DEVELOPMENTS IN MULTI-RATE PREDICTIVE CONTROL

J.A. Rossiter [*] **Tongwen Chen and Sirish L. Shah** [**]

[*] *Dept. of Automatic Control & Systems Engineering, Sheffield University, Mappin Street Sheffield, S1 3JD,*
email: J.A.Rossiter@sheffield.ac.uk
[**] *University of Alberta, Edmonton, AB, Canada, T6G 2V4.*
tchen@ee.ualberta.ca,sirish.shah@ualberta.ca

Abstract: Much of the work on predictive based multi-rate control has been based on the GPC algorithm (Clarke *et al*, 1987). However academic practitioners in single rate predictive control tend to favour approaches with better stability and performance guarantees. This paper demonstrates how those approaches might be deployed in a multi-rate framework and discusses some issues that arise.

Keywords: Multi-rate control, predictive control, constraint handling, feedforward

1. INTRODUCTION

A system is considered multi-rate (MR) when the inputs and outputs of a system are sampled at different rates. Typically this would be necessary if there were no restrictions on the speed at which the input is updated (denote this the fast rate (FR)), but output measurements are available only at a relatively slow rate (SR), for instance where a laboratory test is needed. MR systems take many forms depending on the system dimensions and the sampling rates used. Although quite common in industry, such systems have recieved relatively little study from process control (Li *et al*, 2001; Sheng, 2002) academics and hence this paper is preliminary work and we will adopt the simplest case of a dual rate (DR) system where only two sample rates are present. Moroever we assume that the output sample period is a simple multiple of the input sample period. The study of more complex cases constitutes future work.

One reason why MR systems may have recieved little attention is that one cannot easily use all the tools of linear control design. Single rate control assumes that an output measurement is available every sampling instant, then using z-transform theory one can analyse the behaviour of the nominal loop. However, such linear theory is not applicable when output measurements are available only periodically and hence at first appearances conventional design approaches cannot be used. There are two popular solutions to this difficulty: (i) inferential control (IC) (Lee *et al*, 1992) and (ii) lifting (Kranc, 1957).

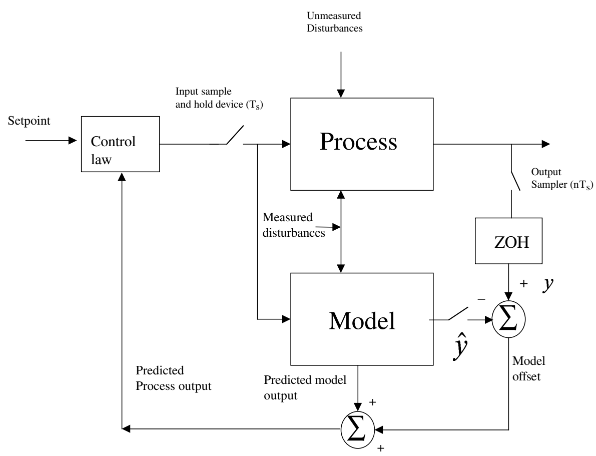

Figure 1: Internal model structure

IC makes use of an internal (see figure 1 (Garcia *et al*, 1982)) process model which operates at the fast rate (FR). This model is used to supply output estimates at the fast sample rate much like a state estimator supplies state values to be used in lieu of the actual (and unknown) state. However, this approach needs more study as there are several obvious weaknesses: (i) the state estimator/internal model receives actual output updates very slowly and this could have repercussions on accuracy and (ii) the approach relies on knowledge of a fast SR model which would have to be identified from MR data; recent work (Li *et al*, 2001) has shown that this is possible in some cases but a clear understanding of the robustness of these models constitutes work in progress.

A more popular alternative (Kranc, 1957) has been to use lifting. In essence this transforms a MR single input single output system (SISO) to a single rate multi input multi output (MIMO) system or if the system were already MIMO it increases the dimension. As the lifted system is SR (the slow rate at which the output is updated) one can use linear design and analysis methods. However: (i) there is the price of working in a significantly increased dimension and hence the design itself maybe far more complex and (ii) there is the so called causality constraint (Chen *et al*, 1994; Sheng *et al*, 2002) whereby one must ensure that the structure of the controller does not make current controls dependent on future controls. This implies a structure constraint that the feedthrough term in the controller is block lower triangular. For both IC and lifting based schemes there is also the issue of intersample ripple (Tangirala *et al*, 2001); to avoid this requires additional constraints in the controller structure.

This paper will contrast two alternative model predictive control (MPC) methods in both the lifted and IC frameworks with the aim of giving the reader a clear summary of what they gain and lose with each scenario. Section 2 will describe the necessary notation and background information. Section 3 will discuss a finite horizon algorithm (denoted FHMPC) and section 4 will develop and discuss an infinite horizon algorithm denoted (IHMPC). Section 5 will discuss the impact of constraints and section 6 presents the conclusions.

2. BACKGROUND AND OBSERVATIONS

2.1 Model predictive control

For simplicity of notation the following is restricted to single input single ouput systems, however the results are equally applicable to MIMO processes. Assume for now a single rate process. Design a finite horizon predictive control (FHMPC) law at the FR along the lines of GPC (Generalised predictive control (Clarke *et al*, 1987)) or Dynamic matrix control (DMC,(Cutler *et al*, 1980)), that is at every sample instant minimise a performance index of the form:

$$\min_{u_0,\dots,u_{n_c-1}} \quad J = \sum_{i=0}^{n_y}(r - y_{i+1})^2 + \sum_{i=0}^{n_c-1}\lambda(u_i - u_{ss})^2 \qquad s.t. \begin{cases} u_i = u_{ss}, \ i \geq n_c \\ \text{constraints} \end{cases} \tag{1}$$

where u_{ss} is the current estimate of the input required to remove steady-state offset [1] . The signals u, y, x, r are the inputs, outputs, states and set point respectively. The constraints include limits on the input, input rate and states and are assumed affine in the degrees of freedom (d.o.f.).

The weakness of FHMPC is that there are no guaranteed a priori stability results, largely because of the mismatch between the prediction assumption and the closed-loop behaviour. For computational reasons one requires the no. of d.o.f. (n_c) to be small but as a consequence the implied constraint (see (1), $u_i = u_{ss}, \ i \geq n_c$, is not close to the closed-loop evolution that is desired. This inconsistency can result in the performance being poor because the minimisation is ill posed; that is one is minimising predicted performance subject to an artifical prediction constraint that is never invoked. Hence the minimum may lie a good distance from the minimum that would arise without the artificial constraint. The effect is much less marked for larger n_c but can cause a significant degradation when n_c is small.

2.2 Infinite horizon MPC

In order to improve the properties of MPC, many authors have proposed the use of infinite costing horizons. One of the most popular of the IHMPC algorithms is given in (Scokaert *et al*, 1996). It can be summarised as at every sample minimise a cost w.r.t n_c degrees of freedom (d.o.f.),

$$\min_{u_0,\dots,u_{n_c-1}} \quad J = \sum_{i=0}^{\infty}(r - y_{i+1})^2 + \lambda(u_i - u_{ss})^2 \qquad s.t. \begin{cases} u_i - u_{ss} = -K(x_i - x_{ss}), \ i \geq n_c \\ \text{constraints} \end{cases} \tag{2}$$

K is an optimal state feedback; that is the optimal control minimising J in the absence of constraints.

The strength of IHMPC is that the open-loop predictions match the expected closed-loop behaviour, for the nominal case. Hence the optimisation is well posed and one can guarantee, a priori, stability and good performance. The main issue with this method is a possible inconsistency between the terminal constraint $u_i - u_{ss} = -K(x_i - x_{ss}), \ i \geq n_c$ and constraints, but that is not a topic of this paper.

---

[1] The control law takes a slightly different form if one uses input increments as the control variables.

2.3 Inferential control and lifting

The above algorithms were summarised for the SR case. However the context of this paper is MR systems or in particular dual rate (DR) processes where the input is updated every T seconds, but a measurement is taken every nT seconds. The algorithms need modifying to fit into this scenario. How this modification can be performed depends upon what model is available.

Inferential control (IC) requires a FR model. This assumption is a weakness but one should also state that if such a model exists, then it is to be expected that a control design using this model should outperform one based on a slow rate model.

Lifting based approaches use a DR model. There is a need to show how the IHMPC algorithm can be reformulated for this scenario and moreover to analyse its behaviour. In particular one should note (Rossiter *et al*, 2003) as discussed in section 2.1 that the restriction to DR models can give quite poorly performing control laws when one uses FHMPC. It will be shown how the move to IHMPC can overcome this weakness.

2.4 Dual rate and single rate models

Consider a FR state space model of the form

$$x_{k+1} = Ax_k + Bu_k; \quad y_k = Cx_k \tag{3}$$

The DR equivalent to this system could be written down as

$$x_{k+n} = \Gamma x_k + \Theta U_k; \quad y_k = Cx_k; \quad U_k = \begin{bmatrix} u_k \\ \vdots \\ u_{k+n-1} \end{bmatrix} \tag{4}$$

where $\Gamma = A^n$, $\Theta = \begin{bmatrix} A^{n-1}B & \cdots & AB & B \end{bmatrix}$. In many scenarios (Li *et al*, 2001) one may be able to identify Γ, Θ (or equivalent model form) from input/output data fairly easily but not A, B. Model (4) will be denoted the lifted model as the input has been lifted from u_k to U_k. Also the output/state is updated only every n samples of the FR. Effectively this gives a SR model with a lifted input.

IC assumes knowledge of the FR model whereas lifted control will make use of the lifted model and assumes the FR model is unknown.

3. FINITE HORIZON MPC

3.1 FHMPC in the lifted environment

This section will illustrate how the FHMPC control laws must be modified to cope with DR signals. First define the performance index to take the form:

$$\min_{u_0,\ldots,u_{n_c-1}} J = \sum_{i=1}^{n_y} (r - y_{k+ni})^2 + \sum_{i=0}^{n_c-1} \lambda (u_{k+i} - u_{ss})^2$$
$$s.t. \begin{cases} u_{k+i} = u_{ss}, \ i \geq n_c \\ \text{constraints} \end{cases} \tag{5}$$

Define the corresponding prediction vectors as:

$$\underset{\rightarrow}{y} = \begin{bmatrix} y_{k+n} \\ y_{k+2n} \\ \vdots \\ y_{k+n_y n} \end{bmatrix}; \ \underset{\rightarrow}{u} = \begin{bmatrix} \underset{\rightarrow}{u}_1 \\ Z \end{bmatrix}; \quad \underset{\rightarrow}{u}_1 = \begin{bmatrix} u_k \\ u_{k+1} \\ \vdots \\ u_{k+n_c-1} \end{bmatrix}$$

where Z is a vector of zeros and it is noted that the output can only be predicted every nth sample due to the limitations of the model (4). However, the input can be updated every sample. Assuming the state x is available (via an observer) the prediction model takes the form

$$\underset{\rightarrow}{y} = \underbrace{[H_1 | H_2]}_{H} \begin{bmatrix} \underset{\rightarrow}{u}_1 \\ Z \end{bmatrix} + Px_k;$$
$$H = \begin{bmatrix} \Theta & 0 & 0 & \ldots \\ \Gamma\Theta & \Theta & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}; \quad P = \begin{bmatrix} \Gamma \\ \Gamma^2 \\ \vdots \end{bmatrix} \tag{6}$$

where the partition of H is conformal with that of $\underset{\rightarrow}{u}$. One can now subsitute this prediction into (5) to derive the first n_c steps of the optimal control trajectory as:

$$\underset{\rightarrow}{u}_1 - Lu_{ss} = [H_1^T H_1 + \lambda I] H_1^T P (x - x_{ss}) \tag{7}$$

where L is an n_c vector of ones and u_{ss}, x_{ss} depend upon r and a disturbance estimate.

Remark 3.1. The main weakness of this approach (Rossiter *et al*, 2003) is the assumption that in the predictions $u_{k+i} = u_{ss}$. This assumption ensures the number of d.o.f. (n_c) is small. Where $n_c < n$ in particular the input signal has large discontinuities which are not removed by the usual receding horizon arguments as the receding horizon update takes place only every n samples in the lifted framework.

3.2 FHMPC with inferential control

In inferential control, one assumes that a fast rate model is available. Hence one can update the control optimisation at the fast sample rate, albeit the estimates of u_{ss}, x_{ss} are only updated at the slow rate. The advantage of such a change is that one no longer has to deal with the discontinuites within the input signal. What is not obvious is how to compare IC and lifting based approaches. One would expect IC control to be better simply because the receding horizon update is faster and this will be demonstrated. However this may not be a logical comparison:

- Due to modelling restrictions, lifting based MPC can only cost every nth value of the predicted output (5). No account can be taken of the unknown intersample output behaviour and this may be oscillatory (Tangirala *et al*, 2001).
- With IC one can estimate intersample outputs and hence it would be more appropriate to use the cost function of (1).

For simplicity we compare lifting and IC FH algorithms with the cost of (5). However it is noted that in practice if one were to adopt IC methods, then it would be better to use cost (1).

3.3 Example contrasting lifting and inferential control with FHMPC

Consider an example with a fast rate state space model

$$x_{k,l+1} = \begin{bmatrix} 0.3 & 0.5 \\ 0.1 & 0.9 \end{bmatrix} x_{k|l} + \begin{bmatrix} 0.1 \\ 0.2 \end{bmatrix} u_{k|l}; \quad y_k = [1 \;\; 0] x_k \tag{8}$$

For the lifted algorithm one would assume that only the equivalent model of form (4) is known. Assume that the output is sampled 4 times slower than the input, i.e. $n = 4$. The FHMPC algorithm of (7) and (5) is implemented for $n_u = 1$ with $n_y = 8, \;\; \lambda = 1$. The simulations are displayed in figures 2a,b for outputs and inputs respectively; circles and dotted lines are used for the IC algorithm and crosses and solid lines are used for the lifted algorithm. The x-axis has units of the fast sample rate so new output measurements are given only every 4th sample. The corresponding closed-loop runtime costs are given in table 1 for $n_u = 1$.

Lifted algorithm	3.18
Inferential Control	2.21

Table 1: Closed-loop runtime costs J

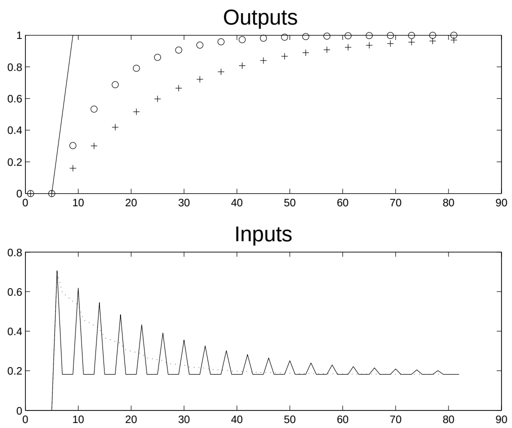

Figure 2: Simulations with $n_u = 1$

It is clear both from the table and the figure that the the use of a fast receding horizon update allowed in IC has given a dramatic improvement in performance, even though there has been not new output measurements. The limitation of the prediction assumption in FHMPC is very clear in figures 2b where it can be seen that the input moves to a *poor* default value, that is u_{ss}, during the later intersample periods. If one uses a fast receding horizon the negative effects of this poor assumption can be alleviated, as the only predicted value actually implemented is the current and the far future is continually updated. In the lifted framework, the first n moves are used and hence one is forced to use a poorly defined input trajectory.

3.4 Summary of FHMPC

FHMPC algorithms typically use input predictions which do not match the expected or desired closed-loop behaviour. This limitation is overcome by the use of the receding horizon concept whereby one updates the predictions at every sample instant so that there is a continual improvement on the initial assumption. Unfortunately in a lifted framework, the receding horizon update only takes place at a slow rate (every n samples) and as a consequence a naive use of FHMPC will cause the control law to inherit a poor input prediction. One obvious solution to this is to use IC, which was popular in some early papers on MR systems (Lee *et al*, 1992). IC allows the use of a fast receding horizon update to improve performance. However it should be emphasised that IC assumes the knowledge of a fast rate model which is not always a realistic assumption. Alternative ways around this are a topic of current research (Rossiter *et al*, 2003).

4. IHMPC IN THE MULTI-RATE ENVIRONMENT

4.1 The motivation for IHMPC

In conventional single rate MPC, there has been a move towards infinite horizons because of the attendant guarantee of stability that can be obtained. However there has been less thought given to understanding what underpins this guarantee as typically it is assumed simply to be a consequence of facilitating the definition of a Lyapunov function. However, there is a more significant change which was made use of in (Scokaert *et al*, 1996) and mentioned in section 2.2.

Ideally one wants the optimised open-loop predictions to match the actual closed-loop behaviour. Then the optimisation is well posed (unlike in FHMPC where one minimises over a class known to be different from the behaviour that will result). The consequence of this change is that the input discontinuities apparent in Fig. 2b should not occur, even in the lifted environment! To rephrase this, in the nominal case, the optimum input trajectory at time k will match exactly the optimum computed at the previous sample (in the absence of constraints). Hence whether one updates the control law at the fast rate or the slow rate, the control inputs will be the same.

We will illustrate this using the example of the previous section and the control implied by the optimisation of (2) [2]. Figure 3 below shows the simulation plots with both a lifted control law (crosses) and an IC control law (circles). Clearly the plots are identical. This implies that if one sets up the infinite horizon algorithm such that only outputs at the same sample rate are costed, then the use of lifting or IC will give the same closed-loop behaviour (in the constraint free case). However this is confusing because one would expect IC control to have more potential due to the faster receding horizon update. This apparent anomaly is discussed in section 5.

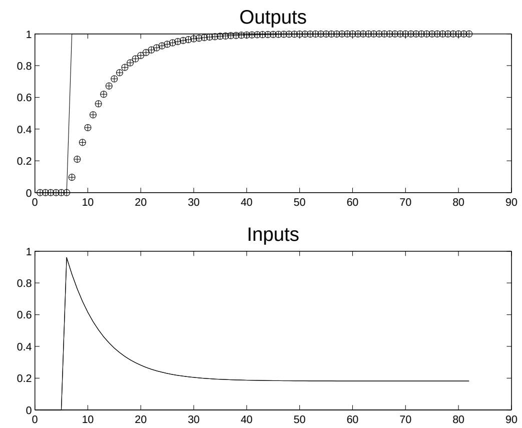

Figure 3: Simulations with optimal control

4.2 Infinite horizons need not imply that IC is equivalent to lifting

The example in the previous section made the assumption that the performance index J was the same for both the IC control and the lifting based control, that is they costed the outputs at the same sample rate, be it fast or slow. Of course in a true multi-rate framework, one does not have access to intermediate output estimates without a fast rate model. So one would use IC if a fast rate model were available and lifting otherwise. These would be based on different performance indices hence giving different control. Logically the lifted approach could not give as tight control over the unmeasured and hence uncontrolled intersample outputs (Tangirala *et al*, 2001). However there is a more noticeable difference which is discussed next.

