 and 7 respectively.

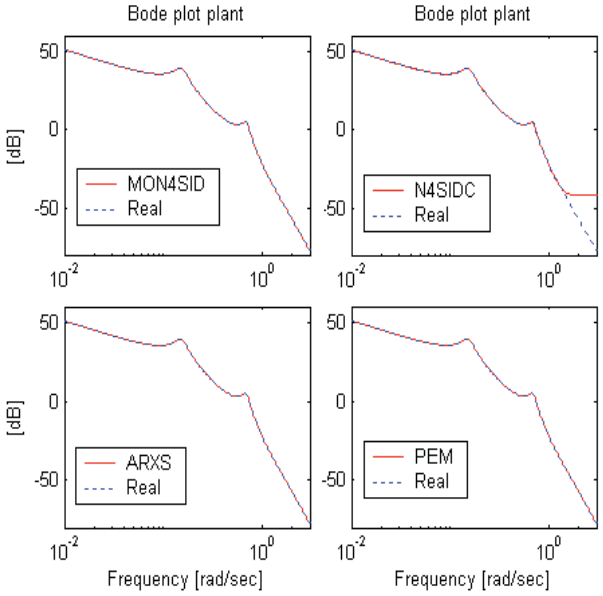


Fig.2. Bode plots of the plant P, to n=5 and no noise.

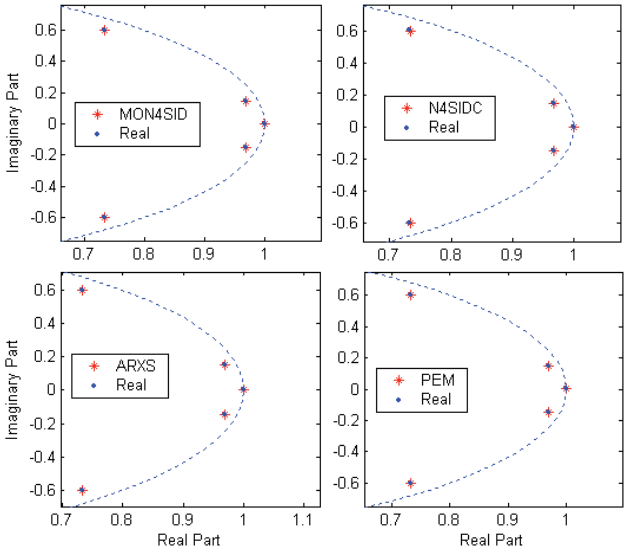


Fig. 3 Poles of the eigenvalues of the estimated A matrix.

From figures 5 and 7 one can see that the algorithm MON4SID performs better in the presence of noise. The order for identification of the plant is $n = 7$. Figures 6 and 8 show the pole estimates and the true poles of the plant. From figure 5 one can see that the algorithm ARXS does not have a good performance.

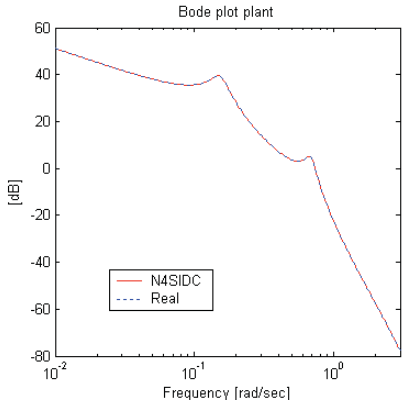


Fig.4. Bode plots of the plant P, to n=7 and no noise.

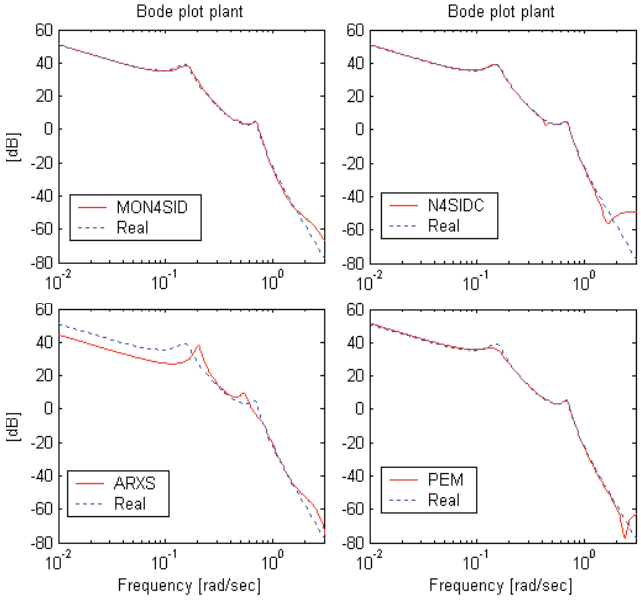


Fig.5. Bode plots of the plant P, to n=7 and little noise.

