15th Triennial World Congress, Barcelona, Spain

ROBUST CONTROL OF NONLINEAR PROCESSES USING MULTIPLE MODELS

Nicolae Constantin, Ion Dumitrache

Automatic Control and System Engineering Department
University "Politehnica" Bucharest
Spl. Independentei, nr.313, 77206, Bucharest, Romania.
Tel./Fax +40-1-4119918, e-mail: nicu@lnx.cib.pub.ro

Abstract: This paper addresses the problem of controlling a nonlinear process when linear models have been identified at different operating points. A multimodel approach is used to to identify and control highly nonlinear process. The internal linear model representation of the nonlinear plant for conventional robust controller is replaced by the linear parameter varying (LPV) model, which introduces transparency while offering distinct advantages for nonlinear model-based control. For this LPV structure a robust controller is then designed based on H∞ techniques. Simulation results for a CSTR process are used to illustrate the performance benefits of multimodel approach.

Keywords: Process control; robust control; multimodel identification;

1. INTRODUCTION

In this paper the control of nonlinear systems subject to multiple operating regimes it is addressed. For most batch and continuous processes in the chemical, biotechnological and power industries, definite regimes can be identified during procedures such as start-up, shut-down and product shifts.

Incorporating simpler models in each operating region can reduce the complexity of the overall model. For example, local state-space and autoregresive moving averaging exogeneous - ARMAX models can be formed using localised perturbation signals and then interpolated to give global non-linear state-space and NARMAX non-linear ARMAX models (Johansen, 1993). The identification of local operating regimes for an unknown plant is difficult. The problem is to identify those variables, which describe the system operating behavior. *A priori* knowledge of the plant can be used at this stage. When little knowledge of the actual regimes exists, however, it may be beneficial to use unsupervised learning methods, such as k-means clustering and nearest neighbors, to give an initial estimate of the interpolation regions.

The main objective of this paper is to design a controller for a nonlinear system that operates in several significantly different modes. This makes it necessary to have a nonlinear model that accurately matches the plant behavior in all operating regimes. The first principles models are usually difficult to

develop if they have to cover a wide range of conditions. The alternative is to identify an empirical model from input-output data. However unmodelled dynamics which are negligible at one operating point may be dominant at another. Therefore it may not be easy to select a model structure that works well in all regimes. In addition, in order to uncover all the necessary plant dynamics, its inputs required by the identification algorithm may not be practically implementable due to their large amplitude and/or large frequency.
In this paper it is pres ented a control structure where multiple local linear models are identified at the different regions of operation, and the robust controller design is carried out using this modes.
The paper is organized as follows. In section 2 the multiple model combination is presented. The controller is designed in section 3. In section 4 the simulation results are presented in section 4 and conclusions in the last section of the paper.

2. COMBINING MULTIPLE LINEAR MODELS

It will be assumed that multiple linear models have been identified to explain plant behavior at different operating points. These local models may have been obtained either through identification, or by linearising a first principles model if one is available. Either continuous or discrete models may be used, though discrete models are most common. The number of models selected is usually related to the number of operating conditions over which the control system is expected to operate. This task must be done in conjunction with the individual controller designs since a larger number of models allows more accurate plant identification and hence the controller based on each model may be more tightly tuned. Using fewer models forces the controller design based on each model to be tuned more robustly.

Model parameters are either determined by linearising a nonlinear model and assuming equilibrium at each operating condition, or , if the plant operates in distinct, well characterized regions, models may be developed from plant data in those operatin g regimes. Input-output models have generally been selected when less is known about the system and a nonlinear model is not available.

Let there be N local linear models (fig.1) with the state space representation :

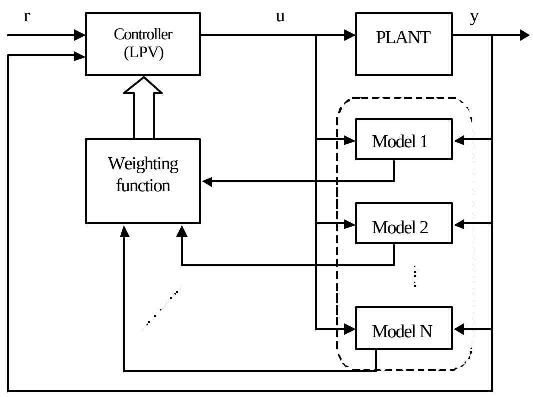

Fig. 1. Basic multiple -model control strategy

$$\begin{aligned} \dot{x} &= A_i x + B_i u \\ y &= C_i x + D_i u \quad , i = 1, \ldots, N \end{aligned} \tag{1}$$

These local model have been obtained either through identification or by linearising a first principles model. The models are combined with the validity functions to obtain a time -varying global model for the plant which will be denoted by M (p, A_i , B_i, C_i, D_i).

