Controller reconfiguration for non-linear systems

S. Kanev*, M. Verhaegen

Department of Applied Physics, Systems and Control Engineering Group, University of Twente, PO Box 217, 7500AE Enschede, The Netherlands

Received 23 November 1999; accepted 2 May 2000

Abstract

This paper outlines an algorithm for controller reconfiguration for non-linear systems, based on a combination of a multiple model estimator and a generalized predictive controller. A set of models is constructed, each corresponding to a different operating condition of the system. The interacting multiple model estimator, extended for systems with offset, is utilized to yield a reconstruction of the state of the non-linear system and mode probabilities of the different models in the model set. Based on this information, a standard cost function in predictive control is optimized under the assumption that the mode probabilities are constant over the maximum costing horizon. The algorithm is illustrated for two different case studies — one with a linear model of one joint of a space robot manipulator, subjected to failures, and one with a non-linear model of the inverted pendulum on a cart.

Keywords: Fault-tolerant systems; Fault detection; Generalized predictive control; Nonlinear systems

1. Introduction

Modern control systems are becoming increasingly complex with more and more demanding performance goals. These complex systems must have the capability for fault accommodation in order to operate successfully over long periods of time. Such systems require fault detection, isolation and controller reconfiguration so as to maintain adequate levels of performance with one or more sensor, actuator, and/or component failures, or a combination of these events. The controller reconfiguration technique that is going to be presented in this paper, though applicable to a general non-linear system, is very suitable for the control of systems subject to failures, since such systems are (naturally) represented by a set of models (Athans et al., 1977; Maybeck & Stevens, 1991; Griffin & Maybeck, 1997; Zhang & Li, 1998).

When dealing with sensor, actuator and component failures, a hybrid dynamic model can be used. The hybrid system is also known as jump linear system: it is linear given the system mode; however it may jump from one such system to another at a random time. Such systems can be used to model situations where the system behavior undergoes abrupt changes, such as system failures (Zhang & Li, 1998). The hybrid dynamic model (Griffin & Maybeck, 1997; Zhang & Li, 1998) consists of a set of (ordinary) discrete-time linear models and a switching logic, determining the switching between these models. The switching between models in the hybrid systems is a consequence of factors, such as failures in its sensors, actuators and components. Different methods for the control of hybrid systems have been proposed in the literature. In Zhang and Jiang (1999), Campo, Bar-Shalom and Li (1996) an interacting multiple model-based control was utilized, a neural adaptive controller is presented in McDowell, Irwin, Lightbody and McConnell (1997), Multiple-model adaptive control (MMAC) is also an important class of control methods with application to the control of jump linear systems (Athans et al., 1977; Griffin & Maybeck, 1997; Narendra & Balakrishnan, 1997), an algorithm based on the Generalized Pseudo-Bayesian method is given in (Watanabe & Tzafestas, 1989). The optimal control of hybrid systems have also been addressed in the literature (Griffiths & Loparo, 1985).

A similar model representation might also be used to represent a non-linear dynamic system when approximating it by a piecewise linear system (PWL system). In the piecewise linear system description it is assumed that the state-space \( \mathcal{X} \) is divided into regions \( \mathcal{X}_i \). Linear dynamics are associated with each such region of the
The switching logic of the hybrid dynamic system representation in this paper is determined by the interacting multiple model (IMM) estimator, adopted from e.g. Zhang and Li (1998), Griffin and Maybeck (1997), Blom and Bar-Shalom (1988), Li (1996), and extended to the case of systems with offset in the state and output. In this approach, the switching logic corresponds to a set of real numbers that determine the convex combination of the models in the model set that is valid at a particular time instant, i.e. the convex combination of the states of the local models that represents the system state.

In the classical approach of gain scheduling using local models (Fabri & Kadirkanathan, 1998b; Hunt & Johansen, 1997) pre-designed controllers are activated when a detection mechanism detects that one (and only one) model from the model set is active at a particular time instant. When making use of this approach in fault-tolerant control, each failure condition should be represented by a single model. Thus, the model set \( \mathcal{M} \) will very quickly grow unboundedly as a consequence of having to model every (partial or total) failure condition. This problem is avoided in this paper by letting the state of the actual system be approximated by a convex combination of the states of the local models. In Griffin and Maybeck (1997) the authors propose a moving bank of filters to reduce the number of active models and the computational burden, i.e. only the models in close proximity to the model of the “real” system are activated.

