Nonlinear Model Predictive Control Utilizing a Neuro-Fuzzy Predictor

Waller, Jonas B.
Process Control Lab, Abo Akademi University | Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University: Visiting Researcher

Hu, Jinglu
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

Hirasawa, Kotaro
Department of Electrical and Electronic Systems Engineering, Faculty of Information Science and Electrical Engineering, Kyushu University

https://doi.org/10.15017/1515714
Nonlinear Model Predictive Control Utilizing a Neuro-Fuzzy Predictor

Jonas B. Waller* , Jinglu HU** and Kotaro HIRASAWA**
(Received December 15, 2000)

Abstract: This paper applies a quasi-ARMAX modeling technique, recently presented in the literature, to a process control framework. The use of this quasi-ARMAX modeling technique in nonlinear model predictive control (NMPC) formulations applied to simple nonlinear process control examples is investigated. The quasi-ARMAX predictor can be interpreted as a neuro-fuzzy predictor, and this neuro-fuzzy predictor is computationally straightforward and has showed excellent prediction capabilities. The predictor is thus well suited for NMPC purposes. Furthermore, the parameters of the neuro-fuzzy model can be argued to have explicit meaning, thus making the procedure of tuning the NMPC system more transparent when using the neuro-fuzzy predictor.

Keywords: Model predictive control, Nonlinear control, Neuro fuzzy models, Nonlinear identification

1. Introduction

Model predictive control (MPC) has received a strong position when it comes to industrially implemented advanced control methodologies, especially in refinery and petrochemical fields\(^1\)\(^2\). One main reason for this is the intuitive way MPC incorporates the process model in the controller design, thus e.g. making it straightforward to take constraints into account at the design stage.

In many problems relevant in the process control field today, the plant under control shows a strongly non-linear behavior. As a means to handle this, non-linear MPC (NMPC) is an often used method. NMPC, simply put, is model predictive control, where a non-linear process model is used for prediction purposes, as opposed to a linear model for basic MPC. See e.g. Camacho and Bordons\(^3\) for an overview of MPC, and Henson\(^4\) and Rawlings et al.\(^5\) for a focus on NMPC.

One main difficulty with non-linear MPC is that the non-linear models needed often are complex and give rise to computationally burdensome optimization problems. Since MPC requires the solution of an optimization problem on-line in every sampling instant, the computational simplicity is crucial.

This paper lists several non-linear process models used in recent literature for NMPC purposes, useful as an indicator of the research activity in the field. A new contribution of this paper is adapting a quasi-ARMAX modeling technique, recently presented in the literature\(^6\)\(^7\), for non-linear MPC purposes. The idea of the quasi-ARMAX modeling scheme is to incorporate a group of certain non-linear non-parametric models (NNMs) into a linear ARMAX structure. The investigated predictor can thus be seen as a linear predictor network consisting of a global linear predictor and several local linear predictors with interpolation. The quasi-ARMAX predictor can in such a way be interpreted as a neuro-fuzzy predictor\(^7\). The neuro-fuzzy predictor is computationally straightforward and has showed excellent prediction capabilities. In Hu et al.\(^(1998)\)\(^6\) and Hu et al.\(^\text{(1999)}\)\(^7\), the predictor is presented mainly from the perspective of system identification. In this paper, we apply the predictor in a nonlinear process control framework. More specifically, we illustrate how the predictor can be used in a NMPC concept, and illustrate preliminary control results.

Furthermore, non-linear MPC control strategies are often difficult to interpret and understand, thus making the choice of tuning-parameters a quite ad hoc procedure. The parameters of the neuro-fuzzy model can be argued to have explicit meaning\(^7\), thus making the procedure of tuning the NMPC system more transparent when using the neuro-fuzzy predictor.

2. Nonlinear Model Predictive Control

Nonlinear model predictive control is an open-loop optimal control sequence calculation where the main characteristics of the controller are\(^3\):

- Explicit use of a nonlinear model to predict the
process output at future time instants (a horizon).
- Calculation of a control sequence minimizing an open-loop objective function.
- Receding strategy so that at each instant the horizon is displaced towards the future, which involves the application of the first control signal of the sequence calculated at each step.

The problem formulation for non-linear MPC can briefly be summarized according to the following. See e.g. Henson\(^4\) for details.