[2] Assume that the lifted algorithm has access to intersample output estimates. One also gets the same result if both algorithms assume the cost of (5).

5. THE IMPACT OF CONSTRAINTS AND COMPUTATIONAL LOAD ON ALGORITHM SET UP

The conclusions of the previous two sections are contradictory. They imply that if one uses FHMPC then there are significant benefits from using IC. However if one uses IHMPC, then there are no benefits, that is one can obtain just as good control with an algorithm updating the control actions only at the slow rate. But, these conclusions apply to the constraint free case only, that is in the presence of constraints the global optimal input trajectory may not be known.

A popular (Rossiter *et al*, 1998) reparameterisation of the IHMPC optimisation (2) is given as

$$\min_{c_i,\ i=0,\ldots,n_c-1} J = \sum_{i=0}^{n_c-1} c_i^T c_i \quad s.t. \begin{cases} u_i - u_{ss} = -K(x_i - x_{ss}) + c_i,\ i < n_c \\ u_i - u_{ss} = -K(x_i - x_{ss}),\ i \geq n_c \\ \text{constraints} \end{cases} \tag{9}$$

Typically the global optimal requires $c_i \neq 0,\ i \geq n_c$ that is the global optimal differs from the unconstrained optimal for p steps where $p > n_c$; this is not allowed for in the prediction class so the global optimal can be reached in the optimisation. In this case it is evident that a fast receding horizon approach will give benefits as the speed of the receding horizon update governs the rate at which new d.o.f., in this case c_i, are introduced into the optimisation. Although no new observations appear at "inter-observation" instants, nevertheless the solution of the optimisation (2) does change, moving closer to the global optimal with each extra d.o.f., and hence there is a major advantage in using IC where that is possible.

5.1 Numerical example

Next a simple simulation study is used to illustrate the point that a IHMPC using IC outperforms a lifting based approach in the presence of constraints. Consider a model represented by the state equation:

$$x_{k+1} = \begin{bmatrix} 1.4 & -0.105 & -0.108 \\ 2 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_k + \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} u_k \qquad y = [0.5 \;\; 0.75 \;\; 0.5] x_k \tag{10}$$

and n is taken to be 5. For a unit set point change simulations are displayed in Figs. 4a, 4b for the constraint free case and Figs. 4c, 4d with constraints $|u| \leq 0.06$ and $|u_i - u_{i-1}| \leq 0.03$. The solid lines are with a lifted algorithm and the dotted lines represent the IC. The runtime costs J are summarised in table 2.

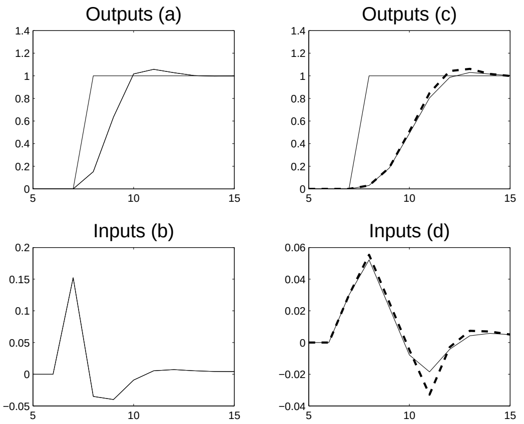

Figure 4: Simulations with optimal control

Algorithm	Figs 4a,b	Figs 4c,d
Lifted	1.11	1.946
IC	1.11	1.926

Table 2: Runtime costs

It is clear that IC has outperformed the lifted algorithm when constraints are active, although in this case by only a small amount. Larger differences will occur when the prediction class available is further from the global optimal, such as may arise with non-minimum phase and unstable systems.

5.2 Conclusion

The differences between FHMPC and IHMPC for MR systems: It was shown that the assumption, usual in FHMPC, that the predicted input move to a fixed value after n_c steps does not mesh well with MR control design. This is because the assumption is made to reduce computation not to improve control and does not match expected closed-loop behaviour well enough. Good control is recovered only by applying the receding horizon concept at a fast enough update rate. Conversely IHMPC techniques are setup to ensure a good match between predictions and expected closed-loop behaviour. Hence in this case the slow rate algorithm moves across to the MR case with a far smaller (zero for some algorithms) deterioration in performance.

The advantages of updating control with a fast receding horizon based on a FR internal model: IHMPC is identical with a FR update or lifting only in the case where the global optimal is in the class of allowable predictions. Usually restrictions to the number of d.o.f. imply this is not the case and hence one can improve performance by introducing more d.o.f. Clearly the faster the rate of receding horizon update the more quickly extra d.o.f. can be introduced to improve performance. Hence IC will always outperform lifting during constraint handling, even for IHMPC in the nominal case.

The weakness of these conclusions is the implicit assumption that one should use IC control as it gives better control for FHMPC and IHMPC. Also there is also an implication that IHMPC should always be prefered. However this is a simplistic. FR models are not always available and there is still study required to analyse their reliability. Also work in progress (Rossiter *et al*, 2003) is looking at means of obtaining control of similar quality to that obtained with IC, but based only on a lifted model. The argument of finite or infinite horizons is well known in the single rate literature of MPC and will not be repeated here.

REFERENCES

Chen, T., and L. Qiu (1994), $\mathcal{H}_\infty$ design of general multirate sampled-data control systems, Automatica, 30(7), pp. 1139-1152.

Clarke, D.W., C. Mohtadi and P.S. Tuffs (1987), Generalized predictive control, parts 1 and 2, Automatica, 23, pp. 137-160.

Cutler, C.R. and B.L. Ramaker (1980), Dynamic matrix control - a computer control algorithm, Proc. ACC, San Fransisco

Garcia, C.E. and M. Morari (1982), Internal Model control 1. A unifying review and some new results, I&EC Proc. Design and Dev., 21, pp. 308-323.

Kranc, G.M. (1957) Input output analysis of multi-rate feedback system, IRE Trans, Automatic Control, 3, pp. 21-28.

Lee, J.H., and M. Morari, Robust inferential control of multi-rate sampled data systems, Chem. Eng Sci., 1992, 47, pp. 865-885.

Li, D., S.L. Shah, and T. Chen (2001), Identification of fast-rate models from multirate data, Int. Journal of Control, 74(7), pp. 680-689.

Rossiter, J.A., M.J. Rice, and B. Kouvaritakis (1998), A numerically robust state-space approach to stable predictive control strategies, Automatica, 34, pp. 65-73.

Rossiter, J.A., T. Chen, and S.L. Shah (2003), Improving the performance of dual rate control in the absence of a fast rate model, European Control Conference.

Scokaert, P.O.M., and J.B. Rawlings (1996), Infinite horizon linear quadratic control with constraints, Proc. IFAC'96, vol. M, pp. 109-114.

Sheng, J. (2002), Generalized predictive control of multi-rate systems, Ph.D. thesis, University of Alberta.

Sheng, J., T. Chen, and S.L. Shah (2002), Generalized predictive control for non-uniformly sampled systems, Journal of Process Control, (12), pp. 875-885.

Tangirala, A.K., D. Li, R.S. Patwardhan, S.L. Shah, and T. Chen (2001), Ripple free conditions for lifted multi-rate control systems, Automatica, 37, pp. 1637-1645.

NONLINEAR PREDICTIVE CONTROL IN THE LHC ACCELERATOR

C. de Prada*, E. Blanco, S. Cristea*, J. Casas****

**Dpt. of Systems Eng. and Automatic Control, University of Valladolid, Spain*
fax: +34 983 42 3161, prada@autom.uva.es, smaranda@autom.uva.es
***LHC Division, CERN, CH-1211, Geneva (Switzerland)*
fax: +41 22 767 4620, Enrique.Blanco@cern.ch, Juan.Casas@cern.ch

Abstract: This paper describes the application of a nonlinear model-based control strategy in a real challenging process. A predictive controller based on a nonlinear model derived from physical relationships, mainly heat and mass balances, has been developed and commissioned in the Inner Triplet Heat Exchanger Unit (IT-HXTU) prototype of the LHC particle accelerator being built at CERN, operating at a temperature of about 1.9 K. The development includes a state estimator with a receding horizon estimation procedure to improve the regulator predictions.

Keywords: Nonlinear Predictive control, Nonlinear estimators, Cryogenic

1. INTRODUCTION

MPC strategies have become the preferred control technique for many process control problems, most of the industrial MPC controllers using an internal linear dynamic model. Nevertheless, many common processes exhibit nonlinear behavior and they may be required to operate over a wide range of conditions, therefore, these controllers are often tuned in a conservative way, which can result in serious degradation of controller performance. An alternative is to use a non-linear internal model. Many controllers, using different types of internal models, have been proposed in the literature, but the number of non-linear MPC is still low in industry.

In this paper we present an implementation of a non-linear MPC to a challenging process, which requires to operate under strong requirements. It corresponds to the new particle accelerator, the LHC, which is under construction at CERN, Geneva. In its final form it will expand in a circumference of about 27 Km in France and Switzerland. In order to test the proposed design some prototypes such as the String1 (Casas et al., 1998) and the IT-HXTU (Byrns et al., 1998) were built (Fig.1).

The aim of the LHC is to accelerate particles at speeds close to the one of the light in order to study the results of its collisions. For this purpose, the particles are driven within the LHC accelerator using very strong magnetic fields, which requires high electrical currents of about 12 kA for its magnets. A practical operation of the magnets requires operating with no electrical resistance in the coils, superconductivity condition, that can be maintained only at extremely low temperatures of around 1.9K. The main aim of the control system presented in this paper is to maintain this temperature in a narrow range, 50 mK, in spite of the unknown disturbances acting on the process. This is required in order to avoid a "quench" that will stop the operation.

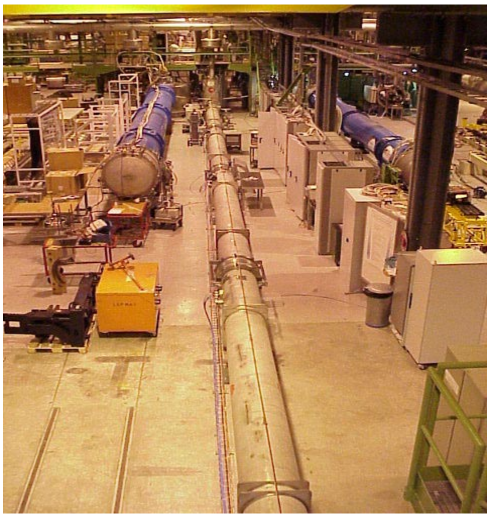

Fig. 1 Inner Triplet Prototype (length: 30 meters)

Several linear control strategies has been tested at String1, including PID and linear MPC (Blanco, 1999; Cristea, 1998) but all of them suffer from the above mentioned problem: their performance is degraded when, due to different heat load charges, the unit must operate in different working conditions. This was the main reason to implement NMPC. This paper presents the non-linear approach, as well as state and disturbance estimations that were not taken into account in previous versions, and it is organized as follows: After the introduction, section 2 describes the String1 and its cryogenic system. Section 3 is devoted to the process model and section 4 to the non-linear controller including the state estimator. Experimental results are given in section 5. The paper ends with some brief conclusions.

2. PROCESS DESCRIPTION

The LHC 1.8 K Cooling Loop represents a structure of four magnets (four quadropoles in the IT-HXTU and one quadrupole and three dipoles in the String1 prototypes) mounted at a slope of 1.4% to match the steepest inclination in the real accelerator tunnel. The superconducting magnets operate below 1.9 K in a bath of pressurized helium.

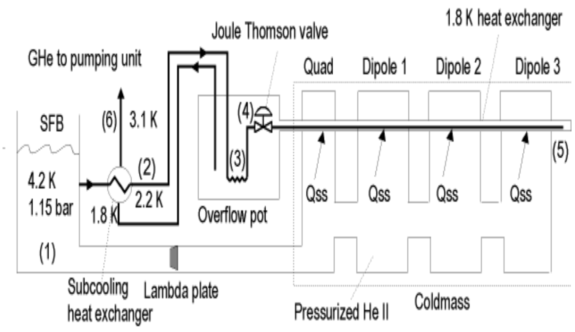

Fig. 2. Process and Schematic Diagram (String1)

Referring to Fig.2, the heat deposited on the bath is extracted by gradual vaporization of saturated superfluid helium flowing along the wetted length of a heat exchanger (HX) tube threading the string of magnets. The tube is only partially wetted, being the wetted length a main control variable. The liquid helium used for cooling at the 1.8K level is taken from the main reservoir (SFB) at 4.2 K and 1.15 bar (1). The helium is subcooled in the subcooling-heat exchanger (2) to 2.2 K, and then it is sent through the heat exchanger in the overflow pot (3). The subcooled liquid is then expanded to saturation at 17 mbar and 1.8 K in the Joule-Thomson valve (4), where a vapor fraction is created as well. The helium is led to the end of String (5). Here, it is let out in the HX, and flows back towards the overflow pot that, in normal operation, is empty. The helium vapor at 17mbar is taken out from the overflow pot (3) and through the subcooling-heat exchanger (6), thus providing the subcooling for the incoming pressurized liquid at 4.2 K.

The regulation goal is keeping the temperature of the superconducting magnets as constant as possible within strict operating constraints imposed by the maximum temperature at which the magnets can operate, the cooling capacity of the cryogenic system, the heat loads, and at last, the accuracy of the instrumentation. A small margin of a few mK is allowed before the superconductivity of the magnet coils is lost. If this happens, a potentially dangerous situation (quench) is created because of the heat released in the new conditions and the sudden helium vaporization it implies.

The Joule-Thomson valve opening is the manipulated variable, and the temperature sensors located in the cold mass (two per magnet) provide the controlled variables, the warmest temperature is taken as controlled variable at every time step. Disturbances are of two different types: general heat loads and variations in the flow through the Joule-Thomson valve. Heat loads are produced mainly by heat inleaks from the higher temperature levels, current magnet ramping and particle beam losses (simulated by electrical heaters). The set point is the saturation temperature plus a certain ΔT, typically 0.03 K.

This process has shown difficult dynamic behavior, being a non-self regulating process (integrating response), with variable dead time (transportation lag between 6-12 minutes) and exhibiting inverse response.

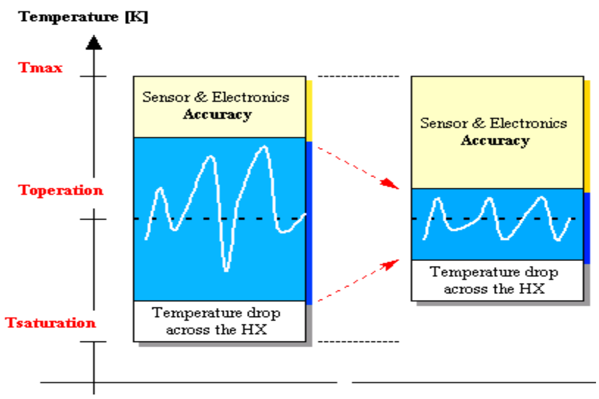

Fig. 3 - Advanced control motivation

An additional aim in implementing MPC on the plant was optimizing its operation. As can be seen in Fig. 3 reducing the variance in the magnets temperature will allow, either operating at a higher temperature setpoint without violating the upper constraints (which implies money savings because of the reduced demand on the cryogenic system) or admitting less instrumentation accuracy (which also implies savings in design and construction).

The fact that trying to squeeze as much as possible the control band is a strong constraint and a full justification for the choice of a MPC technology. The violation of this constraint would imply an eventual high-cost shutdown during normal operation.

3. PLANT MODELLING

Nonlinear predictive control (NMPC) is a natural extension of the linear MPC technique. The algorithm is again based in the use of an internal plant model, this time a nonlinear one which captures the main process characteristics. A key element in NMPC is the nature of the internal model. Several alternatives are possible including first-principle models, neural nets, Volterra series, etc. In our case a physical model was used trying to balance the capture of the process dynamics under several operating conditions and the simplicity of the representation.

A non-linear model based on physical laws and balances has been developed and validated using real experimental data obtained in the IT-HXTU installation. The implementation of this first principles model provides precise predictions over a very different operational conditions having into account changes in the saturation pressure and existing dynamic heat loads. Nevertheless, for control purposes, a simplified model is considered, based on some assumptions and equations (Blanco, 2001).