In this paper it is assumed that the non-linear system is represented by a convex combination of a set of linear discrete-time models \( \mathcal{M} = \{ M_1, M_2, \ldots, M_N \} \). This set will be called the model set. The decision about which convex combination of these models is in effect at the current moment of time is made by the Interacting Multiple Model estimator. It runs a bank of Kalman filters in parallel, each based on a particular model from the model set \( \mathcal{M} \). It calculates the probability of each mode to be in effect. The overall state estimate is then computed as a probabilistically weighted sum (a convex combination) of the state estimates obtained from the different Kalman filters. A model of the system that is assumed to be currently in effect is constructed as a convex combination of the models in the model set. A bank of generalized predictive controllers (GPC) is designed, each corresponding to one of these local models. The optimal GPC control law for the model that is assumed to be in effect is calculated at each sample to minimize a standard cost function. This optimal control is not a convex combination of the local GPC control laws, optimized for each individual model in the model set.

In Maybeck and Stevens (1991) the authors have applied a multiple model adaptive control (MMAC) to a STOL F-15 aircraft. They combine a non-interacting MM algorithm with a bank of LQG controllers, each designed for one particular model. The overall control action is computed as a convex combination of the outputs of the different controllers, i.e.,

\[
\mathbf{u}(k) = \sum_{i=1}^{N} \mathbf{u}_i(\mathbf{x}_i(k)) \cdot \mu_i(k),
\]

where \( \mathbf{u}(k) \) is the overall control action, \( \mathbf{u}_i(\mathbf{x}_i(k)) \) is the output of the \( i \)th LQG controller (dependent on the state estimate of the \( i \)th Kalman filter, \( \mathbf{x}_i(k) \)) and \( \mu_i(k) \) is the probability that model \( M_i \) is in effect at the moment of time \( k \). However, although such a mixing of the controller outputs seems reasonable and intuitive, it does not guarantee optimality of the performance objective used in the design of the local controllers when the model in effect is not contained in the model set. A similar approach was followed in Athans et al. (1977). An illustration will be presented to show that in the case of unanticipated failures the closed-loop stability can no longer be guaranteed by such a control action.

The remaining part of this paper is organized as follows. In Section 2 the general descriptions of the hybrid dynamic system and the piecewise linear system are summarized. Some issues are also given on the model set design. Section 3 gives an overview of the interacting multiple model estimator, extended to the case when the model set consists of systems with offset, and makes some comments on the design of the transition probability matrix. In Section 4 the controller reconfiguration scheme is outlined by presenting a predictive controller strategy for a set of models. An illustration of this approach is made in Section 5 by means of two realistic simulation studies, one with a linear model of one joint of a space robot manipulator (SRM), and one with a non-linear model of the inverted pendulum on a cart. Finally, Section 6 is dedicated to some concluding remarks.

2. The model set

In this section the hybrid dynamic system and the piecewise linear system will be described. It will also pay attention to the model set design.

2.1. Hybrid dynamic model

A hybrid dynamic system can be described as one with both a continuously-valued base state and discretely valued structural/parametric uncertainty. A typical example of such a system is one subject to failures since
fault modes are structurally different from each other and from the nominal mode. By mode a structure or behavior pattern of the system is meant.

Assume that the actual system at any time can be modeled sufficiently accurately by a stochastic hybrid system (Griffin & Maybeck, 1997)
\[
\begin{align*}
x(k + 1) &= A(k, m(k + 1)) \cdot x(k) + B(k, m(k + 1)) \cdot u(k) \\
&\quad + T(k, m(k + 1)) \cdot \xi(k, m(k + 1)) \\
z(k) &= C(k, m(k)) \cdot x(k) + \eta(k, m(k))
\end{align*}
\]
(1)

with the system mode sequence \( m(k) \) assumed to be a Markov chain with transition probabilities
\[
P\{m_j(k + 1) | m_i(k)\} = \pi_{ij}(k), \quad \forall m_i, m_j \in \mathcal{M}
\]
(2)