The process model is assumed to have the form of a non-linear discrete-time input-output mapping, e.g.

\[
y(k) = g(\varphi(k)) \tag{1}
\]

\[
\varphi(k) = [y(k-1), \ldots, y(k-n), u(k-d), \ldots, u(k-m-d+1)]^T
\]

The open-loop optimization problem which is solved in every sampling instant can be formulated as

\[
\min_{u(k), u(k+1), \ldots, u(k+N_u-1)} J
\]

where

\[
J = \phi[y(k+N_y)], + \sum_{j=0}^{N_y-1} L[y(k+j|k), u(k+j|k), \Delta u(k+j|k)] \tag{2}
\]

Where \(N_u\) is used to denote the control horizon and \(N_y\) denotes the prediction horizon.

We thus have a generally expressed loss function consisting of (quite arbitrary) functions of the predicted outputs and future control actions as well as the successive change in the inputs. Often the loss function is specified as a quadratic loss function such that

\[
\phi = [y(k+N_y) - y_s(k)]^T Q [y(k+N_y) - y_s(k)] \tag{3}
\]

\[
L = [y(k+j|k) - y_s(k)]^T Q [y(k+j|k) - y_s(k)] + [u(k+j|k) - u_s(k)]^T R [u(k+j|k) - u_s(k)] + [\Delta u(k+j|k)]^T S [\Delta u(k+j|k)] \tag{4}
\]

Where \(y_s(k)\) and \(u_s(k)\) are steady-state targets for \(y\) and \(u\), and \(Q\), \(R\) and \(S\) are weighting matrices. The predicted outputs are obtained from the non-linear process model, equation (1).

The capability to handle constraints is considered one of the main motivations to use (nonlinear) MPC. The non-linear optimization problem formulated in equation (2) is in the case of constraints present solved subject to the input inequality constraints

\[
u_{\text{min}} \leq u(k+j|k) \leq u_{\text{max}}, \quad 0 \leq j \leq N_u - 1 \tag{5}
\]

\[
\Delta u_{\text{min}} \leq \Delta u(k+j|k) \leq \Delta u_{\text{max}}, \quad 0 \leq j \leq N_u - 1
\]

and the output inequality constraints

\[
y_{\text{min}} \leq y(k+j|k) \leq y_{\text{max}}, \quad 1 \leq j \leq N_y \tag{6}
\]

For the case with a linear time-invariant process model, a quadratic loss function and in the absence of constraints, an analytical solution to the optimization problem can be found.

If there are constraints present, and the constraints are convex, the optimization problem is convex and can easily be solved\(^5\). When the model is non-linear and non-convex constraints are present, a non-convex optimization problem must be solved iteratively at every sampling time. This is the most common case for non-linear MPC.

If no analytical solution can be found, a numerical optimization method has to be used. One could address the solution of nonlinear optimization problems for the purpose of model predictive control for example through successive linearization of model equations, sequential model solution and optimization, and simultaneous model solution and optimization.

Successive linearization of model equations corresponds to the use of the current operating point to linearize the (non-linear) model before each calculation of the control action of the NMPC controller. Different approaches can be taken in an attempt to reduce the error caused by this linearization. For example, in quadratic dynamic matrix control (QDMC) the process model is linearized once for the entire trajectory and in non-linear quadratic matrix control (NLQDMC), the process model is linearized iteratively in each control interval\(^6\). Sequential model solution and optimization in turn use a standard NLP solver to compute the manipulated inputs, and an ODE solver is used to integrate the non-linear model equations. This is performed iteratively until sufficient accuracy is achieved. Simultaneous model solution and optimization requires that the process model equations are discretized as orthogonal collocation of finite elements\(^7\), in order to solve the nonlinear program. Successive (or
sequential) quadratic programming (SQP) is considered the method of choice for solving non-linear programs. A good treatment of relevant optimization methods is offered in e.g. Gopal and Biegler\textsuperscript{11}.

3. Quasi-Armax Modeling

A key issue to address when designing a NMPC controller is the choice of process model, i.e. the structure of the (non-linear) process model to be used. Here we can roughly divide the area into two (or three) kinds of models. Firstly there are models based on fundamental relationships and secondly there are models based on empirical data. The third kind of model would be one combining fundamental and empirical modeling techniques (which usually is referred to as a hybrid model).

With main focus on empirical non-linear models, one decision to make is the selection of suitable model form. The types of discrete-time non-linear models utilized for NMPC in the recent literature include e.g.