All magnets are assumed to operate at equal temperature T. Considering a single cold mass temperature simplifies the model and the heat transfer calculation through the interconnections. An energy balance leads to:

$$\frac{d}{dt}(m_{cm}C_p(T)T) = Q_{ss} + Q_{tr} - q_{cool} \qquad (1)$$

where Q are heat loads and the cooling provided by the heat exchanger, q_{cool} , is calculated through

$$q_{cool} = HA_{ws}(T - T_s) \qquad (2)$$

where the heat transfer coefficient, H, is estimated experimentally, the saturation temperature, T_s, is obtained by direct measurement of the saturation pressure, and the wetted area A_{ws} is estimated from the helium II mass accumulated in the heat exchanger tube by the following calculation

$$A_{ws} = f_1 L_{ws} \qquad (3)$$

where L_{ws} is the wetted length on the HX and it is calculated through.

$$L_{ws} = f_2 \frac{m_{HX}}{\rho(T_s)} \qquad (4)$$

f_1 and f_2 being strong non-linear functions of mass depending of the geometry of the pipe.
The accumulated helium II mass in the heat exchanger is calculated by

$$\frac{dm_{HX}}{dt} = lfjt - Fr - Fv \qquad (5)$$

where the l_{fjt} is the liquid flow passing through the Joule-Thomson valve, Fr, the helium II liquid overflow and Fv the helium which evaporates in the HX, having also into account the vapour fraction, flash, produced by the Joule-Thomson valve.

$$l_{fjt} = m_{fjt}(1 - vflash) \qquad (6)$$

where m_{fjt} represents the total mass flow passing through the JT valve which depends on the valve characteristic, and *vflash* the vapour fraction produced. This is computed from an enthalpy balance between the incoming high-pressure stream and the two coexisting phases at saturation pressure at the output of the JT valve.

$$vflash = \frac{H_{liq}(before) - H_{liq}(after)}{H_{gas}(after) - H_{liq}(after)} \qquad (7)$$

Finally, the vapour on the heat exchanger is composed by two components, vapour fraction produced by the JT valve, and evaporation of the He II liquid.

$$Fv = g_{fjt} + \frac{q_{cool}}{h_{fg}} \qquad (8)$$

g_{fjt} represents the gas flow produced by the JT valve and it is calculated by (9), being h_{fg} the latent heat of vaporization of liquid helium.

$$g_{fjt} = m_{fjt} \cdot vflash \qquad (9)$$

The JT valve is characterized by a calibration curve and its opening represents the input of the model, v_{open}, and the constants, cti, give the valve characteristic.

$$mfjt = ct1 \cdot v_{open}{}^2 + ct2 \cdot v_{open} + ct3 \qquad (10)$$

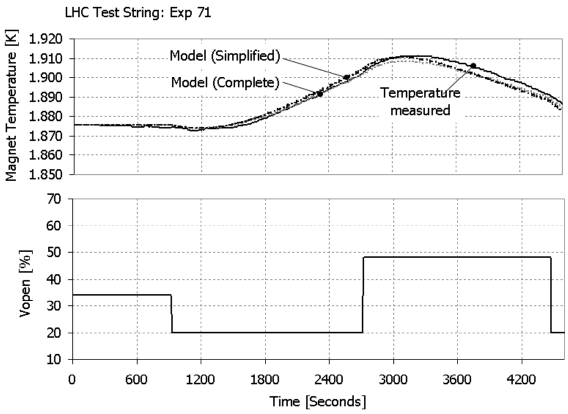

Fig. 4 Simplified vs. Complete first principles model

A comparison between the simplified model versus the complete model shows the magnet temperature when the JT valve is moved (Fig.4). A good trade off between complexity and quality of the model is obtained.

The value of some parameters, i.e. the cold mass and the heat transfer, was estimated by off-line optimization of the model errors.

4. PREDICTIVE CONTROLLER

The objective of the non-linear model predictive control (NMPC) is finding the future optimal manipulated variable sequence in order to minimize a function based on a desired output trajectory over a prediction horizon. The cost function is the integral of the sum of squares of the residuals between the model predicted outputs and the setpoint values over the prediction horizon N_2, plus a penalty term. A typical formulation is

$$\min_{u(k/k),\ldots,u(k+Nu-1/k)} \left(\int_{t_k}^{t_k+N_2} [\gamma(y(t) - r(t))]^2 dt + \sum_{j=0}^{Nu-1} \beta[\Delta u(k+j)]^2 \right) \qquad (11)$$

where y and u are the process output and input. The minimization (11) is done subject to the continuous model equations and the typical restrictions applied on the manipulated and controlled variables:

$$\begin{aligned} & u_{\min} \le u(k) \le u_{\max} \\ & \Delta u_{\min} \le \Delta u(k) \le \Delta u_{\max} \\ & y_{\min} \le y(k) \le y_{\max} \end{aligned} \qquad (12)$$

Of the *Nu* moves optimal control sequence, only the first component is implemented. The optimization is solved using the scheme of Fig.5. Within this schema, the model equations are not explicit restrictions to the optimisation problem, being the manipulated variables the only decision variables. The simulation package will integrate the model equations along the prediction horizon taking as initial conditions the current process state and evaluating the formulated objective at the end of the integration. Path constraints are implemented as penalty functions. A simultaneous solution approach was also tested, but convergence and computation time did not improve the sequential one.

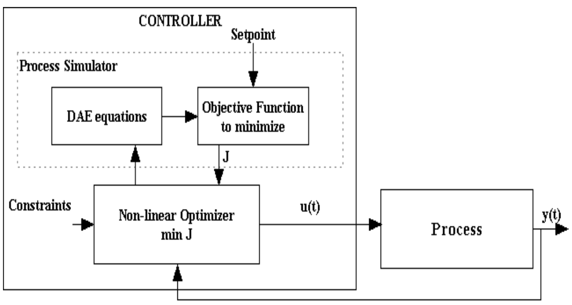

Fig. 5 Nonlinear controller – Continuous implementation framework

4.1 Nonlinear State Estimator

In our plant, the liquid helium II accumulated in the HX tube is not measurable. This is a critical factor of the model predictions, not only because it is a state but because it provides the wetted area in the heat exchanger from where the heat in the pressurized helium is removed. So, as we have an incomplete state vector, in order to apply the NMPC, a method of reconstructing the current state of the system from the measured outputs must be included. There are different approaches to the state estimation problem. We have chosen a receding horizon one (Muske and Edgar, 1997) because it match very well within the predictive control framework, it allows easy extensions to the non-linear case maintaining the same model as in the controller and with explicit inclusion of constraints in the variables and, finally, because state disturbances can be computed as a sub-product of the estimation. In our case it is really important estimating the non measurable overall *heat load* because it highly influences the model predictions. LHC prototypes operation has shown the variance on this disturbance in short periods of time.

In analogy to the model predictive control concept, the estimation problem is formulated as an optimal control problem on a finite horizon into the past. In the framework of the receding horizon estimation a quadratic cost function penalizing, among other things, model and measurement errors, is minimized. The optimisation problem is subject to model equations. Physical limits on the process variables are incorporated through inequality constraints.

More precisely, the problem is to estimate the initial conditions at time instant *t-N*, and the state disturbances, which have driven the process to its present state applying the past control sequence, by minimizing the difference between the outputs given by the evolution of the system from its initial conditions and the actual measured outputs (Fig. 6). In our case this can be translated into estimating the liquid helium mass in the HX tube at time instant t-N and the heat loads in the range [t-N, t-1], so that, if the JT valve were operating as in the real plant in this time interval, the computed temperature would approach the measured one and the heat loads are as small as possible.

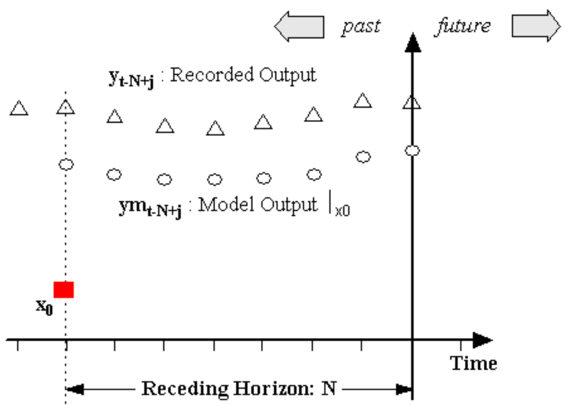

Fig. 6 Receding Horizon State Estimation

The standard approach can be synthesized as the optimization problem:

$$\min_{m_0, Q_{k-N}, \ldots \hat{w}_k} J = \gamma_1 \sum_{j=0}^{N} \left(ym_{k-N+j} - y_{k-N+j} \right)^2 + \gamma_2 \sum_{j=0}^{N} \left(Q_{k-N+j} \right)^2 \qquad (13)$$

where N represents the receding horizon, one of the parameters to tune. y_{k-N+j} is the real process output and ym_{k-N+j} is the output of the model starting with $m_0 = mHX$ at the time $(k\text{-}N)$ when the past controls and the estimated process disturbances Q_{k-N+j} are applied.

Once the initial state and disturbances are estimated, the unknown state at time t can be computed integrating the model with the optimal values obtained.

Nevertheless, when applying (13) we realize that, due to the particular estructure of our process model, they were many solutions able to provide a perfect fit between the measured and computed temperatures. So, in our case, instead of (13) an alternative criteria was formulated, where, on one hand, the model temperature was equated to the measured one in the interval, and a new cost function was defined based on: (A) minimize the initial difference between the mass and its estimated value in the previous iteration mHX_{k-N} (B) penalization of the mass changes only if the JT valve was smooth during the horizon, in other case, where the JT was active, this contribution is cancelled, and finally (C) minimize the heat load change with respect to the value estimated in the previous iteration which provides a smooth JT valve moves.

With this structure the objective becomes

$$\min_{m_0} J = \gamma_0 [A]^2 + \gamma_1 \sum_{j=0}^{N-1} [B]^2 + \gamma_2 \sum_{j=1}^{N-1} [C]^2 \qquad (14)$$

where γ_i weights the contribution of each factor and A, B, C represent

$$A = m_{t-N} - m_{HX\,k-N}{}^{rec}$$
$$B = \frac{m_{HX\,j+1}{}^{cal} - m_{HX\,j}{}^{cal}}{1 + \left[vopen_{j+1}{}^{rec} - vopen_{j}{}^{rec}\right]} \qquad (15)$$
$$C = Q_{tr_j}{}^{cal} - Q_{tr_j}{}^{rec}$$

where $m^{cal}{}_{HXk-N}$ represents the liquid helium II in the heat exchanger already calculated N periods before, $m_{HXj}{}^{cal}$ the mass calculated along the receding horizon, $vopen_j{}^{rec}$ the manipulated variable (JT valve) moves recorded during the receding horizon, $Qtr_j{}^{cal}$ the calculated dynamic heat load and $Qtr_j{}^{rec}$ the previously calculated and recorded dynamic heat load during the horizon.

Estimating only the initial state and heat load disturbances makes the problem tractable. The advantage of this solution is the smaller number of decision variables for any given horizon length, resulting in less computational time to solve the optimization problem (Gelb, 1974). In this case the state disturbance, the heat load present in the process, is included in the estimation procedure by means of the model equations. A good starting point helps in the estimation, which can be found if the process starts operating in a known steady state like helium overflowing. Besides (14), the optimization includes constraints on the values of the decision variables.

The control structure designed for the nonlinear controller incorporates a nonlinear predictive algorithm and a state estimator. The solution proposed yields a new approach based of an initial state estimate and of a moving horizon algebraic estimator in a combined structure. The state estimator provides the optimal mass accumulated in the heat exchanger tube and the dynamic heat load valuation. A block diagram of the structure can be seen in Fig. 7.

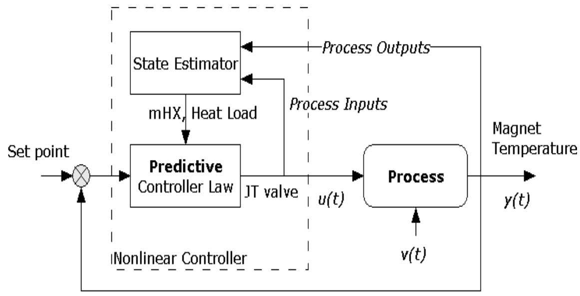

Fig. 7 NMPC proposed control structure

4.2 Simulation results

Simulation studies have been performed in order to verify the performance of the implemented estimator and the improvement on the control. An example is shown in the Fig. 8 where several steps of the heat load (12 Watts) were applied and the values of the estimated state disturbance are compared against the values given by the model showing good agreement. The only tuning parameter, apart from the weights, in the proposed objective structure, is the receding horizon value N (here, N=4). The process is represented by the full simulation model.

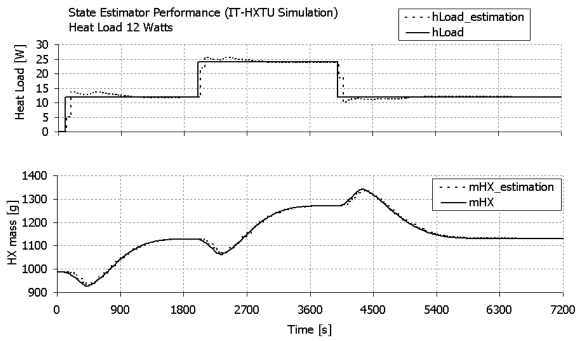

Fig. 8 Performance of the Nonlinear State Estimator

The complexity of the optimization problem is not growing proportionally to the length of the horizon due to the fact that only one initial value is considered as a decision variable and the remainders are algebraic calculations. N has not much influence on neither the controller nor the state estimator performance. In the test, only the parameter γ_0 conditions the way the heat load is estimated, the greater the number, the faster the heat load disturbance estimation.

4.3 Optimization: numerical solutions

Both, the controller law solution and the state estimation problems, presented in the nonlinear predictive controller framework, lead to the same nonlinear programming problem, which could be formulated generically as a real time minimization of a nonlinear function subject to constraints. These constraints could be simple bounds on the variables and both, linear and nonlinear constraints. In the case of the LHC 1.8 K Cooling Loop the method used is a SQP one, due to its ability to solve problems with nonlinear constraints.

5. EXPERIMENTAL RESULTS

The validation of the state estimator module based on the receding horizon was done experimentally by powering the electrical heaters located in the cold mass. These simulate a change in the overall heat load due to a unknown disturbance, then data was stored corresponding to the electrical watts applied and the heat load estimation carried out by the state estimator.

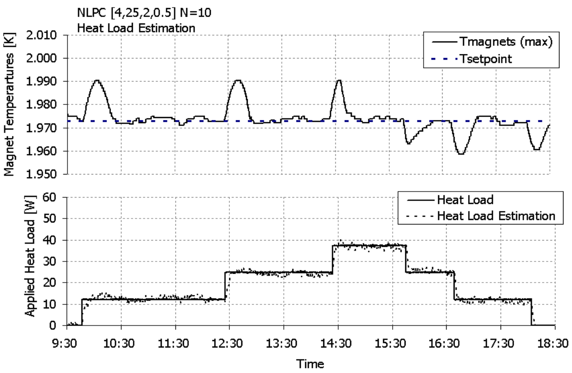

Fig. 9 . RHE performance. Heat load steps

In Fig. 9 several step changes on the heat load were applied to the process. Performance of the state estimator is fast and precise, and the heat load is estimated immediately after its change despite the abrupt jump. This situation could be produced by several factors in the real systems, for example, a degradation of the insulation vacuum which leads to a higher existing overall heat load. The other state estimated, the accumulated helium mass in the HX tube is not shown because is a non-measured variable and no comparison are possible. Performance of the controller is also displayed in the same Fig. 9. The temperature excursions, due to the

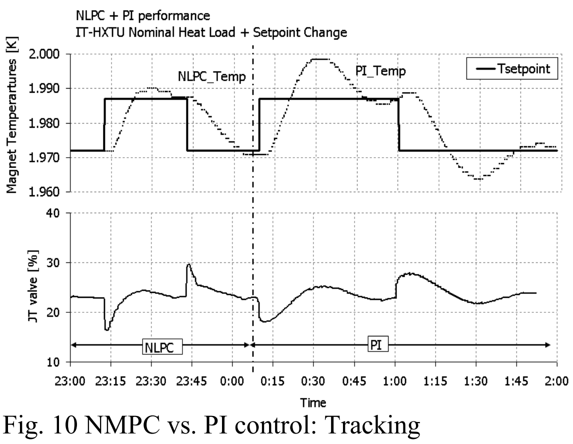

Fig. 10 NMPC vs. PI control: Tracking characteristics

heat load applied, are cancelled around 1.99 K in all the different operational zones showing a robust behavior of the regulator.

Once the state estimator was tested and validated, more experiments were performed in order to validate the nonlinear predictive controller. Changes in the set point were considered to check also for tracking features (Fig. 10). A comparison with a classical PI controller is done in order to show the improvements in terms of control performance and robustness.

6. CONCLUSIONS

The LHC full-scale prototypes were employed as a test-bed of what advanced nonlinear control can do for improving for cryogenic processes regulation. The nonlinear process model construction gave a better understanding of the process, provided the ideas to overcome the usual changes in process dynamics and helped to improve the regulation strategies by means of the simulation. The response has been improved and optimized by the use of the nonlinear predictive controller with a receding horizon state estimator. The regulation structure proposed is based in a nonlinear predictive controller algorithm combined with a state estimator with an initial state estimate approach and a moving horizon algebraic calculation for the disturbance.

ACKNOWLEDGEMENTS

The authors would like to thank the financial support of the project PPQ2000-1075-C02-01 by the Spanish CICYT and are very grateful to the LHC Cryogenics group (ACR) and the String and IT-HXTU teams.

REFERENCES

Blanco, E., C. Prada, J. Casas and S. Cristea (1999). Predictive Temperature Control of the LHC Superconducting Magnets. *European Control Conference ECC'99*, Karlsruhe, Germany.

Blanco, E. (2001). Nonlinear Model-Based Predictive Control Applied to Large Scale Cryogenic Processes. In *Ph.D. Thesis, Universidad de Valladolid*, Valladolid, Spain.

Cristea, S. (1998). Control predictivo en el dominio delta. In *Ph.D. Thesis, Universidad de Valladolid*, Valladolid, Spain.

Casas, J., et al., (1998) Operation Testing and Long Term Behavior of the LHC Test String Cryogenic System. In *ICEC17 Proceedings*, edited by Hughes. D., et al., UK, pp. 747-750.

Gelb, A. (1974). Applied Optimal estimation. MIT Press, Massachusetts

Muske, K.R. and T.F. Edgar (1997). Nonlinear State Estimation. In *Nonlinear Process Control* (M.A. Henson and D.E. Seborg, (Eds)), pp. 311-370. Prentice-Hall, Inc., New Jersey

DISTURBANCE ATTENUATION WITH ACTUATOR CONSTRAINTS BY MOVING HORIZON $\mathcal{H}_\infty$ CONTROL

H. Chen *C.W. Scherer **

* *Dept. Control Science and Engineering, Jilin University*
Renmin Str. 142, Changchun 130025, PR China
chenh@jlu.edu.cn
** *Mechanical Engineering Systems and Control Group*
Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
c.w.scherer@wbmt.tudelft.nl

Abstract: Exploiting the moving horizon strategy, we provide in this paper a solution of the constrained L_2-gain attenuation control problem that is less conservative than a recently suggested switching approach based on off-line controller computations. The main advantage of the presented scheme is its capability of automatically relaxing or tightening the performance specification in order to obey hard control constraints while achieving the best possible performance in a suitable class of LMI-generated control gains.