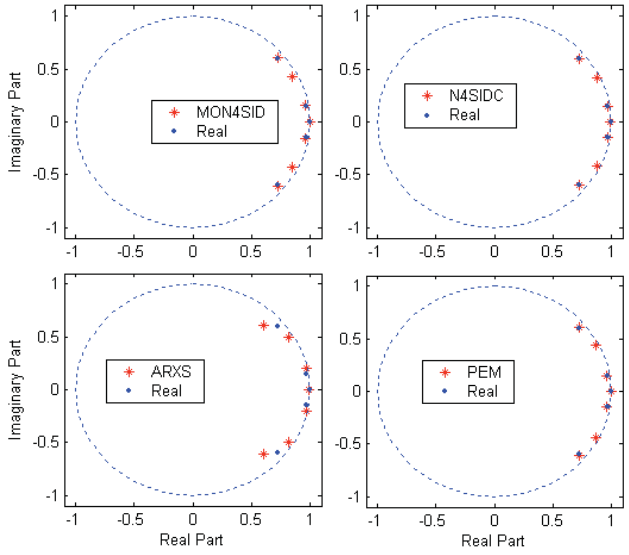


Fig.6 Poles of the eigenvalues of the estimated A matrix.

As can be noticed in figure 6 for the ARXS model, there is a difference between the estimated and real poles, what causes the difference between the real and estimated plots in figure 5.

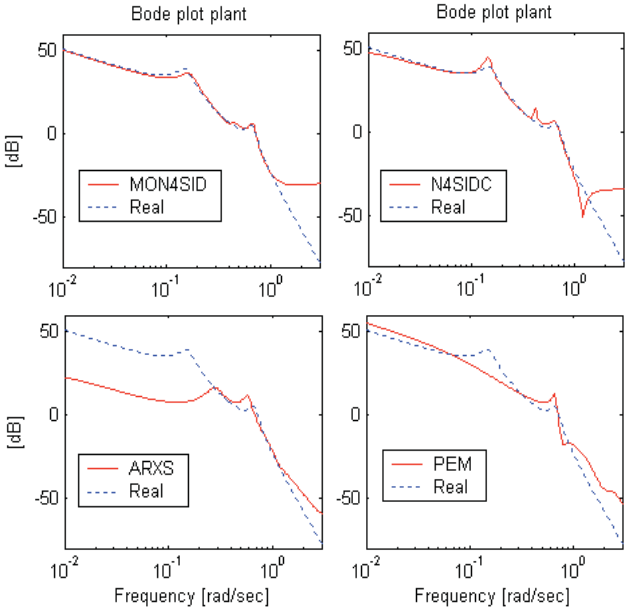


Fig.7. Bode plots of the plant P, to n=7 and high noise.

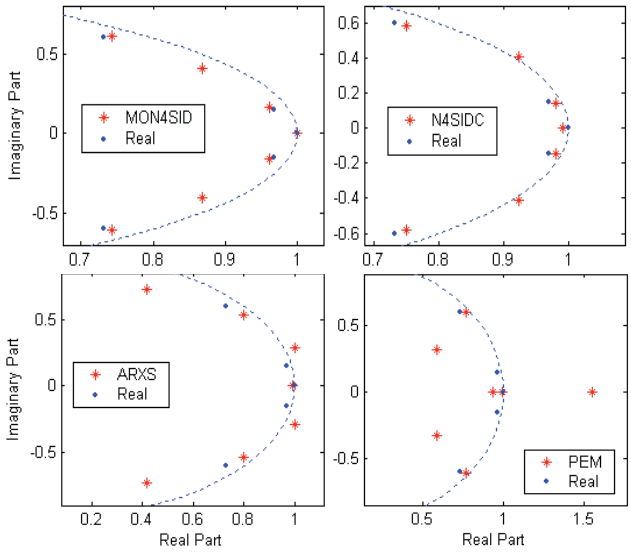


Fig.8 Poles of the eigenvalues of the estimated A matrix.