and
\[
\sum_{i=1}^{N} \pi_{ij}(k) = 1, \quad i = 1, \ldots, N,
\]
(3)

where \( x \in \mathbb{R}^n \) is the state vector; \( z \in \mathbb{R}^m \) is the output of the system; \( u \in \mathbb{R}^p \) is the control input; \( \xi \in \mathbb{R}^q \) and \( \eta \in \mathbb{R}^r \) are independent identically distributed discrete-time process and measurement noises with means \( \xi(k) \) and \( \eta(k) \), and covariances \( Q(k) \) and \( R(k) \); \( m(k) \) is a discrete-valued modal state, i.e. the index of the normal or fault mode, at time \( k \), which denotes the mode in effect during the sampling period ending at \( k \); \( \pi_{ij} \) is the transition probability from mode \( m_i \) to mode \( m_j \). \( \mathcal{M} = \{m_1, m_2, \ldots, m_N\} \) is the set of all possible system modes (the model set).

It can be seen from (1) that the mode information is embedded (i.e., not directly measured) in the measurement sequence \( z(k) \).

In the algorithm outlined in this paper, the models in the model set \( \mathcal{M} = \{M_1, \ldots, M_N\} \) are described by
\[
M_i: \begin{align*}
\begin{cases}
x(k + 1) &= A_i(k) \cdot x(k) + B_i(k) \cdot u(k) \\
&\quad + T_i(k) \cdot \xi_i(k), \\
z(k) &= C_i(k) \cdot x(k) + \eta_i(k),
\end{cases} \\
&\quad i = 1, \ldots, N.
\end{align*}
\]
(4)

The matrices \( A_i, B_i, T_i \) and \( C_i \) may all be different for different \( i \).

2.2. Piecewise linear system

Consider the non-linear system
\[
\begin{align*}
x(k + 1) &= f(x, u, k), \\
z(k) &= g(x, u, k).
\end{align*}
\]

A piecewise linear system (e.g. Johansson, 1999; Johansson & Rantzer, 1998; Rantzer & Johansson, 1997) is a non-linear system, described by an equation whose right-hand side is a piecewise linear function of its arguments. For example, a linear system with saturated inputs results in system equations that are piecewise linear in the input variable. Such systems can also be obtained as a result of piecewise linear approximation of a non-linear system around different operating points.

However, throughout this paper the term piecewise will be interpreted as piecewise in the system state. Therefore, piecewise linear will indicate that the state space is divided into a set of regions \( \mathcal{X}_i \). The dynamics within each region is affine in the state vector \( x \), i.e.
\[
M_i: \begin{cases}
x(k + 1) &= A_i \cdot x(k) + a_i + B_i \cdot u(k) \\
z(k) &= C_i \cdot x(k) + c_i + D_i \cdot u(k)
\end{cases} \quad \text{for} \ x(k) \in \mathcal{X}_i.
\]
(5)

Thus, a piecewise linear system can be described as a collection of ordered pairs
\[
\{(M_i, \mathcal{X}_i)\}_{i=1}^N
\]
that associates a linear dynamics \( M_i \) to each region \( \mathcal{X}_i \). The index set \( I \) is \( \{1, \ldots, N\} \).

2.3. The model set design

The model set design is highly dependent on the particular application considered. However, there are some common features that have to be taken into account. For example, there should be enough separation (distance) between models so that they are identifiable by the IMM estimator. This separation should exhibit itself well in the measurement residuals. Otherwise, the IMM estimator will not be very selective in terms of correct fault detection since it is the measurement residuals that have the most dominant effect on the mode probability computation which in turn affects the accuracy of the overall state estimates. On the other hand, if the separation is too large, numerical problems may occur (Zhang & Li, 1998). The distances between the models should be measured in closed-loop because it is in closed-loop that the IMM estimator will be used. For example, one possible measure for the separation between two models, \( M_1(z) \) and \( M_2(z) \), is the \( \mathcal{H}_\infty \)-norm of the discrepancy between the corresponding closed-loop systems, \( M_{1,CL}(z) \) and \( M_{2,CL}(z) \), i.e. \( \|M_{1,CL}(z) - M_{2,CL}(z)\|_\infty \). Another possible way to define distances between models is the gap-metric (Vinnicombe, 1999). However, it is much more difficult to compute. Thus, a selection of a “good” model set turns out to be an extremely difficult task.

