Proceedings of the
American Control Conference
Seattle, Washington • June 1995



# Robust Model Predictive Control of Nonlinear Systems Using Input-Output Models

Yugender Chikkula and Jay H. Lee [1]

Department of Chemical Engineering, Auburn University, Auburn, AL 36849-5127

Babatunde A. Ogunnaike

E.I. DuPont de Nemours and Company, Wilmington, DE 19880

**Abstract**

In this paper, we present a framework for nonlinear input-output-model-based predictive control from a Bayesian decision theoretic view point. The parameters are modelled as random variables and their probability distributions are computed and used explicitly in the control computation. The framework naturally yields on-line model refinement and "cautious control" against parametric uncertainties. This is important as nonlinear input-output models often contain a large number of parameters and input excitation needed for acceptable parameter estimation is difficult to achieve off-line. The feasibility of the framework is demonstrated by deriving a prototype algorithm for the second order Volterra model. The algorithm is interpreted in the classical Model Predictive Control (MPC) framework and connections to other robust control strategies are discussed.

## 1. Introduction

Most of the chemical processes are inherently nonlinear. Despite this fact, majority of the existing industrial controllers are designed based on linear control techniques. These linear controllers are typically detuned and retuned in order to account for the underlying process nonlinearities, often resulting in conservative control laws that require high maintenance. In some cases the linear control may not even be able to maintain a stable operation. In these situations, a controller which accounts for some of the process nonlinearity can expand the region of operation and improve the closed-loop performance [7, 10].

Despite several important advances made in the nonlinear control system synthesis area, the lack of nonlinear dynamic models has been the major limiting factor for widespread industrial applications of nonlinear control. When a fundamental model is unavailable and prohibitively expensive to develop, empirical modeling offers a practical alternative [11, 6, 15]. Primarily inspired by the linear time-series models, a number of model structures have been suggested that relate the past inputs/outputs to present/future outputs. They range from the relatively simple block-oriented structures to the more general Volterra or NARMAX structures. These models, typically identified from a finite I/O data gathered from a limited experiment, however, have both structural and parametric uncertainties about them and should be used with caution for control purposes.

Structural uncertainty depends on the choice of model structure and operating range within which the model is used. Model structure selection and uncertainty characterization for general nonlinear systems are very difficult issues and solutions do not appear to be within reach any time soon. We believe parametric uncertainty is also an important issue in nonlinear input-output-model-based control since most commonly used model structures contain a large number of parameters and persistent excitation conditions for most of them are either not well understood or difficult to achieve during off-line experiments. For instance, the PRBS signal popular for linear system identification turns out to be a poor choice of input signal for identification of second order Volterra models [14, 12]. We also note that, in the context of developing a model for control, the common practice of reducing the number of parameters based on the information content of the data [1, 16] is an inherently ill-conceived idea, since there is no guarantee that closed-loop input signals generated using the model will lie within the data space used to construct the model.

This paper addresses the issue of *parametric uncertainty* in the setting of nonlinear input-output model based control. We adopt a Bayesian decision theoretic approach and propose to formulate and solve a stochastic optimal con-

[1] Author to whom all correspondence should be addressed: phone (334)844-2060, fax (334)844-2063, e-mail JHL@ENG.AUBURN.EDU

trol problem. In the proposed framework, model parameters are modelled as random variables with given prior distributions. Then their posterior density is computed recursively and used in the control computation. An open-loop optimal control problem is solved at each time step and the solution is implemented in a receding horizon fashion, resulting in a feedback control strategy called "Open-Loop Optimal Feedback Control (OLOFC)." The framework yields a control algorithm that is robust with respect to parametric uncertainty, with features like on-line parameter refinement (and/or adaptation) and "cautious control." A second order Volterra model is used to demonstrate the feasibility of the concept. The prototype algorithm consists of two components: (1) a Kalman filter used for calculation of the posterior densitiy of the parameter vector and (2) an on-line optimizer similar to that appearing in the convetional Model Predictive Control (*i.e.*, MPC based on the *perfect* parameter assumption) but with additional time-varying input penalty terms accounting for the parametric uncertainties. The implications for closed-loop robustness are illustrated by a numerical example.