A linear parameter varying (LPV) is used as the global model. LPV systems are fixed affine functions of time-varying parameter vectors $q(t)$

$$\begin{aligned} \dot{x} &= A(q(t))\, x + B(q(t))\, u \\ y &= C(q(t))\, x + D(q(t))\, u \end{aligned} \tag{2}$$

The system can be interpreted as LTI systems with time-varying parameters, or as linearisations of a nonlinear system along a trajectory of the parameter θ. Here θ is interpreted to be time -varying model validity function vector p(t).
The nonlinearities of the true plant are captured by the functional dependence of the system state-space matrices on p(t). The nominal global model is given by a map

$$\text{H (A(p) , B(p), C(p), D(p)).} \tag{3}$$

The LPV system should reduce to the i-th linear state space model if p_i = 1. The global model has the following state-space representation:

$$\dot{x} = \left[\sum_i^N p_i(t)A_i\right]x + \left[\sum_i^N p_i(t)B_i\right]u$$
$$y = \left[\sum_i^N p_i(t)C_i\right]x + \left[\sum_i^N p_i(t)D_i\right]u \tag{4}$$

In this paper it is proposed to estimate validity function p_i in an on-line mode, whereas in other approaches (Johansen, 1992) these are obtained off-line, and their dependence on the outputs is therefore fixed a priori.

The reason for combining the local models into a single time-varying global model which is to design a controller for the global model which is parameterised by the validity functions.

The nominal global model is coupled with uncertainty description Δ to represent the true unknown plant. The uncertainty incorporates the effects of

- the errors and uncertainties in the local models - these represent the plant-model mismatches around the points where each of the models was identified,
- the errors and uncertainties in the local nominal model - these represent the plant-model mismatches during transition between different regimes.

The uncertainty in the global model, Δ, which may be large during transitions, reduces to the uncertainty in the i-th local model when the plant is within the domain of the i-th model. The controller is designed with a certain degree of robustness against these uncertainties.

Model validity functions. Probability weighting

Model validity functions are estimates of the validity of each of the local models. These functions make up a vector

$$p(t) = \left[p_1(t), p_2(t), \ldots, p_N(t)\right]^T \tag{5}$$

where

$p_i(t) \to 1$ when the I-th model is valid and
$p_i(t) \to 0$ otherwise,

and

$$\sum_1^N p_i(t) = 1 \tag{6}$$

Equation (6) implies that as the plant moves into a region where one of the models become more trustworthy than the others, the other models lose their validity. The idea of using similar functions to compare models has also been used in (Johansen, 1995).

There area a number of ways of interpreting the model validity function, and this is reflected in the different methods that may be used to assign them on-line. Some possible approaches are:

- Fuzzy logic. Here the model validity functions are interpreted as set membership functions, and are estimated off-line. This is the approach taken in (Zhao, 1995).

- Bayesian estimation. Given the conditions on the model validity functions, one way of interpreting them is as model probabilities, i.e. pi(t) is the probability of the i-th model being valid. The values of these may be estimated on-line from plant measurements using Bayes theorem.

- Simultaneous state and parameter estimation. The validity functions can be treated as parameters of the global model and estimated online using a moving horizon based estimator.

In this paper the Bayesian estimation is considered. Let the plant measurements be denoted yi, and the measurement history by Yi = [y1, y2, . . .]T. Let p (j | Yi) denote the probability that the model i best describes the plant given the measurement history till time ti.