Keywords: L_2-gain attenuation, control constraints, dissipation theory, on-line optimization

1. INTRODUCTION

In the past two decades H_∞-control has received considerable interest, in particular for the possibility to beneficially manage the trade-off between high performance requirements and high control action. It is somewhat unfortunate, however, that the designer has only influence onto closed-loop transfer function shapes in the frequency domain and that there is no direct way to enforce hard time-domain constraints on the control inputs. To overcome this drawback, a large variety of approaches have been proposed in the literature such as anti-windup techniques (Kothare *et al.*, 1994), saturation avoidance methods based on maximal output invariant sets (Gilbert and Tan, 1991), model predictive control (Mayne *et al.*, 2000) and switching techniques (Hirata and Fujita, 2000). For a survey we refer to (Scherer *et al.*, 2002) and the references therein.

In this paper, we provide a moving horizon scheme for the L_2-gain attenuation problem with hard control constraints, where a constrained $\mathcal{H}_\infty$ problem is solved on-line and updated by the new measurement. The scheme has the capability to automatically trade-off constraint satisfaction and performance by relaxing or tightening the performance specification, which leads to performance improvements. The feedback gain is determined on-line such that the ellipsoids, where constraints are respected, are shaped according to the actual state and hence performance can be further improved, whereas the off-line controller construction in (Scherer *et al.*, 2002) is based on extremal solutions of the Riccati equation corresponding to the $\mathcal{H}_\infty$ problem. Therefore, this paper can be viewed as a direct extension of (Scherer *et al.*, 2002) towards a non-conservative solution of the constrained L_2-gain attenuation control problem. In a similar fashion, it is suggested in (Kothare *et al.*, 1996) to use the moving horizon strategy in order to ensure robust stability while minimizing an upper bound of a quadratic cost, whereas our scheme explicitly strives for L_2-gain performance guarantees for the overall closed-loop system.

The paper is organized as follows. In Section 2 we describe an off-line solution to the constrained $\mathcal{H}_\infty$ control, using the concept of state-space ellipsoids and reachable sets (Boyd *et al.*, 1994). In Section 3 we derive the crucial condition to guarantee dissipation after briefly showing why the naive implementation of the moving horizon strategy might fail. Then, an extended LMI optimization problem is formulated that will be solved on-line at each sampling time to determine the feedback gain, updated with the actual state. An algorithm for a concrete implementation of the proposed scheme is given in Section 4. Simulation results for the same open-loop unstable continuous stirred tank reactor as in (Scherer *et al.*, 2002) are presented in Section 5.

2. PRELIMINARIES

Consider a discrete system described by

$$x(k+1) = Ax(k) + Bw(k) + B_u(k) \quad \text{(1a)}$$
$$z(k) = Cx(k) + Dw(k) + D_u u(k) \quad \text{(1b)}$$

subject to control constraints

$$|u_i(k)| \le u_{i,max}, \forall k \ge 0, i = 1, 2, \cdots, m_2. \quad (2)$$

Here $x \in \mathbb{R}^n$ denotes the states, $w \in \mathbb{R}^{m_1}$ the external disturbances, $u \in \mathbb{R}^{m_2}$ the control inputs and $z \in \mathbb{R}^p$ the controlled outputs.

With state-feedback control $u = Kx$, the closed-loop system is

$$x(k+1) = A_{cl}x(k) + Bw(k) \quad \text{(3a)}$$
$$z(k) = C_{cl}x(k) + Dw(k) \quad \text{(3b)}$$

where $A_{cl} = A + B_u K$ and $C_{cl} = C + D_u K$. Let us briefly recap the case without control constraints. The discrete time closed-loop L_2-gain from w to z is smaller than γ if and only if there exists a symmetric $P > 0$ such that

$$\begin{pmatrix} P & 0 & A_{cl}^T P & C_{cl}^T \\ 0 & \gamma^2 I & B^T P & D^T \\ PA_{cl} & PB & P & 0 \\ C_{cl} & D & 0 & I \end{pmatrix} > 0 \quad (4)$$

It is easily seen that (4) implies Schur stability of A_{cl}, and with $V(x) = x^T P x$ one easily obtains the dissipation inequality

$$V(x(k)) + \sum_{i=0}^{k-1} \left(\|z(i)\|^2 - \gamma^2 \|w(i)\|^2\right) \le V(x(0)) \quad (5)$$

for any trajectory $x(\cdot), w(\cdot)$ of the closed-loop system (3). Due to $V(x) \ge 0$, for $x(0) = 0$ we can conclude that the discrete L_2-gain of the closed-loop system is not larger than γ. With the substitution $Q = P^{-1}$ and $Y = KQ$ and by performing a congruence transformation with $diag(Q, I, Q, I)$, (4) is equivalent to

$$\begin{pmatrix} Q & * & * & * \\ 0 & \gamma^2 I & * & * \\ AQ + B_u Y & B & Q & * \\ CQ + D_u Y & D & 0 & I \end{pmatrix} > 0 \quad (6)$$

which is an LMI in γ^2, Q, Y. Let γ_{opt} denote the infimal value for which (6) with $P > 0$ is feasible.

Let us now come back to the case with control constraints. For this purpose we assume that the disturbance energy is bounded as

$$\sum_{i=0}^{\infty} \|w(i)\|^2 \le \alpha^2. \quad (7)$$

Due to (5), the output energy is bounded as

$$\sum_{i=0}^{\infty} \|z(i)\|^2 \le r \quad (8)$$

and the state trajectory remains in the ellipsoid

$$\mathcal{E}_1(P, r) := \{x \in \mathbb{R}^n : V(x) \le r\} \quad (9)$$

if the initial state $x(0)$ is contained in the ellipsoid

$$\mathcal{E}_2(P, r, \alpha) := \{x \in \mathbb{R}^n : \gamma^2\alpha^2 + V(x) \le r\}. \quad (10)$$

Exploiting $u = YQ^{-1}x$, we infer (Boyd *et al.*, 1994)

$$\max_{k \ge 0} |u_i(k)|^2 = \max_{k \ge 0} \left|\left(YQ^{-1}\right)_i x(k)\right|^2$$
$$\le \max_{x \in \mathcal{E}_1} \left|\left(YQ^{-1}\right)_i x\right|^2 \le r \left\|\left(YQ^{-\frac{1}{2}}\right)_i\right\|_2^2. \quad (11)$$

Therefore the control constraints (2) can be enforced by guaranteeing that Q and Y also satisfy

$$\begin{pmatrix} \frac{1}{r}X & Y \\ Y^T & Q \end{pmatrix} \ge 0, \quad X_{ii} \le u_{i,max}^2 \quad (12)$$

for some X. We note that (12) is an LMI in X, Y, Q for fixed r, and that the constraint $\xi \in \mathcal{E}_2(P, r, \alpha)$ can as well be re-formulated as

$$\begin{pmatrix} r - \gamma^2\alpha^2 & \xi^T \\ \xi & Q \end{pmatrix} \ge 0 \quad (13)$$

which is an LMI in γ^2 and Q for fixed α.

This leads to an algorithm for solving the constrained L_2-gain attenuation problem. For fixed α, r and $\xi = x(0)$ solve the LMI optimization problem

$$\min_{\gamma^2, Q, Y, X} \gamma^2 \text{ subject to (6), (12) and (13).} \quad (14)$$

Suppose that (γ_0, Q_0, Y_0) is an (almost) optimal solution of (14). If the system is controlled with the state-feedback gain $K_0 = Y_0 Q_0^{-1}$, we conclude with $P_0 = Q_0^{-1}$ from the above discussion that

- the control constraints (2) are respected for all disturbances satisfying (7);
- the disturbances are attenuated in the sense of

$$\sum_{i=0}^{\infty} \left(\|z(i)\|^2 - \gamma_0^2 \|w(i)\|^2\right) \leq x(0)^T P_0 x(0).$$

Remarks.

- In this construction the bound α reflects the a priori knowledge on the disturbance, whereas both the output energy bound r and the corresponding optimal value $\gamma_0 = \gamma_0(\alpha, r)$ are measures for disturbance attenuation. It is simple to extract various limits of these two parameters for feasibility of (14), such as $r \geq \gamma_{opt}^2 \alpha^2$ (where the choice with equality is too ambitious due to control constraints).
- If $\xi = x(0) = 0$, the optimal value γ_c of problem (14) satisfies $\gamma_{opt} \leq \gamma_c$, reflecting a performance degradation due to control constraints. Moreover it follows from (13) that $\gamma_c \leq \gamma_0$ which relates to a further performance degradation due to non-zero initial conditions.

The above construction is a pretty standard approach to guaranteeing disturbance attenuation by constrained control. It clearly reflects an inherent trade-off between satisfying the constraints and achieving high controller performance. If having to be prepared for unforeseen large disturbances one has to choose a large value of α (and hence large r) which leads to large γ or low performance, even if the actual disturbance affecting the system is rather mild and admits a smaller bound on its energy. On the other hand, enforcing high performance levels (small γ) requires to either reduce α or r, which might result in control constraint violation in case that the system is affected by unexpectedly large disturbances. This motivates an on-line scheme to trade-off the satisfaction of constraints and the level of performance. To this end, the moving horizon strategy, which is well-known in the literature of model predictive control, serves as a candidate.

3. MOVING HORIZON STRATEGY

The basis of the moving horizon strategy in model predictive control is solving an optimal control problem on-line at each sampling time, updated by the new measurement (Mayne *et al.*, 2000).

Exploiting the moving horizon strategy, one would solve on-line the LMI optimization problem (14) with the actual state $x(k)$ at each time k, which contains the past information on internal dynamics, external disturbances and controls. In this scheme, the current state $x(k)$ serves as "feedback" not only for the computation of control values but also for the choice of the feedback gain. The latter provides an opportunity to trade-off constraint satisfaction and performance. By minimizing the performance level γ on-line, one obtains the best possible performance, while keeping constraints satisfied. Unfortunately, this simple implementation of the moving horizon strategy might fail to guarantee dissipation for the controlled system, as shown in detail in (Scherer *et al.*, 2002). In the same reference it is also shown how to recover dissipation, and we repeat in this paper the key points for the discrete-time problem and in view of its implementation in a moving horizon scheme.

Assume that the LMI optimization problem (14) admits a solution for the closed-loop state $x(k)$ at each sampling time k, denoted as (γ_k, Q_k, Y_k). The feedback control is defined by

$$u(k) = K_k x(k), \; k = 0, 1, 2, \cdots$$

with $K_k = Y_k P_k$ and $P_k = Q_k^{-1}$.

At time $k = 0$, according to the principle of moving horizon strategy, $u(0) = K_0 x(0)$ will be applied to the system until the next sampling instant $k = 1$. With the actual state $x(1)$ as initial condition, the LMI optimization problem (14) will be solved again. Let us investigate whether the solution at time $k = 1$ keeps the closed-loop system dissipative. We first observe that

$$\|z(0)\|^2 - \gamma_0^2 \|w(1)\|^2 \leq x(0)^T P_0 x(0) - x(1)^T P_0 x(1)$$
$$\|z(1)\|^2 - \gamma_1^2 \|w(1)\|^2 \leq x(1)^T P_1 x(1) - x(2)^T P_1 x(2)$$

and hence

$$\sum_{i=0}^{1} \|z(i)\|^2 - \max\{\gamma_0, \gamma_1\}^2 \|w(i)\|^2 \leq x(0)^T P_0 x(0) - \\ - \left[x(1)^T P_0 x(1) - x(1)^T P_1 x(1))\right] - x(2)^T P_1 x(2).$$

If $\left[x(1)^T P_0 x(1) - x(1)^T P_1 x(1)\right] \geq 0$, dissipation holds with level $\max\{\gamma_0, \gamma_1\}$. The solution of the LMI optimization problem (14) at the time $k = 2$ with $x(2)$ leads in a similar fashion to

$$\sum_{i=0}^{2} \left[\|z(i)\|^2 - \max\{\gamma_0, \gamma_1, \gamma_2\}^2 \|w(i)\|^2\right] \leq \\ \leq x(0)^T P_0 x(0) - \left[x(1)^T P_0 x(1) - x(1)^T P_1 x(1))\right] - \\ - \left[x(2)^T P_1 x(2) - x(2)^T P_2 x(2))\right] - x(3)^T P_2 x(3).$$

To guarantee dissipation one requires

$$x(0)^T P_0 x(0) - \left[x(1)^T P_0 x(1) - x(1)^T P_1 x(1))\right] - \\ - \left[x(2)^T P_1 x(2) - x(2)^T P_2 x(2))\right] \le x(0)^T P_0 x(0).$$

In general, just solving the LMI optimization problem (14) at times $k = 1$ and $k = 2$ with respective initial conditions does not guarantee this inequality to hold. Therefore, the naive implementation of the moving horizon strategy will generally fail. However, the discussion reveals the crucial strategy to guarantee dissipation as follows: Define

$$p_{k-1} := x(0)^T P_0 x(0) - \\ - \sum_{j=1}^{k-1} [x(j)^T P_{j-1} x(j) - x(j)^T P_j x(j)]. \quad (15)$$

For dissipation one has to enforce at iteration k that

$$p_{k-1} - \left[x(k)^T P_{k-1} x(k) - x(k)^T P_k x(k)\right] \le p_0. \quad (16)$$

Moreover p_k can be recursively updated as

$$p_k := p_{k-1} - \left[x(k)^T P_{k-1} x(k) - x(k)^T P_k x(k)\right].$$

It is easy to include the dissipation constraint (16) in the optimization problem (14) to end up with the following extended LMI problem at time k with the actual state $x(k)$:

$$\min_{\gamma^2, Q, Y, X} \gamma^2 \quad (17)$$

subject to (6), (12), (13) for $\xi = x(k)$, and

$$\begin{pmatrix} p_0 - p_{k-1} + x(k)^T P_{k-1} x(k) & x(k)^T \\ x(k) & Q \end{pmatrix} \ge 0. \quad (18)$$

The implementation of this on-line scheme is possible since P_{k-1} and p_{k-1} have been determined at the previous time instant $k-1$ and are held fixed. Let us suppose that (17) admits an (almost) optimal solution (γ_k, Q_k, Y_k) and define the feedback gain $K_k = Y_k Q_k^{-1}$ as well as $P_k = Q_k^{-1}$. Controlling the system with $u(k) = K_k x(k)$ then implies that

- the control constraints (2) are respected;
- the controller automatically relaxes the performance requirement if necessary not to violate constraints and it enhances the performance level if possible and in such a manner that the closed-loop system is guaranteed to obey the dissipation inequality

$$\sum_{i=k}^{l} \|z(i)\|^2 - \gamma^2 \|w(i)\|^2 \le x(k)^T P_k x(k)$$

for $0 \le k \le l$ and with $\gamma = \max\{\gamma_k, \dots, \gamma_l\}$.

Let us stress that the feature of automatic performance adaptation is viewed to be the most relevant progress over (Scherer *et al.*, 2002). Moreover we recall that the off-line controller construction in (Scherer *et al.*, 2002) was based on extremal solutions of the Riccati equation corresponding to the $\mathcal{H}_\infty$ problem, whereas the present scheme picks the solution (shapes of ellipsoids) depending on the individual system state which leads to performance improvements.

For the actual on-line implementation of this scheme it is essential that the LMI optimization problem (17) is feasible at each time-instant k, which gives rise to the need for an on-line adaptation of the parameters α and r as suggested in the algorithm in the next section. If the LMI's are not feasible for all combinations of α and r one could either relax the control constraint to enforce feasibility (which is always successful for stabilizable systems but which might not be practically possible) or one could switch to a standard MPC scheme with quadratic cost which incurs a loss of guaranteed disturbance suppression properties.

4. ALGORITHM FOR MOVING HORIZON IMPLEMENTATION

Let us now discuss a concrete implementation of the suggested scheme, together with one out of a multitude of possibilities how to adapt the parameters α and r. In fact we keep α fixed while we try to enforce feasibility of (17) by increasing r (from a given r_0) whenever necessary. Moreover, for the given α, r_0 and with $x(0) = 0$, we consider the controller $K_c = K_c(\alpha, r_0)$ - defined with $P_c = P_c(\alpha, r_0)$ - as the one with best performance, and at each time k we first check whether this best gain guarantees dissipation and constraint satisfaction in order to avoid unnecessary on-line computations.

Algorithm

Step 1 Initialization. Let α and r_0 be given. Solve the LMI optimization problem (14) with $\xi = x(0) = 0$ and compute $K_c = YQ^{-1}$ and $P_c = Q^{-1}$.

Step 2 At time $k = 0$, set $r = r_0$. If $x(0) = 0$, set $K_0 = K_c$, $P_0 = P_c$, $p_0 = 0$ and go to Step 6. If $x(0) \ne 0$, solve the LMI optimization problem (14) with $\xi = x(0)$. If it admits a solution, compute $K_0 = YQ^{-1}$, $P_0 = Q^{-1}$, $p_0 = x(0)^T P_0 x(0)$ and go to Step 6. If not feasible, increase r until feasibility is retained.

Step 3 At time k, set $r = r_0$. If $x(k) \in \mathcal{E}_2(P_c, r_0, \alpha)$ and $p_{k-1} - x(k)^T P_{k-1} x(k) + x(k)^T P_c x(k) \le p_0$, then set $K_k = K_c$, $P_k = P_c$, and go to Step 5.

Step 4 Solve the LMI optimization problem (17) with $\xi = x(k)$. If it admits a solution, compute $K_k = YQ^{-1}$, $P_k = Q^{-1}$, and go to Step 5. If not feasible, increase r and repeat Step 4.

Step 5 Prepare for the next computation:
$$p_k = p_{k-1} - \big[x(k)^T P_{k-1} x(k) - x(k)^T P_k x(k)\big].$$
Step 6 Apply $u(k) = K_k x(k)$ to control the system. Replace k by $k+1$ and continue with Step 3.

5. EXAMPLE: CONTROL OF AN UNSTABLE CSTR

We take the same example as in (Scherer *et al.*, 2002) for demonstrating the proposed moving horizon scheme. This is a continuous stirred tank reactor, in which the substance B is produced from the initial reactant A in the main reaction, and unwanted parallel and consecutive reactions form by-products D and C, as A $\xrightarrow{r_1}$ B $\xrightarrow{r_3}$ C and A $\xrightarrow{r_2}$ D. The reaction velocities r_i are assumed to depend on the concentration and/or the temperature nonlinearly. The inflow of the CSTR contains only the substance A and is assumed to come from an upstream unit. Therefore, the concentration and temperature in the inflow can be viewed as external disturbances. The control objective is to maintain the concentration of the main product B despite these inflow variations. As control inputs we may choose the inflow rate normalized by the reactor volume and the heat removal, which suffer saturation. A more detailed description of the CSTR can be found in (Allgöwer, 1996).