## 2. Control of Uncertain Systems

Design of feedback systems in presence of parametric uncertainty has been widely investigated over the last few decades. Depending on the uncertainty descriptions considered, proposed methods can be broadly classified into two categories: the *hard-bound* approach and the *soft-bound* approach.

In the hard-bound approach, *deterministic* bounds (known as set-membership description of uncertainty) are assigned to the parameters. Proposed control formulations for this uncertainty description include: 1) Design of a fixed parameter controller with constraints given by robust stability/performance conditions (*e.g.*, those from the Kharitonov's theorem) [3, 4]; 2) On-line optimization of which objective is to minimize the predicted "worst-case" error, *i.e.* the largest possible error over the specified set of models [5, 19]. The drawbacks of the former type methods are that the optimization to be solved is often nonconvex and the conditions for robust stability and performance (that can be incorporated into the optimization) are often very conservative. In addition, these methods lack flexibility and are difficult to generalize (to cases involving input constraints, nonlinear models, *etc.*). The latter type methods, on the other hand, require solving a computationally demanding min-max optimization problem on-line. Moreover, the worst-case minimization often results in a very conservative control.

In the soft-bound or *stochastic* approach, the parameters are modelled to be random variables and the optimal control problem is formulated as minimization of the expectation of the future output error variance. The *dual* control [8] and various sub-optimal dual control strategies such as the OLOFC [17] are examples of this approach. These approaches can potentially lead to algorithms that are computationally less demanding and less conservative compared to the deterministic approach.

With a few exceptions [9, 13, 18], most of the research effort in this area has been directed at feedback design for linear uncertain systems. On the other hand, we believe that parametric uncertainty is a much more serious and relevant issue in the nonlinear control context. In this paper, we examine the viability of the soft-bound approach for *nonlinear* input-output-model-based control. The specific model structure that we choose to demonstrate the concept is the second order Volterra model. The recent work of Genceli and Nikalaou[9] also deals with parametric uncertainty of a second order Volterra model in the nonlinear MPC setting. They provide sufficient conditions for robust stability of nonlinear MPC in the face of a particular parametric uncertainty description. The uncertainty description they consider, *independent* hard bounds on each of the parameters (the probabilistic equivalent of which would be a mutually independent truncated uniform distribution for each of the parameters), however, can potentially result in a very conservative closed-loop. The algorithm proposed in this paper takes advantage of the correlation among the uncertainties in different parameters and addresses the performance issue directly through minimization of the "expected error." Although it does not come with any guaranteed stability margin, it is conceivable that some constraints can be added to the optimization to guarantee closed-loop stability for all parameter combinations within a chosen probability level.

## 3. Prototypical Algorithm for Second Order Volterra Model

### Model

Consider the following second order Volterra model structure, which is a natural extension of linear convolution models to nonlinear systems [7]:

$$y_k = \sum_{i=1}^{N} a_{i,k} u_{k-i} + \sum_{i=1}^{N} \sum_{j=1}^{N} b_{i,j,k} u_{k-i} u_{k-j} + d_k$$

or equivalently

$$y_k = \left[\mathcal{U}_{k-1}^T \; 1\right] \theta_k \tag{1}$$

where

$$\begin{aligned} \theta_k^T &= [a_{1,k} \; \ldots \; a_{N,k} \; b_{1,1,k} \;\; b_{1,2,k} \; \ldots \; b_{N,N,k} \; d_k] \\ \mathcal{U}_k{}^T &= [u_k \;\; u_{k-1} \;\; \ldots \;\; u_{k-N+1} \\ & \qquad u_k^2 \;\; u_k u_{k-1} \;\; \ldots \;\; u_{k-N+1}^2] \end{aligned}$$

$d_k$ represents the disturbance signal entering the output.