- Hammerstein and Wiener models\textsuperscript{12,14}
- Volterra models\textsuperscript{13,15}
- Polynomial ARMAX\textsuperscript{15}
- Artificial neural networks\textsuperscript{16,9}
- Fuzzy logic models\textsuperscript{17,8}

A Hammerstein model consists of a static non-linearity followed by a linear dynamical system, whereas a Wiener model has the order reversed, i.e. the linear dynamics are followed by a static non-linearity. Concerning neural network or fuzzy logic models, several different structures can come in question. Arahal et al.\textsuperscript{16} use a multilayer perceptron-based model, and a model based on RBF (radial basis function) networks whereas Rohani et al.\textsuperscript{9} mainly focus on a recurrent neural network and on a feed-forward neural network based model. Concerning the fuzzy models, both Fisher et al.\textsuperscript{17} and Sousa et al.\textsuperscript{8} mainly focus on the Takagi-Sukeno fuzzy model.

Here, we shall look at a quasi-ARMAX (or neuro-fuzzy) modeling scheme as a model to be used for in the NMPC formulation. We formulate the problem as a general non-linear, time-invariant system

\begin{equation}
y(t) = g(\varphi(t)) + v(t)
\end{equation}

\begin{equation}
\varphi(t) = \left[y(t-1),\ldots,y(t-n), u(t-d),\ldots,u(t-m-d+1)\right]^T
\end{equation}

\begin{equation}
v(t) = e(t) + c_1 e(t-1) + \ldots + c_p e(t-p)
\end{equation}

where \(y(t)\) is the output at discrete time-intervals, \(u(t)\) is the input, \(v(t)\) is a disturbance, \(d(t)\) is white noise, \(d\) is the time delay, \(\varphi(t)\) is the regression vector and \(g\) is a non-linear function.

A hybrid quasi-ARMAX model can be described according to the following. Perform a Taylor expansion of the nonlinear function \(g\), and use two coefficient vectors according to

\begin{equation}
y(t) = g(0) + g'(0)\varphi(t) + \frac{1}{2} g''(0)\varphi(t) + \ldots
\end{equation}

\begin{equation}
\theta = g'(0) = \left[a_1 \ldots a_n \ b_1 \ldots b_m\right]^T
\end{equation}

\begin{equation}
\Delta \theta_i = \left(\frac{1}{2} g''(0) + \ldots\right)^T
\end{equation}

\begin{equation}
= \left[\Delta a_{1,i} \ldots \Delta a_{n,i} \ \Delta b_{1,i} \ldots \Delta b_{m,i}\right]^T
\end{equation}

(where \(g(0)\) is ignored for the sake of simplicity.) The system can thus be expressed as a combined structure

\begin{equation}
y(t) = \varphi^T(t) (\theta + \Delta \theta_i) + v(t)
\end{equation}

The parameters \(\Delta \theta_i\) can be approximated\textsuperscript{9} as

\begin{equation}
\Delta a_{i,t} = f_i(\varphi(t)), \quad i = 1, \ldots, n
\end{equation}

\begin{equation}
\Delta b_{j,t} = f_{j+n}(\varphi(t)), \quad j = 1, \ldots, m
\end{equation}

where

\begin{equation}
f_i(\varphi(t)) = \sum_{j=1}^{M} \omega_{ij} N_f(p_j, \varphi(t))
\end{equation}

where \(N_f(p_j, \varphi(t))\) are fuzzy basis functions, \(\omega_{ij}\) are the coordinate parameters and \(p_j\) are the scale and position parameter vectors.

\begin{equation}
N_f(p_j, \varphi(t)) = \frac{\sum_{j=1}^{L} \mu_{A_{k,j}}(x_k(t))}{\sum_{j=1}^{L} \mu_{A_{k,j}}(x_k(t))}
\end{equation}

We use \(\Lambda\) as the minimum operator, \(r = \dim(\varphi(t))\), \(L\) is the number of fuzzy rules, \(x_k(t)\) are the elements of \(\varphi(t)\), and \(\mu_{A_{k,j}}\) is the membership function of fuzzy set \(A_{k,j}\). \(\mu_{A_{k,j}}\) may be a Gaussian function defined by

\begin{equation}
\mu_{A_{k,j}}(x_k(t)) = \exp\left[-\frac{1}{2} \frac{(x_k(t) - \bar{x}_{k,j})^2}{\sigma_j^2}\right]
\end{equation}
In this case, the parameter vector is defined by

\[ p_j = [x_1^j, x_2^j, \ldots, x_r^j, \sigma_j]^T, \quad j = 1, \ldots, L \]  \quad (12)

Our model can then be reformulated as a combined structure

\[ y(t) = \varphi^T(t)\theta + v(t) + \sum_{j=1}^{M} \varphi^T(t)\Omega_j N_j(p_j, \varphi(t)) \]

where \( \Omega_j = [\omega_{1j} \ldots \omega_{rj}] \).