We discretize the linearized model given in (Scherer *et al.*, 2002) with a sampling time of $\delta = 0.1\,\text{min}$. We obtain a system in the form of (3) with

$$A = \begin{pmatrix} 0.9739 & -0.0942 & -0.4378 \\ -0.0012 & 1.0321 & 0.1567 \\ -0.0162 & 0.0640 & 1.0648 \end{pmatrix}$$
$$(B_1|B_u) = \left(\begin{array}{cc|cc} 0.0592 & -0.0017 & 0.0022 & 0.0502 \\ 0 & 0.0006 & -0.0008 & -0.0103 \\ -0.0005 & 0.0082 & -0.0103 & -0.0028 \end{array}\right),$$

where $x \in \mathbb{R}^3$ represent the normalized concentrations of substances A and B, and the normalized reaction temperature, respectively; $w \in \mathbb{R}^2$ and $u \in \mathbb{R}^2$ denote the normalized disturbances and controls, respectively. It is assumed that controls are bounded as $|u_i(k)| \leq 1,\ \forall k \geq 0,\ i = 1, 2$. We further choose the same controlled output $z = \text{col}(Hx, Eu)$ as in (Scherer *et al.*, 2002) with $H = \text{diag}(0.5, 1, 1)$ and $E = \text{diag}(0.1, 0.1)$. An (almost) optimal attenuation level for the unconstrained $\mathcal{H}_\infty$ problem is $\gamma_{opt} = 0.1819$.

Let us assume that the disturbance could be occasionally very large and the energy is bounded as $\sum_{i=0}^{i=\infty} \|w(i)\|^2 \leq 6$. Following the algorithm given in Section 4, we implement a moving horizon controller with $\alpha = 0.1$ and $r_0 = 200\gamma_{opt}^2\alpha^2$. A much smaller α is chosen, since it is allowed for the moving horizon controller and leads to better performance. According to the discussion in Section 3, the moving horizon controller respects the control constraints while keeping the closed-loop system dissipative. For reasons of comparison, we design a fixed controller by solving LMI optimization problem (14) with $\alpha_f^2 = 6$, $r_f = 4.6\gamma_{opt}^2\alpha_f^2$ and $\xi = x(0) = 0$. The subscript f is affixed for the fixed controller. This design ensures that for any disturbance with energy bounded by 6, the fixed controller satisfies the control constraints and admits a performance level of $\gamma_f = 0.3893$.

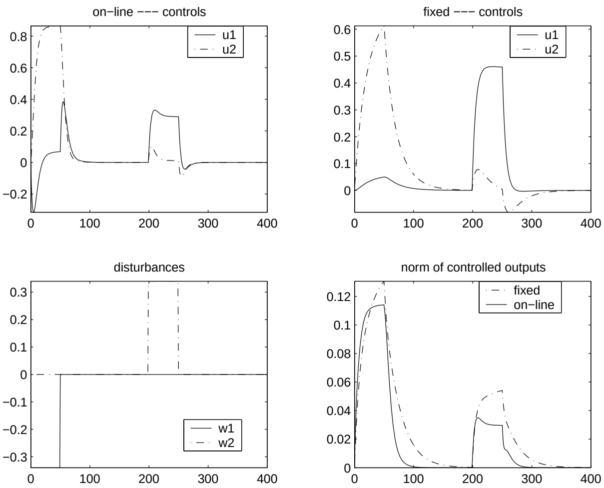

Fig. 1. Comparison of disturbance attenuation. Both moving horizon and fixed controllers guarantee dissipation and constraint satisfaction.

Fig. 1 shows the results of attenuating an impulse variation in inflow concentration and inflow temperature, respectively. The impulse is with a width of 50 sampling periods and an energy of about 6. For these disturbances, no on-line adaptation of γ happens in the moving horizon controller, nevertheless performance improvement over the fixed controller can be clearly seen in the bottom-right picture of Fig. 1, which is achieved by allowing to choose smaller α.

When unexpected stronger disturbances affect the systems, the fixed controller may violate the hard constraints. In this case, we just clip the control signals to keep them within bounds, which implies the loss of dissipation guarantee for the fixed controller. Fig. 2 and Fig. 3 present the results for such disturbances, from both the moving horizon controller and the fixed controller. The disturbances consist of a sinusoidal variation and an impulse with high intensity as plotted in the bottom-left picture of Fig. 2 and as defined by

$$w_1(k) = \begin{cases} s_1 + \bar{w}_1(k) & \text{for } 0 \leq k \leq 50 \\ \bar{w}_1(k) & \text{for } k > 50 \end{cases}$$
$$\bar{w}_1(k) = a_1 \sin(-0.024(k+100)) \sin(0.2(k+100))$$
$$w_2(k) = \begin{cases} \bar{w}_2(k) & \text{for } 0 \leq k < 200 \\ s_2 + \bar{w}_2(k) & \text{for } 200 \leq k \geq 250 \\ \bar{w}_2(k) & \text{for } k > 250 \end{cases}$$
$$\bar{w}_2(k) = a_2 \sin(-0.024(k+110)) \cos(0.2(k+110))$$

with $a_1 = a_2 = 0.02$, $s_1 = -0.41$ and $s_2 = 0.75$. Automatic performance adaptations are indicated clearly in the bottom-right picture of Fig 2, which leads to better performance of the moving horizon controller as illustrated in the bottom-right picture of Fig 3. More precisely, the moving horizon controller makes the best of the control constraints to achieve the improvement during the first impulse; relaxing on-line performance so as to avoid actuator saturation leads to the improvement during the second impulse. Moreover, the performance improvement around $k = 100$ and $k = 300$ is obtained by tightening the performance specification when the impulse variations in the inflow are removed.

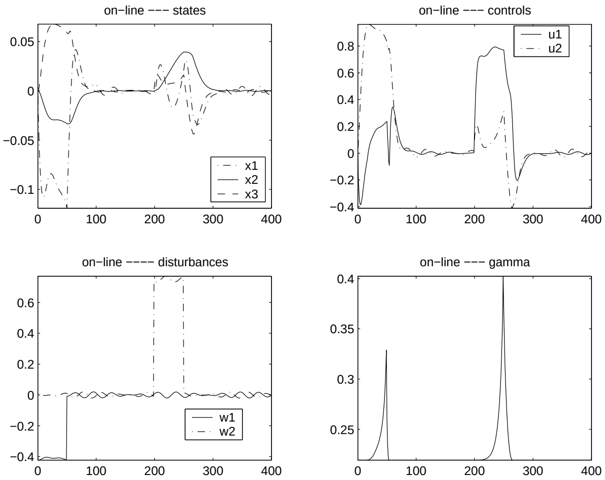

Fig. 2. Responses for moving horizon controller.

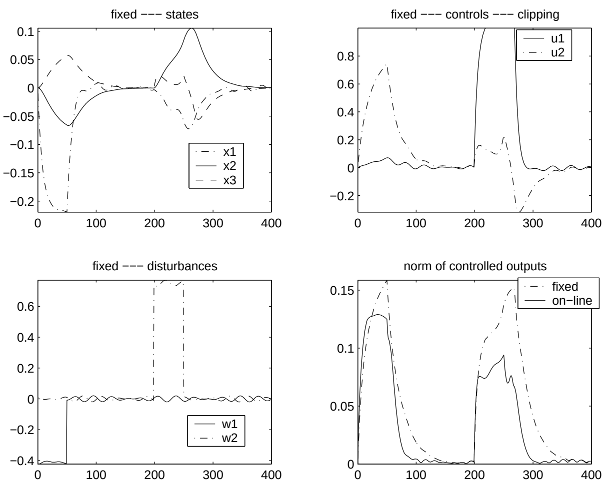

Fig. 3. Responses for fixed controller and comparison of disturbance attenuation with moving horizon controller.

6. CONCLUSIONS

In combining the moving horizon paradigm with dissipation theory, we proposed in this paper an on-line optimization scheme to solve the L_2-gain attenuation problem with hard control constraints. Technically, the feedback gain is determined on-line by solving a constrained $\mathcal{H}_\infty$ control problem updated by the actual state, while a dissipation constraint is introduced to guarantee disturbance attenuation for the closed-loop system. This scenario automatically manages the trade-off between satisfying constraints and achieving high performance, which is viewed as the most relevant progress over (Scherer *et al.*, 2002) with corresponding improvements of performance.

It should be pointed out that the tuning mechanism for the parameters α and r in the proposed scheme requires further investigation. If the disturbances are not directly feed through to the outputs, it is straightforward to include output constraints as well, and other extensions pointed out in (Scherer *et al.*, 2002) are under consideration.

Acknowledgment

The authors thank Prof. F. Allgöwer, Institute for Systems Theory in Engineering, University of Stuttgart, for useful discussions. H. Chen thanks the National Natural Science Foundation of China (No.60234030) and the Ministry of Education of China (Key Project No.00038) for financial supports.

REFERENCES

Allgöwer, F. (1996). *Näherungsweise Ein-/ Ausgangs-Linearisierung nichtlinearer Systeme.* Fortschr.-Ber. VDI Reihe 8 Nr. 582, VDI Verlag. Düsseldorf.

Boyd, S., L. El Ghaoui, E. Feron and V. Balakishnan (1994). *Linear Matrix Inequalities in System and Control Theory.* SIAM. Philadelphia.

Gilbert, E.C. and K.C. Tan (1991). Linear systems with state and control constraints: The theory and application of maximal output admissible sets. *IEEE Trans. Automat. Contr.* **AC-36**, 1008–1020.

Hirata, K. and M. Fujita (2000). Measurement feedback switching control for systems with state and control constraints. In: *Proc. Amer. Contr. Conf.*. Chicago, Illinois. pp. 3321–3325.

Kothare, M. V., P. J. Campo, M. Morari and C. N. Nett (1994). A unified framework for the study of anti-windup designs. *Automatica* **30**(12), 1869–1883.

Kothare, M. V., V. Balakrishnan and M. Morari (1996). Robust constrained model predictive control using linear matrix inequalities. *Automatica* **32**(10), 1361–1379.

Mayne, D.Q., J.B. Rawlings, C.V. Rao and P.O.M. Scokaert (2000). Constrained model predictive control: Stability and optimality. *Automatica* **36**, 789–814.

Scherer, C. W., H. Chen and F. Allgöwer (2002). Disturbance attenuation with actuator constraints by hybrid state-feedback control. In: *Proc. 41th IEEE Conf. Decision Contr.*

AN LMI-BASED CONSTRAINED MPC SCHEME WITH TIME-VARYING TERMINAL COST

B. Pluymers, L. Roobrouck, J. Buijs
J.A.K. Suykens, B. De Moor

Katholieke Universiteit Leuven
Department of Electrical Engineering, ESAT-SCD-SISTA
Kasteelpark Arenberg 10, B-3001 Heverlee (Leuven), Belgium
E-Mail : {bert.pluymers, luc.roobrouck, jeroen.buijs, johan.suykens, bart.demoor}@esat.kuleuven.ac.be

Abstract: Modelbased Predictive Control (MPC) is a control technique that is widely used in chemical process industry. In the past decade, stability of MPC has been an intensive research area, resulting in the general acceptance of a theoretical MPC stability framework introducing a terminal cost and terminal constraint to the classic MPC formulation. Although guaranteeing stability, issues regarding optimality and feasibility remain. In this paper, an LMI-based constrained MPC scheme for linear systems is introduced which guarantees stability by use of a time-varying terminal cost and terminal constraint. The online calculation of the terminal cost results in improved performance and feasibility compared to MPC schemes with fixed terminal cost. Finally, the technique is illustrated on a copolymerization reactor.

Keywords: modelbased predictive control, linear matrix inequalities, feasibility, stability, optimality, constraints, LQR

1. INTRODUCTION

In the past decade, wide consensus has been reached with respect to achieving stability of MPC controllers (Mayne et al. (2000)). The common denominator of stabilizing MPC schemes has been recognized to be the addition of a terminal cost and terminal constraint to the classic MPC formulation. It has also been recognized, however, that the choice of these two additions influences feasibility and local optimality of the resulting controller.

In this paper a new MPC scheme with time-varying terminal cost and constraint will be introduced which results in a controller that is locally optimal while conserving feasibility over a wide range of states. The approach taken here is to use a special case of the LMI formulations introduced in Kothare et al. (1996) and later used in Lee et al. (1998) to ensure stability for linear, time-varying systems. It is then shown that these LMI's can be merged with the classic MPC optimization problem, which results in an LMI-based MPC scheme that calculates in each time step both the optimal inputs and an optimal, stabilizing terminal cost and constraint.

This paper is organized as follows. In a first part a general introduction to linear MPC will be given, after which, more specifically, the established stability theory of MPC will be discussed. The third part explains how a stabilizing terminal cost and terminal constraint can be calculated using LMI's and how these stabilizing ingredients influence optimality. In a fifth part, the MPC scheme with time-varying terminal cost will be introduced, which is then illustrated on a copolymerization reactor in the final section.

2. MODELBASED PREDICTIVE CONTROL

Modelbased Predictive Control (MPC) is a control scheme that calculates in each time step an optimal future input sequence $u_{k,i}$, $i = 0 \dots P-1$ by solving an optimization problem. In this paper only linear MPC (i.e. using linear models) will be discussed, in which case the problem reduces to a Quadratic Program (QP) :

$$\min_{x,u} \; J_k = \min_{x,u} \left(\sum_{i=0}^{M-1} u_{k,i}^{\mathrm{T}} R u_{k,i} + \sum_{i=1}^{P} (x_{k,i} - x_{k,i}^{\star})^{\mathrm{T}} Q (x_{k,i} - x_{k,i}^{\star}) \right) \quad \text{(1a)}$$

subject to

$$x_{k,i+1} = Ax_{k,i} + Bu_{k,i}, \quad i = 0, \dots P-1. \quad \text{(1b)}$$

$x_k \equiv x_{k,0}$ denotes the states measured at time k. $x_{k,i}, i > 0$ denote the states at time $k+i$ as predicted at time k, $x_{k,i}^{\star}, i \geq 0$ denote the future states at time $k+i$ as desired at time k and $u_{k,i}, i \geq 0$ denotes the input sequence as calculated at time k. Q and R are positive definite weighting matrices indicating the relative importance of the states and inputs in the control problem. M and P denote respectively the number of future inputs and states that are calculated in each time step and are called the *control horizon* and the *input horizon*. M and P are generally given equal values, which is what will be assumed in this paper. Equality constraints (1b) represent the system behavior, with given system matrices A and B.

From the calculated input sequence only the first input $u_k \equiv u_{k,0}$ is applied to the system, after which the calculation restarts in the next time step with a new measured state $x_{k+1,0}$.

Typically additional linear inequality constraints are incorporated in the QP, representing either physical system limitations, safety margins or economical constraints. Depending on the variables involved these are called *input constraints* or *state constraints* and will be denoted by $u_k \in \mathbb{U}$ and $x_k \in \mathbb{X}$ respectively.

3. STABILITY THEORY

Although MPC calculates in each time step an *optimal* input sequence, there is no guarantee that the controller is stable. In each time step only the first input of the calculated sequence is applied to the system and there is no guarantee that this sequence of first inputs is optimal in any way. The fundamental cause of this problem is the fact that the horizon P is finite, and the behavior of the system for $i \geq P$ is not taken into account in the optimization problem. Several MPC variants were developed to address this issue (Mayne et al. (2000); Rawlings and Muske (1993); Lee et al. (1998)). Most of these methods fit into the following stability framework.

Assuming $x_{k,i}^{\star} \equiv 0$ (or a previous transformation accomplishing this), a stabilizing MPC controller can be obtained by modifying the weight matrix of the terminal state $x_{k,P}$ into Q_P (called the *terminal cost*), the total control cost of a locally stabilizing feedback controller $\kappa(x) = Kx$ for $i \geq P$, and adding a terminal state constraint $x_{k,P} \in \mathbb{X}_P$ (also called *terminal constraint*) corresponding to the region in which the terminal cost is valid. More formally, asymptotic stability is achieved when following conditions are satisfied :

a. $\mathbb{X}_P \subset \mathbb{X}$ (2a)

b. $\kappa_N(x) \in \mathbb{U}, \forall x \in \mathbb{X}_P$ (2b)

c. $f(x, \kappa(x)) \in \mathbb{X}_P, \forall x \in \mathbb{X}_P$ (2c)

d. $F(x) - F(f(x, \kappa(x))) - l(x, \kappa(x)) \geq 0,$
$\forall x \in \mathbb{X}_P$ (2d)

given there exists a feasible solution to (1) supplemented with the terminal constraint and terminal cost. $f(x,u) \equiv Ax + Bu$ denotes the system state transition function, $l(x,u)$ denotes the cost function $x^{\mathrm{T}}Qx + u^{\mathrm{T}}Ru$ for all $i = 1, 2, \dots, P-1$ and $F(x)$ the cost term $x^{\mathrm{T}}Q_P x$ for the terminal state.

Note that these conditions are sufficient but not necessary. Most stabilizing MPC schemes make different choices for $\mathbb{X}_P$ and K to satisfy these conditions. When they are satisfied $J_k^o = \min J_k$ can be proven to be monotonically descending to 0, so it can be used as a Lyapunov function to prove asymptotic stability.

In the unconstrained case (no input or state constraints), the above conditions can be easily satisfied by choosing $\mathbb{X}_P = \mathbb{R}$ and $K = K_{\mathrm{lqr}}$. The solution S to the corresponding Riccatti equation (Kalman (1960)) can then be used as terminal cost (Bitmead et al. (1990)).