### Parameter Estimation

We consider the following state-space form of the model structure:

$$\begin{aligned} \theta_{k+1} &= \Phi\theta_k + \nu_k \\ \hat{y}_k &= \Xi_k\theta_k + w_k \end{aligned} \tag{2}$$

In the above equation, $\Xi_k$ denotes the regressor vector $[\mathcal{U}_{k-1}^T\ 1]$ and $w_k$ represents the measurement noise. The statistics of the vector $\theta$ (and its variation) can be shaped arbitrarily by choosing the transition matrix $\Phi$ and statistics of $\nu_k$ appropriately. $\nu_k$ can be described most generally as an output of a stochastic difference equation. In addition to $\nu_k$, the noise $w_k$ can also be chosen as a general random process by augmenting the state vector.

For simplicity of presentation, however, we will assume in this paper that both external signals are independent, identically distributed gaussian random vectors, *i.e.*,

$$\omega_k \sim N(0, Q^w); \quad \nu_k \sim N(0, Q^\nu)$$

Furthermore, we will assume that $\Phi$ is chosen as an identity matrix, implying a random-walk type parameter and disturbance variation. Note that, because the second order Volterra model contains a large number of parameters, one cannot choose a very large value of $Q^\nu$ (which leads to a high degree of "forgetting"), as this would lead to a very ill-conditioned estimation and consequently little control (*i.e.*, too high level of uncertainty causing a phenomenon called "controller turn-off"). If all the elements of $Q^\nu$ corresponding to the system parameters are chosen to be zero, the estimation will lead to on-line parameter refinement, but will not be able to follow parameter changes.

The on-line computation of the conditional mean and covariance of $\theta_k$ can then be carried out by the following Kalman filter:

$$\begin{aligned} \theta_{k|k} &= \theta_{k|k-1} + L_k\epsilon_k \\ L_k &= P_{k|k-1}\Xi_k^T\{\Xi_k P_{k|k-1}\Xi_k^T + Q^w\}^{-1} \\ \theta_{k|k-1} &= \theta_{k-1|k-1} \\ P_{k|k-1} &= P_{k-1|k-1} + Q^\nu \\ \epsilon_k &= \hat{y}_k - \Xi_k\theta_{k|k-1} \\ P_{k|k} &= P_{k|k-1} - P_{k|k-1}\Xi_k^T \\ &\quad \{\Xi_k P_{k|k-1}\Xi_k^T + Q^w\}^{-1}\Xi_k P_{k|k-1} \end{aligned}$$

Here the pair $(\theta_{k|k}, P_{k|k})$ defines the posterior density of $\theta_k$ and gives the best estimate and the associated uncertainty of the parameter vector at time $k$. Along with the deterministic state vector $\mathcal{U}_{k-1}$, it constitutes the information state vector at time $k$ (denoted as $\mathcal{I}_k$ hereafter).

## Control

The idea of considering the parameter uncertainty in control evaluation can be traced back to the concept of *dual control*, where the interdependence between future parameter estimates (plus their covariance) and future inputs was realized [8]. Rigorous consideration of this interplay between the controller and the estimator in the stochastic optimal control setting leads to the following dynamic program:

$$\begin{aligned} V_{k+\ell-1,k+p} = \min_{u_{k+\ell-1}} \; & E\{e_{k+\ell}^T Q_\ell e_{k+\ell} \\ & + u_{k+\ell-1}^T W_u u_{k+\ell-1} + V_{k+\ell,k+p} \\ & \mid \mathcal{I}_{k+\ell-1}\} \end{aligned} \tag{3}$$

where $V$ is the cost-to-go, $\mathcal{I}$ is the information state and $e(\triangleq y_{k+\ell} - r_{k+\ell})$ is the traking error. (3) is a dynamic program, that must be solved sequentially starting form $\ell = p$ to $\ell = 1$. Unlike the linear quadratic gaussian (LQG) problem, an analytical solution of the above dynamic program is not possible because of the nonconvex dependence of $E\{V_{k+\ell,k+p} \mid \mathcal{I}_{k+\ell-1}\}$ on $u_{k+\ell-1}$. Note that, in order to derive an optimal *feedback* control law, the future inputs must be treated as stochastic variables since they depend on the future outputs (yet to be measured). This dependence, however, is very complex and cannot be evaluated analytically. The numerical solution is not a practical option either, due to the computational complexity involved and the "curse of dimensionality."