Or, alternatively, the system can be expressed as a linear regression structure:

\[ y(t) = \varphi_{NL}^T(t)\Theta + e(t) \]  \quad (13)

where

\[ \Theta = [\theta^T \omega_{11} \ldots \omega_{rL} c_1 \ldots c_p]^T \]
\[ \varphi_{NL}(t) = [\varphi^T(t) \varphi^T(t) \otimes \varphi^T_{N_j}(t)]^T \]
\[ \varphi^T_{N_j}(t) = [N_j(p_j, \varphi(t)), \quad j = 2, \ldots, L] \]
\[ \theta = [a_1 \ldots a_n b_1 \ldots b_m] \]

Kronecker production is indicated by \( \otimes \). This gives us an implementable algorithm, where the parameters \( \Theta \) can be evaluated using e.g. a recursive least squares algorithm. This is the form of the algorithm implemented in the simulations below. This form is also usable for adaptive control.

The investigated predictor can be seen as a learning network based predictor interpreted as a linear predictor network consisting of a global linear predictor and several local linear predictors with interpolation. This predictor should thus be well suited for control purposes in e.g. process control, when the plant in question has to run at several different operating points. Furthermore, the predictor is computationally efficient, thus making it exceptionally well suited for use in NMPC, where the computational simplicity is crucial, since a non-linear optimization problem has to be solved in every sampling instant.

4. Quasi-Armax NMPC

The predictor presented above has, for control purposes, only been illustrated using a one-step ahead predictive controller\(^7\), often referred to as a myopic controller.

The myopic controller can be seen as the most simple case of model predictive control, i.e. the case with a prediction horizon and the control horizon = 1. It can quickly be illustrated according to the following. Choose a controller which minimizes the cost function

\[ J = \frac{1}{2}(y(k + 1) - y_r(k + 1))^2 + \lambda \frac{1}{2}u(k)^2 \]

where \( \lambda \) is a tuning factor. Substitute \( y(k + 1) \) with its prediction \( \hat{y}(k + 1) \), where the prediction is given by the model used.

We look at the example used in\(^7\), where the system under control is given by

\[ y(k) = G(z^{-1})z(k), \]
\[ G(z^{-1}) = \frac{-0.7z^{-1} - 0.68z^{-2}}{1 - 1.72z^{-1} + 0.74z^{-2}} \]

where \( z(k) \) is a nonlinear function of \( u(k) \) given by

\[ z(k) = \begin{cases} 
  u(k) - 1.75, & u(k) > 2 \\
  0.0625\text{sign}(u(k))u(k)^2, & |u(k)| \leq 2 \\
  u(k) + 1.75, & u(k) < -2 
\end{cases} \]  \quad (15)

which behaves like a dead zone. Assume that a linear ARX-model is used to model the system

\[ y(k + 1) = -a_1y(k) - a_2y(k - 1) \]
\[ +b_1u(k) + b_2u(k - 1) \]
\[ = 1.72y(k) - 0.74y(k - 1) \]
\[ +0.7u(k) - 0.68u(k - 1) \]

Taking the derivative of the loss function \( J \) with respect to \( u(k) \) and equaling it to zero gives an analytical expression for the control action

\[ u(k) = \frac{b_1}{(b_1^2 + \lambda)} [a_1y(k) + a_2y(k - 1) \]
\[ -b_2u(k - 1) + y_r(k + 1)] \]  \quad (16)

The myopic controller, although useful for illustrative reasons due to its simplicity, is not practically very useful for process control applications. A myopic controller e.g. requires the system to be minimum-phase and its I/O delay exactly known. This is a severely limiting assumption, especially in process control.