4. DETERMINING TERMINAL COST USING LMI'S

The forementioned method to calculate a terminal cost cannot be used in the constrained case, in which case the constraints have to be accounted for explicitly. As proposed in Boyd et al. (1994) and Kothare et al. (1996), a linear, stabilizing feedback controller, which is optimal in an LQR

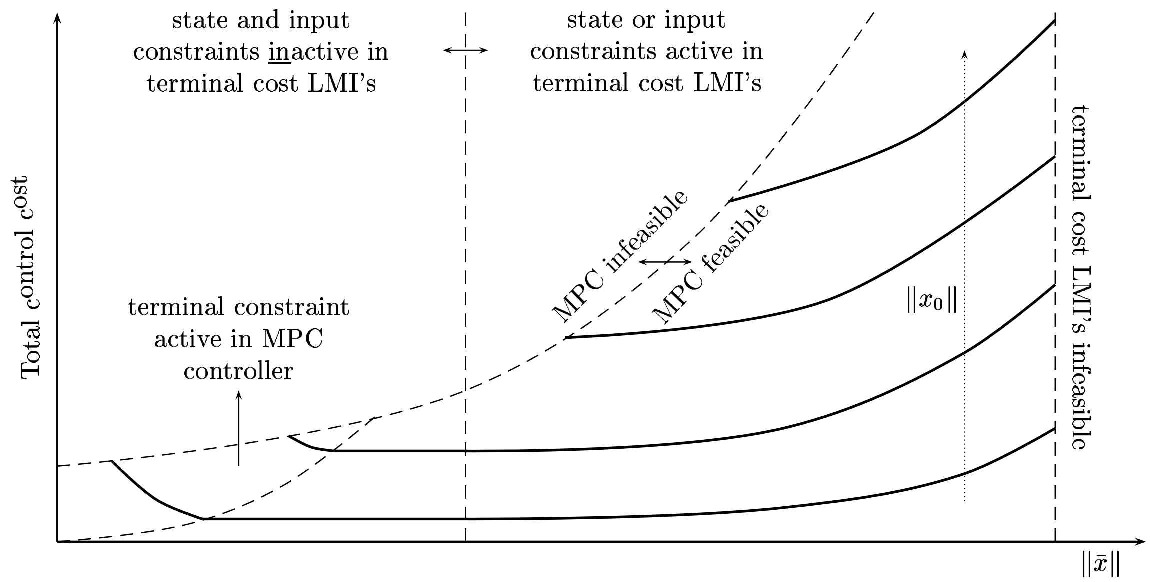

Fig. 1. Total control cost (full line) of recovering from an initial disturbance x_0 for a typical system, in function of $\bar{x}$ and for different disturbance sizes. Different regions of operation are delimited by dashed lines. It is clear that performance is very dependent on the choice of $\bar{x}$. A compromise has to be made between feasibility (larger $\bar{x}$) and optimality (smaller $\bar{x}$).

sense and which respects input and state constraints, can be found by solving a linear optimization problem with LMI constraints :

$$\min_{\gamma,Z,Y,X} \gamma \tag{3a}$$

subject to stability constraints

$$\begin{bmatrix} 1 & \bar{x}^{\mathrm{T}} \\ \bar{x} & Z \end{bmatrix} \geq 0 \tag{3b}$$

$$\begin{bmatrix} Z & (AZ+BY)^{\mathrm{T}} & ZQ^{\frac{1}{2}} & Y^{\mathrm{T}}R^{\frac{1}{2}} \\ AZ+BY & Z & 0 & 0 \\ Q^{\frac{1}{2}}Z & 0 & \gamma I & 0 \\ R^{\frac{1}{2}}Y & 0 & 0 & \gamma I \end{bmatrix} \geq 0 \tag{3c}$$

$$Z > 0, \tag{3d}$$

input constraints

$$\begin{bmatrix} X & Y \\ Y^{\mathrm{T}} & Z \end{bmatrix} \geq 0 \tag{3e}$$

$$\begin{bmatrix} u_{1,\max}^2 & & & \\ & u_{2,\max}^2 & & \\ & & \ddots & \\ & & & u_{n_u,\max}^2 \end{bmatrix} \geq X \tag{3f}$$

and state constraints

$$\begin{bmatrix} Z & (AZ+BY)^{\mathrm{T}}C_j^{\mathrm{T}} \\ C_j(AZ+BY) & y_{j,\max}^2 \end{bmatrix} \geq 0$$

$$j = 1 \ldots n_{\mathrm{sc}} \tag{3g}$$

after which the feedback matrix and terminal cost can be calculated as

$$K = YZ^{-1} \tag{4a}$$

$$Q_P = \gamma Z^{-1}. \tag{4b}$$

$\bar{x}$ is an arbitrary state, representing typical excitations, to be chosen by the user.

Remark 1. The resulting linear feedback controller minimizes $\bar{x}^{\mathrm{T}}Q_P\bar{x}$, representing the total control cost of the controller to achieve equilibrium from initial condition $\bar{x}$, while respecting input and state constraints

$$|[u]_i| \leq u_{i,\max} \quad i = 1 \ldots n_u$$
$$|C_j x| \leq y_{j,\max} \quad j = 1 \ldots n_{\mathrm{sc}}$$

for all initial states x satisfying

$$||x||_{Q_P} \leq ||\bar{x}||_{Q_P} \tag{5}$$

with n_u denoting the number of inputs and n_{sc} denoting the number of state constraints.

Remark 2. For sufficiently small values of $\bar{x}$, where input and state constraints are not active, the resulting K and Q_P can be shown to be identical to those obtained by calculating an LQR controller with the help of Riccatti equations. In this case, the terminal cost exactly represents the remaining control cost beyond the horizon which leads to a locally optimal controller.

Remark 3. For sufficiently large values of $\bar{x}$, it will not be possible to find a stabilizing linear feedback controller which still respects the input and state constraints. In this case optimization problem (3) won't have a feasible solution.

Remark 4. (5) can be shown to satisfy conditions (2a), (2b) and (2c) for $\mathbb{X}_P$, so it can be used as terminal constraint. A more thorough proof is given in Pluymers et al. (2003). This constraint is not linear but quadratic, which can complicate the optimization problem. Two arguments can be given however to relativate this. First of all, the constraint can be approximated by a set of linear constraints, which again reduces the problem to a QP. Secondly, because of the fact that Q_P is a positive definite matrix, the terminal constraint is elliptic, thus convex, so the use of efficient convex optimization algorithms is still possible.

5. OPTIMALITY

In the previous section an approach is explained to calculate a stabilizing terminal cost and terminal constraint for linear MPC with input and state constraints using LMI's. One aspect that has not been clarified, however, is the choice of $\bar{x}$. It is clear that, on the one hand, $\bar{x}$ should be chosen small enough to make sure (3) has a feasible solution, while, on the other hand, $\bar{x}$ should still be chosen large enough to make sure (5) isn't overrestrictive which can result in an infeasible MPC optimization problem.

Apart from feasibility, which is a necessary condition for stability, another important feature which is influenced by the choice of $\bar{x}$ is the performance of the controller. This is shown in figure 1. It can be observed that feasibility (large $\bar{x}$) and optimality (small $\bar{x}$) cannot be achieved simultaneously for all $\|x_0\|$. One should thus choose the smallest $\bar{x}$ that still results in a feasible MPC controller for all disturbances that can be expected.

It is clear that the use of a time-varying terminal cost and terminal constraint could result in improved performance, while preservering feasibility of the controller. A technique accomplishing exactly this is proposed in the next section.

6. MPC WITH TIME-VARYING TERMINAL COST

As shown in the previous section, significant performance improvements can potentially be achieved by adaptively choosing $\bar{x}$ depending on the current state of the system. The strategy chosen in this paper, is to incorporate the choice of $\bar{x}$ directly into the MPC optimization problem

$$\min_{x,u,\bar{x}} J_k = \min_{x,u,\bar{x}} \left(\sum_{i=0}^{P-1} u_{k,i}^{\mathrm{T}} R u_{k,i} + \sum_{i=1}^{P-1} x_{k,i}^{\mathrm{T}} Q x_{k,i} + x_{k,P}^{\mathrm{T}} Q_P(\bar{x}) x_{k,P} \right) \quad (6a)$$

subject to

$$x_{k,i+1} = A x_{k,i} + B u_{k,i} \quad (6b)$$
$$u_{k,i} \in \mathbb{U} \quad (6c)$$
$$x_{k,i} \in \mathbb{X} \quad (6d)$$
$$x_{k,P}^{\mathrm{T}}\ Q_P(\bar{x})\ x_{k,P} \leq \bar{x}^{\mathrm{T}}\ Q_P(\bar{x})\ \bar{x} \quad (6e)$$

where $i = 0, \ldots P-1$ and $Q_P(\bar{x})$ explicitly denotes the dependence of the terminal cost and terminal constraint on $\bar{x}$. This way in each time step $\bar{x}$ is chosen to minimize J_k.

The above optimization problem cannot be implemented as such, due to the fact that evaluating $Q_P(\bar{x})$ in turn requires the solution of (3), which is not efficient. The approach taken here is to convert the above optimization problem into LMI's, after which these can be merged with (3), resulting in a unified (and convex) optimization problem.

Before doing this, another simplification can be made. It can be rigourously proven that in the optimum, the equality $Q(\bar{x}) = Q(x_{k,P})$ holds. In this way $\bar{x}$ can be eliminated from the optimization problem by replacing $Q_P(\bar{x})$ with $Q_P(x_{k,P})$. Consequently, the terminal constraint can be removed because it is trivially satisfied. The validity of this elimination is proven in Pluymers et al. (2003).

To convert the MPC optimization problem into an LMI problem, we first eliminate the equality constraints, because these cannot be efficiently converted to LMI's. After elimination, the MPC problem can be written as

$$\min_u u^{\mathrm{T}} K_{\mathrm{quad}} u + k_{\mathrm{lin}} u + (C_P u + D_P)^{\mathrm{T}} Q_P(C_P u + D_P)(C_P u + D_P) \quad (7a)$$

subject to

$$A_{\mathrm{ineq}} u \leq B_{\mathrm{ineq}} \quad (7b)$$

where $C_P u + D_P$ is an expression for the terminal state $x_{k,P}$, so $Q_P(C_P u + D_P)$ again expresses the dependency of Q_P on this state. This QP, without equality constraints, can be converted to a linear optimization problem with LMI constraints

$$\min_{u,\gamma_1,\gamma_2} \gamma_1 + \gamma_2 \quad (8a)$$

subject to

$$\begin{bmatrix} \gamma_1 - k_{\mathrm{lin}} u & u^{\mathrm{T}} \\ u & K_{\mathrm{quad}}^{-1} \end{bmatrix} \geq 0 \quad (8b)$$

$$\begin{bmatrix} \gamma_2 & (C_P u + D_P)^{\mathrm{T}} \\ C_P u + D_P & Q_P(C_P u + D_P)^{-1} \end{bmatrix} \geq 0 \quad (8c)$$

$$B_{\mathrm{ineq}} - A_{\mathrm{ineq}} u \geq 0. \quad (8d)$$

It can be easily seen that γ_1 represents the cost of states and input up to $i = P-1$, while γ_2 represents the cost of the terminal state $x_{k,P}\ Q_P(x_{k,P})\ x_{k,P}$. Because of the equivalence of $Q_P(x_{k,P})$ and $Q_P(\bar{x})$ this is exactly the same expression as the objective function of (3) as mentioned in remark 1. This is illustrated by the fact that the objective function of (3) is represented by (3a), which can be made equivalent with (8c) by applying (4b). The dependence of Q_P on $x_{k,P}$ can thus be explicitly incorporated into (8) by adding constraints (3b)-(3g). This results in the following optimization problem :

$$\min_{u,\gamma_1,\gamma_2,Z,Y,X} \gamma_1 + \gamma_2 \quad (9a)$$

subject to

$$\begin{bmatrix} \gamma_1 - k_{\text{lin}} u & u^{\text{T}} \\ u & K_{\text{quad}}^{-1} \end{bmatrix} \geq 0 \tag{9b}$$

$$\begin{bmatrix} 1 & (C_P u + D_P)^{\text{T}} \\ C_P u + D_P & Z \end{bmatrix} \geq 0 \tag{9c}$$

$$B_{\text{ineq}} - A_{\text{ineq}} u \geq 0. \tag{9d}$$

and (3c)-(3g).

As will be shown in the next section, this MPC scheme with time-varying terminal cost achieves better performance than the traditional MPC scheme with fixed terminal cost and terminal constraint, while preserving feasibility of the controller. The optimization problem is converted from a QP into a linear problem with LMI contraints. This is still a convex optimization problem, for which efficient algorithms exist. The disadvantage, however, is the increased number of optimization variables, which causes a significantly higher computational complexity. A more detailed analysis of the computational complexity and a way to largely eliminate this disadvantage is given in Pluymers et al. (2003).

7. EXAMPLE

To illustrate the concepts introduced in the previous sections, control of a continuously stirred copolymerization reactor is considered. The model used in this paper has already been discussed in Congalidis et al. (1989, 1986). Although the model is stable at the operating point used in this paper, the introduction of an terminal cost and the adaptive, online calculation hereof results in improved performance.

The reactor consists of a continuously stirred tank to which the reagents are fed with a feed rate chosen by the controller. The reaction product (polymers), solvent and residual reagents are simultaneously drained from the tank, after which the first is separated from the latter. From an engineering point of view, the most important measured variables are the reactor temperature $T_{\text{r}}(K)$, the polymer production rate, $G_{\text{p}}(kg/h)$, the mass fraction Y_{ap} of monomer A in the polymer and the average molar mass $M_{\text{p}}(g/mole)$ of the polymer. See table 1 for an overview of inputs and outputs and their steady state values.

The differential equations of the nonlinear reactor model as derived in Congalidis et al. (1989) were implemented in MATLAB. The model was discretized in time ($T_{\text{s}} = 300\text{s}$)by using a MATLAB differential equation solver and linearized around the operating point described in the same reference. Resulting in a linear state-space model with 6 inputs and 15 states, among which the 4 forementioned variables. This model was then normalized with respect to the steady state values ($x_{\text{n}} = (x - x_{\text{ss}})/x_{\text{ss}}$ and $u_{\text{n}} = (u - u_{\text{ss}})/u_{\text{ss}}$) to avoid numerical problems. All states were assumed to be measured.

input	s.s.-value
monomer A feed rate (G_{af})	18.00 kg/h
monomer B feed rate (G_{bf})	89.99 kg/h
initiator feed rate (G_{if})	0.18 kg/h
solvent feed rate (G_{sf})	36.02 kg/h
chain transfer agent feed rate (G_{tf})	2.70 kg/h
inhibitor mass feed rate (G_{zf})	0.0003 kg/h
output	**s.s.-value**
polymer production rate (G_{p})	23.31 kg/h
monom. A mass fract. in polym. (Y_{ap})	0.56
polymer molar mass (M_{p})	35003.48 g/mole
reactor temperature (T_{r})	353.00 K

Table 1. Overview of the input and output variables of the reactor model and their steady state values.

To compare the performance of the different MPC schemes discussed in this paper, the recovery of the system from an initial disturbance in the reactor temperature was investigated. More specificly, initial disturbances of the reactor temperature of respectively 1%, 2%, 3% and 4% were applied and the behaviour of the system was observed.

The following observations can be made. As shown in table 2, standard MPC is feasible for all disturbances, but has a worse control cost, compared to some of the controllers with fixed terminal cost. The controllers with fixed terminal cost, however, impose a trade-off between local optimality and feasibility. This is overcome by the controller with time-varying terminal cost which is feasible for all disturbances and has superior control cost. The different input and output trajectories for the 3% disturbance can be observed in fig. 2.

8. CONCLUSION

In this paper a linear LMI-based MPC scheme using a time-varying terminal cost was introduced and asymptotic stability was proven. Existing LMI-formulations for calculating a stabilizing terminal cost were combined with the classical MPC formulation to obtain a single optimization problem, leading to this scheme. The advantages are improved feasibility and conservation of local optimality, which are illustrated using a model of a continuously stirred copolymerization reactor.

The disadvantage of the proposed scheme is the increase in computational complexity. A further analysis and a new scheme with time-varying terminal cost with reduced computational complexity is discused in Pluymers et al. (2003), as well as a generalization towards nonlinear systems, using linear, time-varying models.

$b \setminus a$	/	0.01	0.02	0.03	0.04	time-varying $\bar{x}$
0.01	5.46	4.83	4.83	5.57	6.71	4.83
0.02	21.84	/	19.32	22.29	26.86	19.31
0.03	49.15	/	/	50.16	60.43	43.78
0.04	87.37	/	/	/	107.32	79.97

Table 2. Control cost of MPC without terminal cost (column 1), MPC with fixed terminal cost (column 2 to 5) for different $\bar{x}$ (with $T_{\mathrm{r,n}} = a$) and MPC with time-varying terminal cost (column 6) for different initial disturbances x_0 (with $T_{\mathrm{r,n}} = b$) on the copolymerization reactor model. A slash (/) means the controller resulted in an infeasible optimization problem. The parameters used were $Q = I$, except $Q_{(7,7)} = Q_{(13,13)} = Q_{(14,14)} = Q_{(15,15)} = 10$, $R = I$ and $P = 3$. Component-wise bounds $[-1, 1]$ were applied to the (normalized) inputs and states.

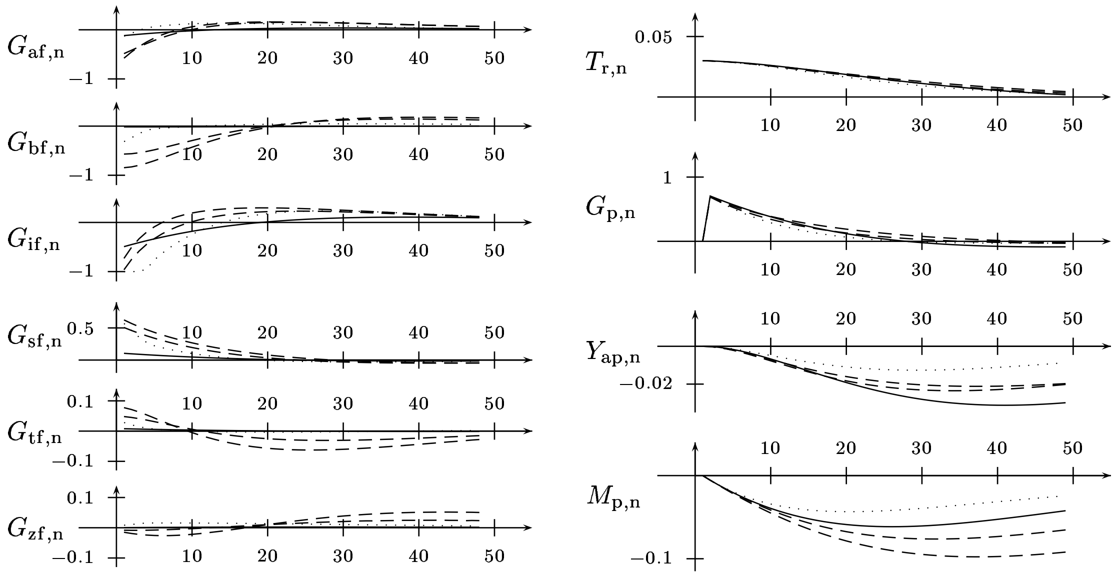

Fig. 2. Input(left) and output sequences (right) for the simulations with MPC without terminal cost (solid), MPC with fixed terminal cost (dashed) and MPC with time-varying terminal cost (dotted) for an initial disturbance of 3% in the reactor temperature. The same parameters as in table 2 were used. The method with time-varying terminal cost clearly performs better than the other methods.