Then applying Bayes theorem

$$\begin{aligned} p(i \mid Y_k) &= p(i \mid y(t_k, Y_{k-1}) = \\ &= \frac{f(y_k \mid i, Y_{k-1})p(i \mid Y_{k-1})}{p(y_k)} = \\ &= \frac{f(y_k \mid i, Y_{k-1})p(i \mid Y_{k-1})}{\sum_{j=1}^N f(y_k \mid j, Y_{k-1})p(j \mid Y_{k-1})} \end{aligned} \tag{7}$$

where f (y_k | i , Y $_{k-1}$) is the probability distribution function of the outputs of the i-th model at time t_k given the measurement history Y $_{k-1}$.

Equation (7) describes how an incoming plant measurement changes the belief about model validity, i.e. how the measurement relates the a posteriori probability to the a priori probability. If the i-th model exactly matches the plant, the model residuals ε_i will be zero-mean, and their covariance will be given by $S_i = C_i P_i C_i + R_i$ where Pi is the state error covariance of the ith Kalman filter, based on the ith model, and, Ri is the covariance of the measurement noise in the i-th regime.

Then assuming stationarity,

$$f(y_k \mid i, Y_{k-1}) = f(y_k \mid i) = f(e_k \mid i) = \\ = \frac{\exp\left\{-\frac{1}{2} e_{ki}(j) S_k^{-1}(j) e_{ki}(j)^T\right\}}{\left[(2p)^N \mid S_k(i) \mid\right]^{1/2}} \quad (8)$$

Therefore equation (8) can be substituted into (7) to obtain an algorithm for estimating model validity.
The probability that the i-th model output represents the plant at time-step k is given by

$$p_{i,k} = \frac{\dfrac{1}{(2p)^{\frac{1}{2}} \mid S_{i,k} \mid^{\frac{1}{2}}} \exp\left(-\frac{1}{2} e_{i,k}^T S_{i,k}^{-1} e_{i,k}^T\right) p_{i,k-1}}{\sum_{j=1}^{N} \left[\dfrac{1}{(2p)^{\frac{1}{2}} \mid S_{j,k} \mid^{\frac{1}{2}}} \exp\left(-\frac{1}{2} e_{i,k}^T S_{i,k}^{-1} e_{i,k}^T\right) p_{j,k-1}\right]} \quad (9)$$

The measurements residuals $e_{k,i}$ (measured outputs at time k minus estimated outputs from i-th model at time step k) are already available from the Kalman filters and the residual covariance matrices S $_{i,k}$ are generated using the covariances matrices, also available from the Kalman filters. Through only the most recent outputs are required, identification is not based solely on that single time sample of information. Past information on the residuals is contained within former probabilities p $_{i,k-1}$; p $_{i,k-2}$; etc.

An important addition to equation (9) is a lower bound preventing $p_{i,k}$ from becoming zero

$$\begin{aligned} P_{i,k} &= p_{i,k} \quad , \quad p_{i,k} > d \\ P_{i,k} &= d \quad , \quad p_{i,k} \le d \end{aligned} \quad (10)$$

This bounding prevent the wind-up phenomenon and limits the number of past observations contained in the current probability estimate. Large values of d yield faster model switching because noncontributing model probabilities are kept artificially high. Low d values require the probabilities in the equation (9) to go through more iterations before i-th controller contributes a relevant portion of the overall control signal.

Inclusion of d affects the probabilities and for this reason very often it find a single model which represents the plant, and may not always blend them well. This means that it shows convergence to a single model rather than a combination of more than one model.

3. CONTROLLER DESIGN

The linear local models are described by the state space matrices: [A_i, B_i, C_i, D_i], i = 1..N. These are then combined with the model validity functions to construct a time-varying global model. The true plant is supposed to lie in the family given by M = M(p) + Δ (t) where M(p) is the LPV model. There are two approaches to robust controller design for such systems:

1. Traditional robust control. Here pi are treated as uncertainties, and a single LTI controller for all regimes is designed .

2. Self-scheduled control. The controller adjusts to variations in plant dynamics by estimating values of p_i (t) on line and using them in the control law to adjust to variations in plant dynamics.

As the control law has to include p(t) the controller is an LPV system denoted by K(p)

$$\begin{aligned} \dot{x}_c &= \left[\sum_i^N p_i(t) A_{k,i}\right] x_c + \left[\sum_i^N p_i(t) B_{k,i}\right] e \\ u &= \left[\sum_i^N p_i(t) C_{k,i}\right] x_c + \left[\sum_i^N p_i(t) D_{k,i}\right] e \end{aligned} \quad (11)$$

where e is the tracking error.