The approach taken in Model Predictive Control is to treat the future inputs as deterministic variables. When the parameters are exact, the two approaches yield identical results. However, in the case of stochastic parameters, the two approaches produce very different results, the latter approach giving the following tractable *single-stage* optimization:

$$\min_{u_k,\cdots,u_{k+m-1}} E\{\sum_{l=1}^{p}\left(y_{k+\ell|k} - r_{k+\ell}\right)^2 \mid \mathcal{I}_k\} \tag{4}$$

Now expand each term in the open-loop objective given in (4) as follows:

$$\begin{aligned} & E\ \left[(y_{k+\ell} - r_{k+\ell})^2 \mid \mathcal{I}_k\right] \\ = \; & E\{\left[\left(y_{k+\ell|k} - r_{k+\ell}\right) + \left(y_{k+\ell} - y_{k+\ell|k}\right)\right]^2 \mid \mathcal{I}_k\} \\ = \; & \left[y_{k+\ell|k} - r_{k+\ell}\right]^2 + \Xi_{k+\ell}^T P_{k+\ell|k}\Xi_{k+\ell} \end{aligned}$$

Note that only $u_k$ among the computed control moves is implemented and the whole optimization (with the window shifted by one time step) is repeated at the next time step, giving rise to a suboptimal feedback control strategy called "Open-Loop Optimal Feedback Control (OLOFC)."

Thus the optimal open-loop control problem to be solved at each time step is the following constrained NLP:

$$\begin{aligned} \min_{u_k,\cdots,u_{k+m-1}} \sum_{l=1}^{p} \Big[ & \left(y_{k+\ell|k} - r_{k+\ell}\right)^2 + \\ & \Xi_{k+\ell}^T P_{k+\ell|k}\Xi_{k+\ell}\Big] \end{aligned} \tag{5}$$

with

$$\begin{aligned} y_{k+\ell|k} &= \Xi_{k+\ell}^T\theta_{k|k} \\ P_{k+\ell|k} &= P_{k|k} + \ell \cdot Q^\nu \end{aligned}$$

Note that the above is a minimization of a multi-step objective, similar to that found in the classical MPC formulation. The difference, however, is that various combinations of input terms contained in $\Xi_{k+\ell}$ are now penalized in proportion to the corresponding parametric uncertainty characterized by the matrix $P_{k+\ell|k}$. The changes in the model quality are reflected as changes in the covariance matrix $P_{k|k}$, which is recomputed at each time step using the Kalman filter. In the context of conventional MPC, it has been observed that the input weighting has a significant effect on the closed-loop robustness and a constant diagonal input weighting matrix is not always the best choice. However, there has been no systematic way of choosing a non-diagonal input weighting matrix, which varies according to changing uncertainty levels. The proposed formulation results in an automatic, time-varying and directional tuning of input weights, based on the conditional covariance of the parameter vector computed under the assumed stochastic parameter model.

In adaptive control context, the formulation represents a way to relax the "certainty equivalence"(CE) assumption. In CE control, the estimated parameters are treated as though they are the true system parameters, *i.e.*, the uncertainties in the estimates are not considered. This assumption is reasonable in cases where the number of parameters is small and the parameter variation is slow compared to the system states. However, it is not justifiable during the initial phases of the adaptation (which are expected to be significant because of the large number of parameters in typical nonlinear input output models) when the parameter variances are large or when rapid and sudden process changes occur requiring reidentification.