9. ACKNOWLEDGEMENTS

Bert Pluymers is a research assistant with the I.W.T. (Flemish Institute for Scientific and Technological Research in Industry).

Jeroen Buijs is a research assistant with the Katholieke Universiteit Leuven.

Johan Suykens is a postdoctoral researcher with the FWO Flanders and a professor at the Katholieke Universiteit Leuven.

Dr. Bart De Moor is a full professor at the Katholieke Universiteit Leuven, Belgium.

Research supported by Research Council KUL: GOA-Mefisto 666, several PhD/postdoc & fellow grants;

Flemish Government: FWO: PhD/postdoc grants, projects, G.0240.99 (multilinear algebra), G.0407.02 (support vector machines), G.0197.02 (power islands), G.0141.03 (Identification and cryptography), G.0491.03 (control for intensive care glycemia), G.0120.03 (QIT), research communities (ICCoS, ANMMM); AWI: Bil. Int. Collaboration Hungary/ Poland; IWT: PhD Grants, Soft4s (softsensors);

Belgian Federal Government: DWTC (IUAP IV-02 (1996-2001) and IUAP V-22 (2002-2006)), PODO-II (CP/40: TMS and Sustainibility); EU: CAGE; ERNSI; Eureka 2063-IMPACT; Eureka 2419-FliTE;

Contract Research/agreements: Data4s, Electrabel, Elia, LMS, IPCOS, VIB;

REFERENCES

R. R. Bitmead, M. Gevers, and V. Wertz. *Adaptive Optimal Control - The Thinking Man's GPC*. Englewood Cliffs, NJ: Prentice Hall, 1990.

S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM Publications, 1994.

J.P. Congalidis, J.R. Richards, and W.H. Ray. Modeling and control of a copolymerization reactor. *Proceedings of ACC*, 3:1779–1793, 1986.

J.P. Congalidis, J.R. Richards, and W.H. Ray. Feedforward and feedback controler of a solution copolymerization reactor. *AIChE Journal*, 35 (6), 1989.

R.E. Kalman. Contributions to the theory of optimal control. *Bull. Soc. Lath. Mex.*, pages 102–119, 1960.

M.V. Kothare, V. Balakrishnan, and M. Morari. Robust constrained model predictive control using linear matrix inequalities. *Automatica*, 32:1361–1379, 1996.

J.-W. Lee, W. H. Kwon, and J. Choi. On stability of constrained receding horizon control with finite terminal weighting matrix. *Automatica*, 34:1607–1612, 1998.

D.Q. Mayne, J.B. Rawlings, C.V. Rao, and P.O.M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36:789–814, 2000.

B. Pluymers, L. Roobrouck, J. Buijs, J. Suykens, and B. De Moor. An LMI-based constrained MPC-scheme with time-varying terminal cost for nonlinear systems. *SISTA Internal Report 03-53*, 2003.

S.J. Qin and T.A. Badgwell. An overview of industrial model predictive control technology. *AIChE Symposium Series 316*, 93:232–256, 1996.

J. B. Rawlings and K. R. Muske. Stability of constrained receding horizon control. *IEEE Transactions on Automatic Control*, 38:1512–1516, 1993.

COMPUTATIONAL DELAY IN NONLINEAR MODEL PREDICTIVE CONTROL

Rolf Findeisen * **Frank Allgöwer** *

* *Institute for Systems Theory in Engineering, University of Stuttgart, 70550 Stuttgart, Germany,*
{findeise,allgower}@ist.uni-stuttgart.de

Abstract: By now a series of NMPC schemes exist that lead to guaranteed stability of the closed-loop. However, in these schemes the computation time to find a solution of the open-loop optimal control problem is often neglected. In practice the necessary computation time is often not negligible, and leads, since not explicitly considered, to a delay between the state information and the input signal implemented on the system. This delay can lead to a drastic performance decrease or even to instability of the closed-loop. In this paper we outline a simple approach how the computational delay can be considered in nonlinear model predictive control schemes and provide conditions under which the stability of the closed-loop can be guaranteed. This allows to employ nonlinear model predictive control even in the case that the necessary numerical solution time is significant. The presented approach is exemplified considering the control of a continuous reactor.

Keywords: nonlinear predictive control, computational delay, stability

1. INTRODUCTION

In many process control problems it is desired to design a stabilizing feedback such that a performance criterion is minimized while satisfying constraints on the controls and the states. From an optimal control point of view one would ideally like to solve the corresponding Hamilton-Jacobi-Bellmann equations to obtain an explicit solution of the corresponding feedback law. However, often the explicit solution of the corresponding partial differential equations can not be obtained. One way to circumvent this problem is the application of model predictive control(MPC) strategies.

The work presented in this paper is concerned with nonlinear model predictive control (NMPC) for continuous time processes and the problems resulting from the often not negligible on-line computation time. While by now a series of NMPC schemes exist that guarantee closed loop stability (see for example (Mayne *et al.* 2000, Rawlings 2000, Allgöwer *et al.* 1999) for an overview), in these schemes the necessary on-line computation time is typically not taken into account.Even though that recent developments in dynamic optimization have lead to efficient numerical solution methods for the open-loop optimal control problem (See e.g. (Bartlett *et al.* 2000, Findeisen *et al.* 2002, Tenny and Rawlings 2001, Diehl *et al.* 2002)), the solution time is often significant. Neglecting the resulting delay is thus of paramount interrest. Otherwise the performance might degrade significantly or even instability of the closed loop can occur.

One of the few works that take the delay into account is the work presented in (Chen *et al.* 2000). In this paper we outline a similar, rather simple method on how the occuring delay can be taken into account in sampled-data NMPC. In comparison to (Chen *et al.* 2000), the derived results allow to stabilize a wider class of systems and to consider more general cost functions. We furthermore exemplify the importance of the consideration of the delay via a small example system.

The paper is structured as follows: In Section 2 we discuss the difference between the so called sampled-data and the instantaneous approach to NMPC. Section 3 contains a description and the proof of stability for the proposed NMPC approach that takes the delay into account. The properties of this approach are discussed in Section 4. Before we conclude in Section 6 we present in Section 5 a small example considering the control of a simple CSTR.

2. SAMPLED-DATA NMPC

We consider the stabilization of continuous time nonlinear systems described by

$$\dot{x}(t) = f(x(t), u(t)), \quad x(0) = x_0 \tag{1}$$

subject to the input and state constraints

$$u(t) \in \mathcal{U}, \quad x(t) \in \mathcal{X}, \quad \forall t \geq 0, \tag{2}$$

where $x(t) \in \mathcal{X} \subseteq \mathbb{R}^n$ and $u(t) \in \mathcal{U} \subset \mathbb{R}^m$ denote the vector of states and inputs, respectively. The set of feasible inputs is denoted by $\mathcal{U}$ and the set of feasible states is denoted by $\mathcal{X}$. We assume that $\mathcal{U} \subset \mathbb{R}^m$ is compact, $\mathcal{X} \subseteq \mathbb{R}^n$ is connected and $(0,0) \in \mathcal{X} \times \mathcal{U}$. With respect to the vector field $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ we assume that it is locally Lipschitz continuous and satisfies $f(0,0) = 0$.
Model predictive control is based on the repeated solution of an open-loop optimal control problem subject to the system dynamics and the constraints. Based on the system state at time t, the controller predicts the behavior of the system over a prediction time T_p in the future [1] such that an open-loop performance objective functional is minimized. To incorporate feedback that counteracts possible disturbances, the optimal open-loop input is implemented only until the next *recalculation instant*. Based on the new system state information, the whole procedure – prediction and optimization – is repeated, moving the control and prediction horizon forward.
Mathematically the open-loop optimal control problem that is solved at the recalculation instants can be formulated as:

$$\min_{\bar{u}(\cdot)} J(\bar{u}(\cdot), x(t)) \tag{3a}$$

$$\text{s.t. } \dot{\bar{x}} = f(\bar{x}, \bar{u}), \quad \bar{x}(t) = x(t) \tag{3b}$$

$$\bar{u}(\tau) \in \mathcal{U}, \; \bar{x}(\tau) \in \mathcal{X}, \; \tau \in [t, t+T_p], \tag{3c}$$

$$\bar{x}(t+T_p) \in \mathcal{E} \tag{3d}$$

where the cost function J is typically given by

$$J(\cdot) = \int_t^{t+T_p} F(\bar{x}(\tau), \bar{u}(\tau)) d\tau + E(\bar{x}(t+T_p)). \tag{3e}$$

The bar denotes internal controller variables, $\bar{x}(\cdot)$ is the solution of (3b) driven by the input $\bar{u}(\cdot) : [0, T_p] \to \mathcal{U}$ with initial condition $x(t)$. We assume that the "stage cost" $F : \mathcal{X} \times \mathcal{U} \to \mathbb{R}$ is locally Lipschitz continuous with $F(0,0) = 0$ and $F(x,u) > 0 \; \forall \; \mathcal{X} \times \mathcal{U} \ni (x,u) \neq (0,0)$. The end penalty E and the terminal region constraint $\mathcal{E}$ are often used to enforce stability of the closed-loop (Mayne *et al.* 2000, Allgöwer *et al.* 1999, Fontes 2000).
In the following, optimal solutions of the dynamic optimization problem (3) are marked by $(\cdot)^*$. For example we denote the optimal input for $x(t)$ by $u^*(\cdot; x(t)) : [0, T_p] \to \mathcal{U}$.
The input applied to the system in NMPC is based on the optimal input u^*. Depending on how "often" the open-loop optimal control problem (3) is recalculated, different concepts of NMPC exist. If the open-loop is solved at all time instants, we refer to it as *instantaneous NMPC*. If the dynamic optimization is solved only at disjoint recalculation instants and the resulting optimal control signal is implemented open-loop in between, the resulting scheme is called *sampled-data NMPC*.
Instantaneous NMPC: In instantaneous NMPC the input applied to the system is given by

$$u(x(t)) = u^*(t; x(t)), \tag{4}$$

leading to the *nominal closed-loop system*

$$\dot{x}(t) = f(x(t), u(x(t))). \tag{5}$$

Various instantaneous NMPC schemes exist, see for example (Mayne *et al.* 2000). From a practical point of view instantaneous NMPC schemes are not appealing, since an open-loop optimal control problem must be solved at *all times*, which is certainly not possible in practice.

Sampled-data NMPC: In the remainder of the paper we consider sampled-data NMPC. In difference to instantaneous NMPC, in sampled-data NMPC, the open-loop optimal control problem is only solved at the *discrete* recalculation instants and the resulting optimal input signal is applied open-loop to the system until the next recalculation instant. Thus the applied input is given by

$$u(\tau) = u^*(\tau; x(t_i)), \; \tau \in [t_i, t_{i+1}) \tag{6}$$

where t_i denotes the discrete recalculation instants. The *nominal closed-loop system* under the feedback (6) is given by

$$\dot{x}(t) = f(x(t), u^*(t; x(t_i))). \tag{7}$$

For simplicity and clarity we denote the resulting state by $x(\tau; x(t_i), u^*(\cdot; x(t_i)))$, $\tau \in [t_i, t_{i+1})$.

We assume that the recalculation instants t_i are given by a partition π of the time axis.

Definition 1 (Partition) *Every series $\pi = (t_i)$, $i \in \mathbb{N}$ of positive real numbers such that $t_0 = 0$, $t_i < t_{i+1}$ and $t_i \to \infty$ for $i \to \infty$ is called a partition. Furthermore,*

- $\bar{\pi} := \sup_{i \in \mathbb{N}} (t_{i+1} - t_i)$ *is the upper diameter of* π *(longest recalculation time).*
- $\underline{\pi} := \inf_{i \in \mathbb{N}} (t_{i+1} - t_i)$ *is the lower diameter of* π *(shortest recalculation time).* ◇

For a given t, t_i should be taken as the nearest recalculation instant with $t_i < t$. We denote the time between

[1] For simplicity we assume that the prediction and control horizon coincide.

two consecutive recalculation instants t_i and t_{i+1} as recalculation time $\delta_i^r = t_{i+1} - t_i$. Allowing for varying recalculation times, allows to re-optimize the input more frequently if the system dynamics changes rapidly. Sampled-data NMPC schemes leading to stability of the closed-loop are for example given in (Fontes 2000, Michalska and Mayne 1993, Chen and Allgöwer 1998, Jadbabaie *et al.* 2001, de Oliveira Kothare and Morari 2000, Magni and Scattolini 2002, Chen *et al.* 2000, Findeisen *et al.* 2003).

Even so that in sampled-data NMPC in principle the recalculation time $\delta_i^r = t_{i+1} - t_i$ is available for the solution of the open-loop optimal control problem, most of the existing standard NMPC schemes that guarantee stability do not take the necessary solution time for (3) and the resulting delay into account. One of the few exceptions is the work presented in (Chen *et al.* 2000), in which the computational delay is taken into account by optimizing at every recalculation instant, based on a prediction of the state at the next recalculation instant, the open-loop optimal control problem for the next recalculation instant. The purpose of this work is to expand the results in (Chen *et al.* 2000) and to outline rather general conditions that guarantee that the closed loop is stable. Furthermore we underpin by a simple example the importance of a correct consideration of the occuring computational delay.

We assume in the following that the *maximum time for finding the solution* to the open-loop optimal control problem is known and denoted by $\bar{\delta}^c$. Furthermore, we assume that the lower diameter of the recalculation instant partition $\underline{\pi}$ satisfies $\underline{\pi} \geq \bar{\delta}^c$ and that $T_p \geq \bar{\pi}$.

Remark 1 *Note, that we do not necessarily sample and hold the input in between recalculation instants. The reason for this is twofold: First of all the use of a fixed input does not allow to achieve asymptotic convergence to the origin if it is not considered during the optimization without* decreasing the recalculation time to zero. *Secondly, in practice the recalculation time is either predetermined by the time needed to solve the open-loop optimal control problem or by an external scheduling mechanism. It is typically significantly larger than the sampling time of the process control system. As sampling time δ^s we refer to the time the process control system operates, i.e. the A/D and D/A converter operate. Typically, the sampling time is in the order of seconds, whereas the time needed for solving the open-loop optimal control problem (which often also defines the recalculation time) is typically in the order of tenth of seconds, minutes or even tenth of minutes. Thus the open-loop optimal input signal that is applied to the system during $[t_i, t_{i+1}]$ can be sufficiently well approximated by a sample and hold staircase related to the sampling time δ^s of the D/A converters see Figure 1. Since the sampling time is often significantly faster then the recalculation time, the remaining approximation error can be seen as an (small) input disturbance, which NMPC under certain conditions is able to handle (Findeisen* et al. *2003).* ♢

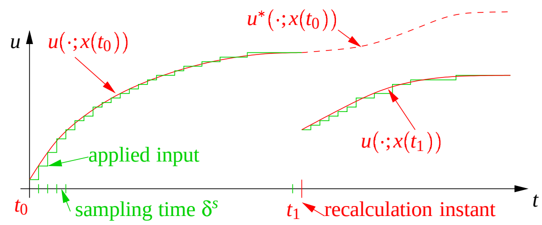

Fig. 1. Recalculation time, sampling time and sample and hold.

The next section outlines a simple approach to consider the necessary solution time in the NMPC problem and gives conditions under which stability of the closed loop can be guaranteed.

3. NMPC AND COMPUTATIONAL DELAY

The approach we propose is based on the idea to continue applying the input from the last recalculation instant t_i also during the (maximum) time $\bar{\delta}^c$ needed for solving the open-loop optimal control problem. In comparison to (6), the open-loop input that is applied to the system is thus given by:

$$u(\tau) = u^*(\tau; x(t_i))\ ,\ \tau \in [t_i + \bar{\delta}^c, t_{i+1} + \bar{\delta}^c).$$

Since the input for the time $[t_i, t_i + \bar{\delta}^c)$ is now given by the previous recalculation, it is not any longer available as degree of freedom in the open-loop optimal control problem (3). Thus problem (3) must be adapted to account for this new situation. In principle one can add the additional constraint

$$\bar{u}(\tau) = u^*(\tau; x(t_{i-1}))\ \tau \in [t_i, t_i + \bar{\delta}^c) \tag{8}$$

to (3) or one can use $u^*(\tau; x(t_{i-1}))\ \tau \in [t_i, t_i + \bar{\delta}^c)$ to predict $x(t_i + \bar{\delta}^c)$ and solve the open-loop optimal control problem for this "initial" state. For simplicity of notation we follow the first approach. For this reason we require additionally that $T_p \geq \bar{\pi} + \bar{\delta}^c$, i.e. the prediction horizon is long enough to at least span to $t_{i+1} + \bar{\delta}^c$. The resulting open-loop optimal control problem that is solved at every recalculation instant t_i is give by

$$\min_{\bar{u}(\cdot)} J(\bar{u}(\cdot), x(t_i)) \tag{9a}$$

$$\text{s.t.}\ \dot{\bar{x}} = f(\bar{x}, \bar{u}),\quad \bar{x}(t) = x(t_i) \tag{9b}$$

$$\bar{u}(\tau) = u^*(\tau; x(t_{i-1})), \tau \in (t_i, t_i + \bar{\delta}^c] \tag{9c}$$

$$\bar{u}(\tau) \in \mathcal{U},\ \bar{x}(\tau) \in \mathcal{X},\ \tau \in [t_i, t_i + T_p], \tag{9d}$$

$$\bar{x}(t_i + T_p) \in \mathcal{E}, \tag{9e}$$

with J given by (3e). Note that the notation $u^*(\tau; x(t_i), t_i, t_i + T_p)$ is not totally correct. The optimal input u^* now also depends on the input at t_{i-1}.

In the following we state a theorem establishing conditions for stability of the closed-loop. The theorem is along the lines of the results in (Fontes 2000, Chen and Allgöwer 1998), which do not consider the computational delay.