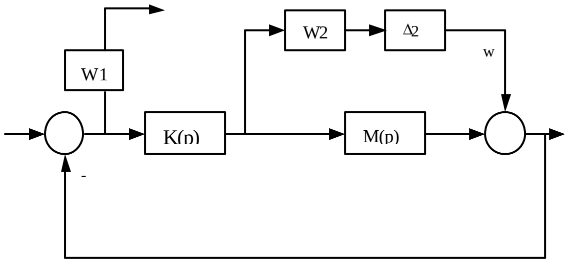

Fig. 2 Feedback loop

The feedback loop is cast into M-Δ form in order to use H_∞. To obtain robust stability and performance for all possible values of the vector p(t).

The trajectory of the model validity function vector p(t) always lies in a polytope whose vertices are

$$[1\ 0\ \ldots\ 0]^T,\ \ldots,\ [0\ 0\ \ldots\ 1]^T \qquad (12)$$

The lower fractional transformation of G(p) and K(p), is also an LPV system whose state-space matrices may be given by

$$[\ A_{cl}(p)\ ,\ B_{cl}(p)\ ,\ C_{cl}(p)\ ,\ D_{cl}(p)\ ] \qquad (13)$$

These matrices evolve in a polytope of matrices whose vertices are:

$$[\ A_{cl,\,i}(p)\ ,\ B_{cl,\,i}(p)\ ,\ C_{cl,\,i}(p)\ ,\ D_{cl,\,i}(p)\ ]$$

and are obtained by substituting (12) in (13).

The closed-loop system is said to have quadratic H_∞ performance if a single quadratic Lyapunov function can be found that establishes global stability. Applying the results of (Apkarian *et. al*., 1994) the Lyapunov function is given by V(x) = xT and it is shown that designing such type controller is equivalent to solving a system of 2N + 1 linear matrix inequalities (LMI).

The controller is LPV and of the form of equation (11). The LPV structure being fixed, the LMIs are solved using LMI-Lab toolbox from Matlab.

4. CASE STUDIES. CSTR EXAMPLE.

In the previous sections a robust multimodel control approach was derived. To demonstrate the applicability and effectiveness of the proposed scheme, in this section we shall implement the proposed scheme to control of a nonlinear first-order exothermic reaction in a continuously stirred tank reactor (CSTR).

The dynamic behavior is described by the following state equations taken from (Ray, 1981)

$$\frac{dx_1}{dt} = -x_1 + D_a(1-x_1)\exp\left(\frac{x_2}{1+x_2/g}\right)$$

$$\frac{dx_2}{dt} = -(1+b)x_2 + BD_a(1-x_1)\exp\left(\frac{x_2}{1+x_2/g}\right) + bu \qquad (14)$$

x is a vector of dimensionless reactant concentration and reactor temperature. The control input u is the dimensionless coolant flow. The physical parameters in the CSTR model equations are Da , γ, B, and β which correspond to Damokhler number, the activated energy, heat of reaction and heat transfer coefficient, respectively. Based on the nominal values of system parameters, Da = 0.072 , γ = 20, B = 8, β = 0.3, the open loop CSTR exhibits three steady states $(x_1,x_2)_A$ = (0.144,0.886), $(x_1,x_2)_B$ = (0.445, 2.750) and $(x_1,x_2)_C$ = (0.765, 4.705) where the upper and lower steady states $(x_1,x_2)_A$ and $(x_1,x_2)_C$ are stable, whereas the middle one, $(x_1,x_2)_B$ is unstable. The plant model is merely used for simulation of the dynamics of the CSTR.

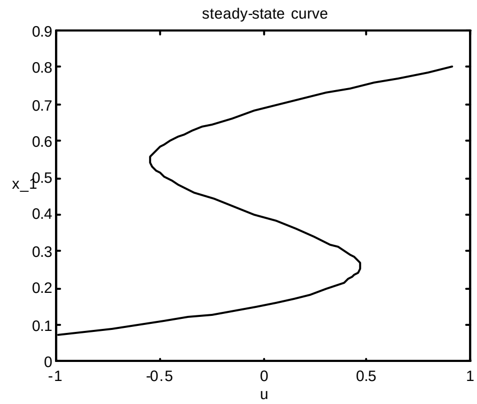

Fig. 3 , Steady state diagram

Since the optimal operation of any chemical reactor is determined by the overall economics of the process, it may be desirable to operate the reactor at several different steady-states. The system is interesting because it exhibits output multiplicity, as can be seen from the steady-state curve given in the figure.