The overall product is a *cautious* controller with a "multi-step" horizon, which can be viewed as an extension of the classical "one-step" cautious control [2] to multi-step predictive control. One well-documented problem with cautious control is the controller *turn-off* caused by a lack of persistent input excitation (needed for good parameter estimation). In the case that the parameter variations are expected to be significant, one would have to combine the proposed algorithm with some probing (active learning) mechanism that maintains the parameter covariance at a reasonable level. The turn-off should not be a serious concern when one starts with a reasonable off-line estimate of the parameter vector and the parameters are assumed to be constant or slowly varying.

## 4. Numerical Example

We consider the following numerical example to illustrate the proposed method.

### 4.1. Example

Consider a SISO second order Volterra system and a second order Volterra model identified from an input-output experiment. The input employed in the identification was a white noise of very small variance which essentially excites the linear dynamics. As can be seen from Figure 1, the estimates for the first-order terms (the first 10 terms) are close to the actual parameters, but the estimates for the second order parameters are substantially off for obvious reasons. This model was implemented in closed-loop using both OLOFC and the CE control strategy. Notice that the CE control strategy can be interpreted as a special case of OLOFC with $P_{k+\ell|k}$ set to a zero matrix in (5).

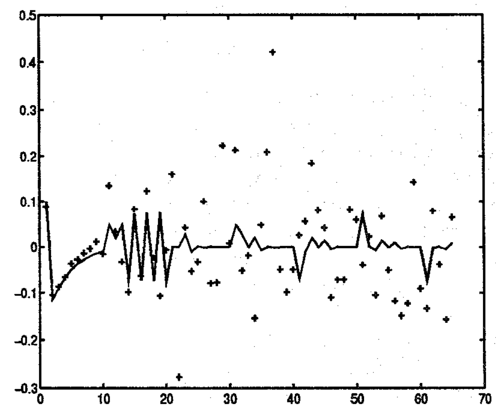


**Figure 1:** Actual (-) vs identified (+) parameters for the second order Volterra system

In order to make the comparison more transparent, we consider a simple scenario. We set $Q^w = 0$ (no measurement noise) and $Q^v = 0$ (constant parameter case). The Kalman filter was initialized with the off-line least-squares estimates of the parameters and the corresponding covariance matrix. The following controller settings were used.

$$p = 17; \quad m = 10$$

The CE controller and the OLOFC are simulated and the results are plotted in Figure 2.

As can be observed, the OLOFC is "cautious" compared to the CE controller, especially as the system moves away from the initial region of identification. The detuning of a CE controller, such as classical MPC, is typically carried out by weighting the manipulated inputs. However, the detuning effects are not transparent and it is not straightforward to detune for the changing level of uncertainties. In OLOFC, however, the in-built and automatic detuning mechanism takes care of this issue and results in a cautious controller.

The OLOFC can be overly cautious depending on the level of uncertainties. In these situations, the objective (5) can be modified as below which gives a relatively more ag-

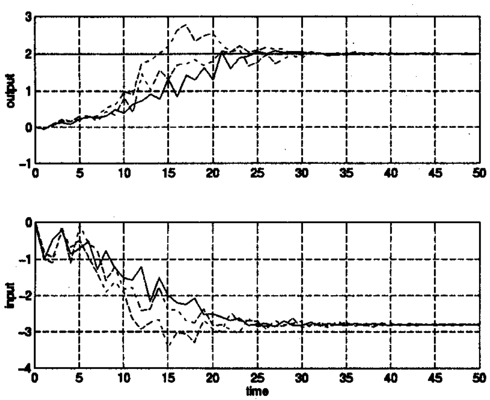


**Figure 2:** Comparison of OLOFC (solid line), modified OLOFC (dashed line) and CE controller (dotdash line) for a second order Volterra system.

gressive controller depending on the choice for $\lambda$.

$$\min_{u_k,\cdots,u_{k+m-1}} \sum_{l=1}^{p} \left[ \left( y_{k+\ell|k} - r_{k+\ell} \right)^2 + \lambda \Xi_{k+\ell}^T P_{k+\ell|k} \Xi_{k+\ell} \right]$$

The above objective interpolates between CE and OLOFC for $0 \leq \lambda \leq 1$. For $\lambda = 0$ we have CE control and for $\lambda = 1$ it results in OLOFC. By choosing an appropriate value for $\lambda$, one can reduce the conservativeness of OLOFC or improve the level of robustness of CE type controllers. The above example for $\lambda = 0.1$ is solved and the closed-loop response is shown in Figure 2. Notice that the level of performance of this *modified* OLOFC is intermediate to OLOFC and CE controller.