Theorem 3.1 (Stability of sampled-data NMPC considering computational delay)
Suppose there exists a set $\mathcal{E}$ and a terminal penalty E such that

(a) $E \in C^1$ *and* $E(0) = 0$,

(b) $\mathcal{E} \subseteq \mathcal{X}$ *is closed and connected with the origin contained in* $\mathcal{E}$,

(c) $\forall x \in \mathcal{E}$ *there exists a input* $u_{\mathcal{E}} : [0, \bar{\pi}] \to \mathcal{U}$ *such that* $x(\tau) \in \mathcal{E}$, $\forall \tau \in [0, \bar{\pi}]$ *and*

$$\frac{\partial E}{\partial x} f(x(\tau), u_{\mathcal{E}}(\tau)) + F(x(\tau), u_{\mathcal{E}}(\tau)) \leq 0 \qquad (10)$$

(d) the NMPC open-loop optimal control problem has a feasible solution for t_0.

Then the state of the nominal closed-loop system defined by (9), (8), *and* (3e) *converges to the origin for all partitions* π *that satisfy* $\underline{\pi} \geq \delta^c$, $T_p \geq \bar{\pi} + \delta^c$. *Furthermore, the region of attraction* $\mathcal{R}$ *is given by the set of states for which the open-loop optimal control problem* (9) *has a solution.*

Note that we achieve stability in the sense of convergence to the origin (=steady state).

Proof.

As usual in predictive control the proof consists of two parts: a feasibility part and a convergence part.

Feasibility: Take any time t_i for which a solution exists (e.g. t_0). After solving the open-loop optimal control problem, the optimal input $u^*(\tau; x(t_i))$ corresponding to $x(t_i)$ is implemented for $\tau \in (t_i + \delta^c, t_{i+1} + \delta^c]$. Since we assume no model plant mismatch and since the open-loop input from the previous recalculation, which is applied during the solution of (9) is taken into account, the predicted open-loop state $\bar{x}(t_{i+1})$ at t_{i+1} coincides with $x(t_{i+1})$. Thus, the remaining piece of the optimal input $u^*(\tau; x(t_i))$, $\tau \in [t_{i+1}, t_i + T_p]$ satisfies the state and input constraints if "applied" to (9b), and $\bar{x}(t_i + T_p; x(t_i), u^*(\tau; x(t_i))) \in \mathcal{E}$. According to Theorem 3.1 (c) $\mathcal{E}$ and E are chosen such that for every $x(t) \in \mathcal{E}$ there exists at least one input $u_{\mathcal{E}}(\cdot)$ that renders $\mathcal{E}$ invariant over $\bar{\pi}$. Consider the following input candidate for t_{i+1},

$$\tilde{u}(\tau) = \begin{cases} u^*(\tau; x(t_i)), \ \tau \in [t_{i+1}, t_i + T_p] \\ u_{\mathcal{E}}(\tau), \tau \in (t_i + T_p, t_{i+1} + T_p] \end{cases} \qquad (11)$$

which is a concatenation of the remaining old input and $u_{\mathcal{E}}(\cdot)$. This input satisfies all constraints and leads to $x(t_{i+1} + T_p; x(t_{i+1}), \tilde{u}(\cdot)) \in \mathcal{E}$. Thus, feasibility at time t_i implies feasibility at t_{i+1}, i.e. if the open-loop optimal control problem has a solution for t_0 it also has a solution afterwards. Furthermore, if one can show that the states for which (9) has a (initial) solution converge to the origin, it is clear that the region of attraction $\mathcal{R}$ consists of the points for which (9) posses a solution. This is established in the next part of the proof.

Convergence: We denote the optimal cost at every recalculation instant t_i as value function $V(x(t_i)) = J^*(u^*(\cdot, x(t_i))$. We show that the value function is strictly decreasing. This allows to establish convergence of the state to the origin. Remember that the value of V at the recalculation instant t_i is given by:

$$V(x(t_i)) = \int_{t_i}^{t_i+T_p} F(\bar{x}(\tau; x(t_i), u^*(\cdot; x(t_i))), u^*(\tau; x(t_i)) d\tau + E(\bar{x}(t_i + T_p; x(t_i), u^*(\cdot; x(t_i))).$$

Consider now the cost resulting from the application of $\tilde{u}$ a starting from $x(t_{i+1})$:

$$J(\tilde{u}(\cdot), x(t_{i+1})) = \int_{t_{i+1}}^{t_{i+1}+T_p} F(\bar{x}(\tau; x(t_{i+1}), \tilde{u}(\cdot)), \tilde{u}(\tau)) d\tau + E(\bar{x}(t_{i+1} + T_p; x(t_{i+1}), \tilde{u}(\cdot))).$$

Reformulating yields

$$\begin{aligned} J(\tilde{u}(\cdot), x(t_{i+1})) &= V(x(t_i)) \\ &- \int_{t_i}^{t_{i+1}} F(\bar{x}(\tau; x(t_i), u^*(\cdot; x(t_i))), u^*(\tau; x(t_i))) d\tau \\ &- E(\bar{x}(t_i + T_p; x(t_i), u^*(\cdot; x(t_i)))) \\ &+ \int_{t_i+T_p}^{t_{i+1}+T_p} F(\bar{x}(\tau; x(t_{i+1}), \tilde{u}(\cdot)), \tilde{u}(\tau)) d\tau \\ &+ E(\bar{x}(t_{i+1} + T_p; x(t_{i+1}), \tilde{u}(\cdot))) \end{aligned}$$

Integrating inequality (10) over $\tau \in [t_i + T_p, t_{i+1} + T_p]$ we can upper bound the last three terms by zero. Thus, we obtain

$$\begin{aligned} &V(x(t_i)) - J(\tilde{u}(\cdot), x(t_{i+1})) \\ &\quad \leq - \int_{t_i}^{t_{i+1}} F(\bar{x}(\tau; x(t_i), u^*(\cdot; x(t_i))), u^*(\tau; x(t_i))) d\tau \end{aligned}$$

Since $\tilde{u}$ is only a feasible, but not the optimal input for $x(t_{i+1})$ it follows that

$$\begin{aligned} &V(x(t_i)) - V(x(t_{i+1})) \\ &\quad \leq - \int_{t_i}^{t_{i+1}} F(\bar{x}(\tau; x(t_i), u^*(\cdot; x(t_i))), u^*(\tau; x(t_i))) d\tau \end{aligned} \qquad (12)$$

This establishes that for any partition with $\underline{\pi} \geq \bar{\delta}^c$ (the time between two recalculations is sufficiently long to allow the solution of the open-loop optimal control problem) and with $T_p \geq \bar{\pi} + \delta^c$ (the prediction horizon spans sufficiently long into the future) the value function is decreasing. Since the decrease in (12) is strictly positive for $(x, u) \neq (0, 0)$ it is possible, similar to (Fontes 2000, Chen and Allgöwer 1998), to employ a variant of Barbalat's lemma to establish that the states converge to the origin for $t \to \infty$. ■

4. DISCUSSION

The conditions for stability in Theorem 3.1 are, similar to the results in (Fontes 2000), rather general. We do not give specific details on how to obtain a suitable terminal region or terminal penalty term, since most NMPC approaches with guaranteed stability and do not take the computational delay into account can be simply adapted. Examples for suitable approaches are the zero terminal constraint approach (Mayne and Michalska 1990), quasi-infinite horizon NMPC (Chen and Allgöwer 1998), control Lyapunov function based approaches (Jadbabaie *et al.* 2001), and the so called simulation approximated infinite horizon NMPC approach (De Nicolao *et al.* 1998).

In comparison to the scheme presented in (Chen *et al.* 2000) that takes the computational delay into account, the outlined approach is applicable to a wider class of systems and does not require to consider a quadratic cost function. Furthermore, the presented conditions even allow to design NMPC controllers that can stabilize systems which can not be stabilized by a feedback that is continuous in the state, compare (Fontes 2000).
The key reason for including the computation time into the open-loop optimal control problem is that if it is neglected, it is strictly not possible to establish stability, as also shown in the example in the next section. Only if the delay due to the numerical solution is sufficiently small, it can be consider it as an disturbance which NMPC under certain conditions is able to handle (Findeisen *et al.* 2003).

5. EXAMPLE

To illustrate the outlined method and the general influence of a neglected computational delay, we consider the control of classical continuous stirred tank reactor (CSTR), for the exothermic, irreversible reaction $A \rightarrow B$ as outlined in. The model under the assumption of constant liquid volume takes the following form (Henson and Seborg 1997):

$$\dot{c}_A = \frac{q}{V}(c_{Af} - c_A) - k_0 e^{\frac{-E}{RT}} c_A$$
$$\dot{T} = \frac{q}{V}(T_f - T) + \frac{-\Delta H}{\rho C_p} k_0 e^{\frac{-E}{RT}} c_A + \frac{UA}{V\rho C_p}(T_c - T),$$

where UA, q, V, c_{Af}, E, RT, ρ, k_0, $-\Delta H$, C_p are constants, c_A is the concentration of substance A, T is the reactor temperature, and T_c is the manipulate variable – the coolant stream. The objective is to stabilize the operating point $T_s = 375K$, $c_{As} = 0.159\text{mol/L}$ via the coolant stream T_c ($T_{cs} = 302.84K$), where T_c is limited to the interval $[220\text{K}, 330\text{K}]$. As NMPC method quasi-infinite horizon NMPC is applied. The terminal penalty term E and the terminal region $\mathcal{E}$ are obtained considering the quadratic "stage cost" $F = x^{\mathrm{T}} \left[\begin{smallmatrix} 4 & 0 \\ 0 & 4 \end{smallmatrix}\right] x + 2(T - T_{cs})^2$, where $x = \left[\begin{smallmatrix} c_a - c_{As} \\ T - T_s \end{smallmatrix}\right]$, using the direct semi-infinite optimization approach as outlined in (Chen and Allgöwer 1998). For simplicity we assume that the recalculation instants are equally apart, i.e. $t_i = i\delta^r$, where $\delta^r = 0.15\text{min}$. Furthermore we assume, that the maximum required solution time $\bar{\delta}^c$ coincides with the recalculation time, i.e. $\bar{\delta}^c = \delta^r$. The prediction horizon horizon is set to $T_p = 3\text{min}$. The open-loop optimal control problem is solved using a direct optimization method (see e.g. (Biegler and Rawlings 1991)) that is implemented in Matlab. For this purpose the input signal is parametrized as piecewise constant with a sampling time that also coincides with the time between the recalculation instants, i.e. $\delta^s = \delta^r$. Figure 2 and 3 show the simulation result for the initial condition $c_A(0) = 0.5\text{mol/L}$ and $T(0) = 350\text{K}$. Shown are

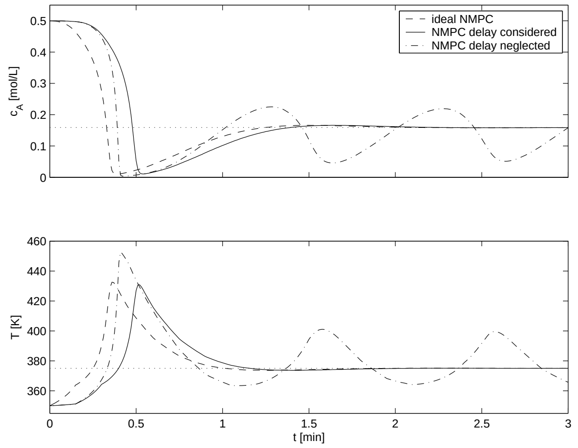

Fig. 2. Resulting states considering an ideal NMPC controller, an NMPC controller that does neglect the delay, and an NMPC controller that accounts for the delay.

the results for an ideal (theoretical) NMPC controller (ideal NMPC), i.e. assuming that the optimal control problem can be solved immediately, an NMPC controller in which the computational delay is not taken into account (NMPC delay neglected), and the scheme outlined in Section 3 (NMPC delay considered). As expected the best performance is achieved for the ideal NMPC controller (which can not be implemented in practice). The more realistic setups, in which a delay occurs, show degraded performance. Clearly it can be seen, that if the delay is not taken into account, that the performance degrades dramatically (curve NMPC delay neglected), i.e. no convergence to the desired steady state is achieved, even so that the delay of 0.15min is rather small. This is also clearly visible in implemented input as shown in Figure 3. Notably,

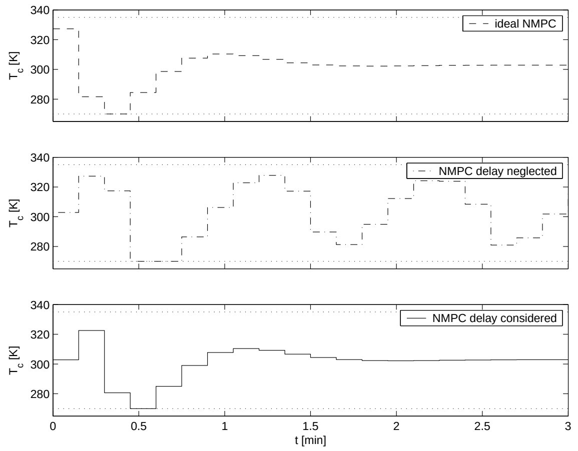

Fig. 3. Resulting input signals.

the NMPC controller that takes the delay into account, achieves very similar performance if compared to the ideal NMPC controller. The remaining difference is mainly due to the initial delay time up to $t = 0.15\text{min}$, in which the old steady state input is applied to the system. Overall it becomes clear, how important the correct consideration of the never avoidable computational delay for stability and good performance.

6. CONCLUSIONS

In this paper we considered the sampled-data NMPC of continuous time systems taking the necessary solution time of the open-loop optimal control problem directly into account. As shown, if the computational delay is not taken into account, the performance of the closed-loop can degrade or even instability can occur. In the approach we outlined, the open-loop input from the previous recalculation is applied until the solution of the optimal control problem is available. Since the "old" input is also taken into account in the open-loop optimal control problem the predicted open-loop trajectory and the closed-loop trajectory coincide in the nominal case. Based on this and suitable assumptions on the terminal region constraint and terminal penalty term, we outlined conditions under which the closed-loop is stable. The assumption on the terminal region constraint and the terminal penalty term are rather general and allow to obtain suitable candidates using various methods such as quasi-infinite horizon NMPC. Overall, the outlined method allows to employ NMPC even in the case that the solution time of the optimal control problem is not negligible.

7. REFERENCES

Allgöwer, F., T.A. Badgwell, J.S. Qin, J.B. Rawlings and S.J. Wright (1999). Nonlinear predictive control and moving horizon estimation – An introductory overview. In: *Advances in Control, Highlights of ECC'99* (P.M. Frank, Ed.). pp. 391–449. Springer.

Bartlett, R.A., A. Wächter and L.T. Biegler (2000). Active set vs. inerior point strategies for model predicitve control. In: *Proc. Amer. Contr. Conf.*. Chicago, Il. pp. 4229–4233.

Biegler, L.T. and J.B Rawlings (1991). Optimization approaches to nonlinear model predictive control. In: *Proc. 4th International Conference on Chemical Process Control - CPC IV* (W.H. Ray and Y. Arkun, Eds.). pp. 543–571. AIChE, CACHE.

Chen, H. and F. Allgöwer (1998). A quasi-infinite horizon nonlinear model predictive control scheme with guaranteed stability. *Automatica* **34**(10), 1205–1218.

Chen, W., D.J. Ballance and J. O'Reilly (2000). Model predictive control of nonlinear systems: Computational delay and stability. *IEE Proceedings, Part D* **147**(4), 387–394.

De Nicolao, G., L. Magni and R. Scattolini (1998). Stabilizing receding-horizon control of nonlinear time varyng systems. *IEEE Trans. Automat. Contr.* **43**(5), 1030–1036.

de Oliveira Kothare, S. and M. Morari (2000). Contractive model predictive control for constrained nonlinear systems. *IEEE Trans. Automat. Contr.* **45**(6), 1053–1071.

Diehl, M., R. Findeisen, S. Schwarzkopf, I. Uslu, F. Allgöwer and J.P. Bock, H.G. Schlöder (2002). An efficient approach for nonlinear model predictive control of large-scale systems. Part I: Description of the methodology. *Automatisierungstechnik* **12**, 557–567.

Findeisen, R., L. Imsland, F. Allgöwer and B.A. Foss (2003). Towards a sampled-data theory for nonlinear model predictive control. In: *New Trends in Nonlinear Dynamics and Control* (W. Kang, C. Borges and M. Xiao, Eds.). Lecture Notes in Control and Information Sciences. Springer-Verlag. New York.

Findeisen, R., M. Diehl, I. Uslu, S. Schwarzkopf, F. Allgöwer, H.G. Bock, J.P. Schlöder and Gilles (2002). Computation and performance assesment of nonlinear model predictive control. In: *Proc. 42th IEEE Conf. Decision Contr.*. Las Vegas, USA.

Fontes, F.A. (2000). A general framework to design stabilizing nonlinear model predictive controllers. *Syst. Contr. Lett.* **42**(2), 127–143.

Henson, M.A. and D.E. Seborg (1997). *Nonlinear Process Control*. Prentice Hall. Upper Saddle River, NJ.

Jadbabaie, A., J. Yu and J. Hauser (2001). Unconstrained receding horizon control of nonlinear systems. *IEEE Trans. Automat. Contr.* **46**(5), 776 –783.

Magni, L. and R. Scattolini (2002). State-feedback MPC with piecewise constant control for continuous-time systems. In: *Proc. 42th IEEE Conf. Decision Contr.*. Las Vegas, USA.

Mayne, D.Q. and H. Michalska (1990). Receding horizon control of nonlinear systems. *IEEE Trans. Automat. Contr.* **35**(7), 814–824.

Mayne, D.Q., J.B. Rawlings, C.V. Rao and P.O.M. Scokaert (2000). Constrained model predictive control: stability and optimality. *Automatica* **26**(6), 789–814.

Michalska, H. and D.Q. Mayne (1993). Robust receding horizon control of constrained nonlinear systems. *IEEE Trans. Automat. Contr.* **AC-38**(11), 1623–1633.

Rawlings, J. B. (2000). Tutorial overview of model predictive control. *IEEE Contr. Syst. Magazine* **20**(3), 38–52.

Tenny, M.J. and J.B. Rawlings (2001). Feasible real-time nonlinear model predictive control. In: *6th International Conference on Chemical Process Control – CPC VI*. AIChE Symposium Series.