From figure 8 it can be seen that the MON4SID method provides a better estimation of the most crucial pole and all the poles are inside the unit circle. This does not happen for the other methods. Based on figure 8 one can say that direct identification models do not have a good performance for closed loop identification in the presence of high noises.

6. CONCLUSIONS

In this work, the MON4SID algorithm is presented, which uses LQ factorization in the same way as the MOESP method, which is used to compute the oblique and orthogonal

projections; these projections are used to compute the state sequence and the extended observability matrix, respectively. The past and future state sequences are computed from the state sequences, which have only one initial state. It does not happen in the N4SID method, because for each oblique projection (Θ_i and Θ_{i+1}) different state sequences ($\widetilde{X}_1$ and $\widetilde{X}_{i+1}$) are computed, generating a problem of bias in the estimates.

This algorithm was compared with three identification algorithms (PEM, N4SIDC, ARXS), when applied to a simulated example, which was used in (Van Overschee; De Moor, 1997) to identify a plant model in discrete time state space. Their results were compared by means of Bode plot and the comparison of the estimated poles with the true poles. The algorithm MON4SID presented good performance in all the cases. This algorithm has an advantage over the N4SIDC algorithm in the sense that it does not need any knowledge about the controller.

7. ACKNOWLEDGMENTS

The authors thank the financial support provided by CNPq (Brazil) for the development of this work.

REFERENCES

Bauer, D. (2003) Subspace algorithms. In *Proceedings of the 13th IFAC SYSID Symposium*, pages 1030-1041, Rotterdam, NL.

Bauer, D.; Ljung, L. (2002) Some facts about the choice of the weighting matrices in Larimore type of subspace algorithms. Automatica, volume 38, pages 763–773.

Chiuso, A.; Picci, G. (2005) Consistency analysis of some closed-loop subspace identification methods. Automatica, volume 41, issue 3, pages 377-391.

Forssell, U and Ljung, L. (1999) Closed loop identification revised, Automatica volume 35, issue 7, pages 1215-1241.

Huang, B.; Ding, S.X. and Qin, S. J. (2005) Closed loop subspace identification an orthogonal projection approach, J. Proc. Cont, volume 15, pages 53-66.

Hakvoort, R. (1990) Approximate identification in the controller design problem, Master Thesis, Delft University of Technology, Measurement and control theory section. Mech. Eng., A-538.

Katayama T.; Kawauchi H. and Picci G. (2002) Subspace identification of closed loop systems by stochastic realization, Proc. IFACWorld Congress, Barcelona.

Katayama, T. (2005) Subspace methods for system identification, Springer, London.

Katayama, T; Kawauchi, T. and Picci, G. (2005) Subspace identification of closed loop systems by the orthogonal decomposition method, Automatica, volume 41, issue 5, pages 863-872.

Larimore, W. (1990) Canonical variate analysis in identification, filtering and adaptive control, Proc. 29th Conference on Decision and Control, Hawaii, USA, pages 596-604.

Ljung, L. (1999) System Identification -Theory for the User, Prentice Hall Englewood Cliffs, NJ.

Ljung, L. and McKelvey, T. (1996) Subspace identification from closed loop data, Signal Processing, volume 52, issue 2, pages 209-215.

Qin, S. J.; Lin, W. and Ljung, L. (2005) A novel subspace identification approach with enforced causal models, Automatica, volume 41, issue 12, pages 2043-2053.

Van den Hof, P. (1997) Closed loop issues in system identification, Proc. 11th IFAC Symp. System Identification, Kitakyushu, Japan, pages 1651-1664.

Van Overschee, P. and De Moor, B. (1997) Closed loop subspace systems identification, Proc. 36th IEEE Conference on Decision and Control, San Diego, pages 1848-1853.

Van Overschee, P. and De Moor, B. (1996) Subspace identification for linear systems: Theory, implementation, applications, Dordrecht, Kluwer Academic Publishers.

Verhaegen, M. (1993) Application of a subspace model identification technique to identify LTI systems operating in closed loop, Automatica, volume 29, issue 4, pages 1027-1040.

Verhaegen, M. (1994) Identification of the deterministic part of MIMO state space models given in innovation form from input-output data, Automatica, volume 30, issue 1, pages 61-74.