For the full NMPC case, we will minimize a cost function given by
\[
N, \quad J(N_{1}, N_{y}, N_{u}) = \sum_{j=N_{1}}^{N_{u}} \sigma(j)[\hat{y}(t+j|t) - y_{r}(t+j)]^2 \\
+ \sum_{j=1}^{N_{u}} \lambda(j)[u(t+j - 1)]^2
\]  

In our example the weights are chosen as \( \sigma = 1 \) and \( \lambda = \text{constant} \). Further, in our example we choose \( N_{1} = 1, N_{y} = 3 \) (the prediction horizon) and \( N_{u} = 3 \) (the control horizon). Also we use the change of the control input instead of the control input itself, in order to smoothen the control input. In process control, the smoothness of the control action is often of significance, in order to e.g. reduce wear on actuators.

\[
J(k) = \sum_{j=1}^{3} [\hat{y}(k+j) - y_{r}(k+j)]^2 \\
+ \sum_{j=1}^{3} \lambda \Delta u(k+j-1)^2
\]

For the example above (given by equation 14 and 15) we can identify a quasi-ARMAX model with e.g. 5 parameters in the global model, and 18 local models, active in different regions of the input space. The identification is performed with a recursive least squares algorithm, and the identified model output compares to the actual process output and to the output of an identified linear model as shown in Fig. 1.

The NMPC controller can then be implemented according to minimizing equation 18, where

\[
\begin{align*}
\hat{y}(k+1) &= \varphi_{NL}(k+1)\Theta \\
\varphi_{NL}(k+1) &= [\varphi^{T}(k+1) \varphi^{T}(k+1) \otimes \varphi_{NL}(k+1)]^{T} \\
\varphi_{NL}(k+1) &= [N_{f}(p_{j}, \varphi(k+1), j = 1, \ldots, L] \\
\varphi_{NL}(k+1) &= [y(k), \ldots, y(k-n+1), \\
&\quad u(k-d+1), \ldots, u(k-m-d+2)]^{T}
\end{align*}
\]

and \( \hat{y}(k+j), j = 2, 3, \ldots \) are calculated recursively. The minimization is performed at every sample instant, and only the first calculated control move \( (u(k)) \) is implemented. The parameter vector \( p_{j} \) is kept fixed. Control results for the deterministic case are illustrated in Fig. 2.

Another example, useful for demonstrating the principles of the proposed controller consists of a process working in a few distinct operating regions. In each region the process can be described by a linear model. Assume dynamics for three regions according to

\[
\begin{align*}
y(z^{-1}) &= 0.0137z^{-1} + 0.0092z^{-2}, \quad y < -1 \\
u(z^{-1}) &= 1 - 1.1866z^{-1} + 0.3012z^{-2}, \quad y < -1 \\
y(z^{-1}) &= 0.0274z^{-1} + 0.0184z^{-2}, \quad |y| < 1 \\
u(z^{-1}) &= 1 - 1.1866z^{-1} + 0.3012z^{-2}, \quad |y| < 1 \\
y(z^{-1}) &= 0.0548z^{-1} + 0.0368z^{-2}, \quad y > 1 \\
u(z^{-1}) &= 1 - 1.1866z^{-1} + 0.3012z^{-2}, \quad y > 1
\end{align*}
\]
i.e. the process gain changes with the operating region. The usefulness of the physical interpretation
of the parameters can be demonstrated through this (oversimplified) example. Let the global model correspond to the parameters in the region $-1 < y < 1$, and the neuro-fuzzy local models consist of the difference between the global model and the model at the current operating region. The fuzzy basis functions are defined so as to correspond to each of the distinct operating regions. Control results with this controller (prediction and control horizon = 3, and a tuning factor equal to 0.05) gives the simulated set-point changes in Fig.3, and as comparison a linear MPC controller is shown in Fig.4.

5. Conclusion

The purpose of this paper is to relate the quasi-ARMAX identification method to a thorough process control framework, thus this paper presents the derivation and application of a nonlinear quasi-ARMAX predictor to nonlinear model predictive control purposes.

The proposed controller shows good results, being able to combine good prediction capability with computational efficiency. The controller proposed should be well suited for use in the NMPC formulation, since the identification algorithm is computationally more efficient than e.g. a neural network based model.

Furthermore, the specific meaning of the parameters makes the predictor well suited for NMPC control, since the model thus can be understood and analyzed based on fundamental properties of the process to be controlled.

Several issues are still open concerning this modeling approach for use in NMPC applications. The robustness of the resulting control system seems promising, but a more rigorous evaluation should be performed. Furthermore, even though the parameters of the neuro-fuzzy model have explicit meanings, the tuning of the controller as a whole is still a quite ad hoc procedure. Also here, more thorough investigations are needed.

Acknowledgments

Jonas Waller gratefully acknowledges financial support from Svenska Tekniska Vetenskapsakademien i Finland.

References