**Acknowledgements:**

Financial support from National Science Foundation (CST-9209808) is gratefully acknowledged.

**References**

[1] Akaike H., "Information theory and an extension of the maximum likelihood principle," *Proc. $2^{nd}$ International Symposium on Information Theory*, pp. 267, 1972.

[2] Åström K.J. and B. Wittenmark, *Adaptive control*, Addison Wesley, 1989.

[3] Barmish B.R., C.V. Hollot, F.J. Kraus and R. Tempo, "Extreme point results for robust stabilization of interval plants with first order compensators," *IEEE Transactions on Automatic Control*, AC-37, pp. 707-714, 1992.

[4] Bernstein D.S. and W.M. Haddad, "Robust controller synthesis using Kharitonov's theorem," *IEEE Transactions on Automatic Control*, AC-37, pp. 129-132, 1992.

[5] Campo P.J. and M. Morari, "Robust model predictive control," *Proc. Automatic Control Conference*, Atlanta, GA, pp.1021-1026, 1987.

[6] Diaz H. and A.A. Desrochers, "Modeling of nonlinear discrete-time systems from input-output data, " *Automatica*, 24, pp. 629-641, 1988.

[7] Doyle III F.J., B.A. Ogunnaike and R.K. Pearson, "Nonlinear model based control using second order Volterra models," *To appear in Automatica*, 1994.

[8] Feldbaum A.A., "Optimal control theory," Academic Press, New York, 1965.

[9] Genceli H. and M. Nikalaou, "Design of robust constrained model predictive controllers with Volterra series," *AIChE Annual Meeting*, San Fransisco, CA, 1994.

[10] Hernandez, E., "Control of nonlinear systems using input-output information," Ph.D. Thesis, Georgia Institute of Technology, Atlanta, GA, 1992.

[11] Leontaritis I.J. and S.A. Billings, "Input-Output parametric models for nonlinear systems: parts I & II," *International Journal of Control*, 41, pp. 303-344, 1985.

[12] McCullough G., B.R. Maner, F.J. Doyle III and R.K. Pearson, "Identification of second order Volterra models from plant data for use in model predictive control," *AIChE Annual Meeting*, San Fransisco, CA, 1994.

[13] Michalska H. and D.Q. Mayne, "Robust receding horizon control of constrained nonlinear systems," *IEEE Transactions on Automatic Control*, AC-35, pp. 1623-1633, 1993.

[14] Pearson R.K., B.A. Ogunnaike and F.J. Doyle III, "Identification of nonlinear input/output models using non-gaussian input sequences," *Proc. American Control Conference*, San Fransisco, CA, 1993.

[15] Pearson R.K., "Nonlinear input-output modeling," *Preprints ADCHEM'94*, Kyoto, Japan, 1994.

[16] Snee R.D., "Validation of regression models: Methods and examples," *Technometrics*, 19, pp. 415, 1977.

[17] Tse E. and M. Athans, "Adaptive stochastic control for a class of linear systems," *IEEE Transactions on Automatic Control*, AC-17, pp. 38-51, 1972.

[18] Tse E., Y. Barsholam and L. Mier III, "Wide-sense adaptive dual control of nonlinear stochastic systems," *IEEE Transactions on Automatic Control*, AC-18, pp. 98-108, 1973.

[19] Veres S.M. and J.P. Norton, "Predictive self-tuning control by parameter bounding and worst-case design," *Automatica*, 29, pp. 911-928, 1993.