Three linear models of the process are identified around the three steady-state points corresponding to u=0 on the three branches of the steady-state curve (fig. 3). As the middle branch is unstable, so therefore is the second model. The dynamics of the system change significantly depending upon the operating point. The Bayesian approach described in section 2.1. is used to calculate model validity functions, but with a lower bound of 0.05 on the model probabilities to prevent wind-up problems.

Figure 4 shows the performance of the controller when a setpoint trajectory through all three regimes was provided. Fig. 5 shows how the model validity functions varied with time.

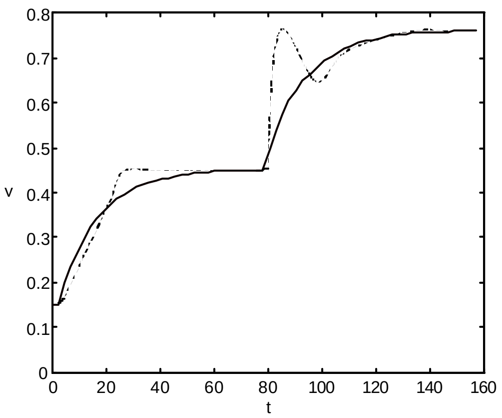

Fig. 4 . Controller performance with three models.

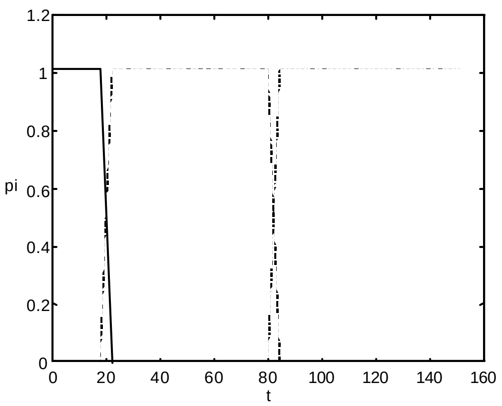

Fig. 5. Model probabilities with three models

5. CONCLUSIONS

In this paper it is presented an approach where different local linear models have been identified at the operating regions. It has been shown that a robust H_∞ controller can be designed for the nonlinear plant based on these local models. This controller is parameterised by time-varying model validity functions that are estimated on-line.

The major advantage of this approach is to bypass most critical aspects associated with gain interpolation or gain scheduling techniques in accounting for the time-varying nature of the plant, and in handling the whole parameter range of the plant in one "shot", that is without extensive simulations.

REFERENCES

Apkarian,P., Gahinet,P , Becker, G., (1994) Self-scheduled H_∞ control of linear parameter-varying systems, Proc. of American Control Conference, pp.3192-3196.

Brown, M.D., Lightbody, G. and Irwin, G.W., Nov. 1997, "Non-linear internal model control using local model networks", *IEE Proc. – Control Theory and Applications*, Vol. 144, No. 6, pp. 505-514.

Constantin, N. (1999). Regularizing Learning in Local Model Networks, Proc. of 12th CSCS Int. Conference, vol. 2, pp.196-201.

Doyle, J.A., A.K.Packard and M.Morari (1989). Robust controller design for a nonlinear CSTR. *Chemical Eng. Science* 44(9), 1929-1947.

Kuipers, B. and K. Astrom (1994) The composition and validation of heterogeneous control laws. *Automatica* 30(2),233-249.

Johansen ,T.A. and B.A. Foss. (1992). Representing and learning unmodelled dynamics with neural networks memories. Proc. of American Control Conference, pp.3037-3043.

Johansen ,T.A. and B.A. Foss (1995). Identification of non-linear system structure and parameters using regime decomposition. *Automatica* 31(2), 233-249.

Ray, W.H. Advances Process Control, McGraw-Hill, New York, 1981.

Zhao, J., V.Wertz and R. Gorez (1995). Design an stabilizing fuzzy controller using LMI method. Proc. of the 3rd European Control Conference, pp.1201-1206.