If systems subject to failures are considered, total actuator failures may be modeled by making zero(s) the appropriate column(s) of the \( B \) matrix. For total sensor failures one needs to annihilate the appropriate row(s) of the \( C \) matrix. Partial actuator or sensor failures are modeled by multiplying the appropriate column (row) of the \( B \) (or \( C \)) matrix by a scaling factor. For example, a partial 40% sensor failure is modeled by multiplying the corresponding row of the \( C \) matrix by 0.4. To prevent
ambiguity, note that in this way 100% failure means no failure at all, and that a 0% failure is a total failure.

However, although sensor and actuator failures can be represented in this manner, the problem of which particular failure conditions should be selected to form a “good” model set still stands.

Since there currently exists no systematic procedure for the choice of $\hat{m}$, in this paper it will be assumed that the model set is given. However, it is reasonable to select the models in the model set to correspond to total failures, or to 5–15% partial failures, since in this case the convex combination of the models would cover a greater set of possible faulty models. If, for example, one wants to be able to represent all sensor (actuator) failures in the interval (10%, 100%), one should build up a model that describes the system with the total (10%) sensor (actuator) failure (in addition to the nominal model). Also such a selection has the potential to make the distance between the models not too small.

3. The IMM estimator for systems with offset

This section will briefly summarize the IMM estimator (Zhang & Li, 1998; Griffin & Maybeck, 1997). The $i$th model from the model set is represented by

$$
\begin{align*}
M_i: & \quad x_i(k + 1) = A_i(k)x_i(k) + a_i(k) + B_i(k)u(k) \\
& \quad + T_i(k)\xi_i(k), \\
& \quad z(k) = C_i(k)x_i(k) + c_i(k) + \eta_i(k), \\
& \quad i = 1, \ldots, N.
\end{align*}
$$

The offsets $a_i(k)$ and $c_i(k)$ do not change the Kalman filters (as well as the IMM estimator) since they are additive to $T_i(k)\xi_i(k)$ and $\eta_i(k)$, respectively. Thus, in the original setting of the IMM estimator (Zhang & Li, 1998), the offsets in the state and in the output should just be added up, $a_i(k)$ and $c_i(k)$, to $T_i(k)\xi_i(k)$ and $\eta_i(k)$, in order to take them into account.

Table 1 presents a complete cycle of the IMM estimator with Kalman filters. The inherent parallel structure of the IMM estimator makes it very attractive for parallel processing.

The design parameters of the IMM algorithm are the transition probability matrix and the model set. Note that the performance of the IMM estimator depends also on the type and magnitude of control input excitation used. However, the design of the transition probability matrix $\pi$ is very important since the sensitivity of

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>One cycle of the IMM estimator for systems with offset</td>
</tr>
<tr>
<td>(1) Mixing of the estimates (for $j = 1, \ldots, N$)</td>
</tr>
</tbody>
</table>

Predicted mode probability:

$$
\mu_j(k + 1) = \sum_{i=1}^{N} \pi_{ij}\mu_i(k)
$$

Mixing probability:

$$
\pi_{ij} = \pi_{ij}\mu_i(k)/\mu_j(k)
$$

Mixing estimate:

$$
\hat{x}_j^0(k|k) = \sum_{i=1}^{N} \pi_{ij}\hat{x}_i(k|k)
$$

Mixing covariance:

$$
\mathbf{P}_j^0(k|k) = \sum_{i=1}^{N} \pi_{ij}\mathbf{P}_i(k|k) + \mathbf{X}_j^0(k|k) - \hat{x}_j^0(k|k)[\mathbf{X}_j^0(k|k)]
$$

(2) Model-conditional filtering (for $j = 1, \ldots, N$)

Predicted state:

$$
\hat{x}_j(k + 1|k) = A_j(k)\hat{x}_j^0(k|k) + a_j(k) + B_j(k)u(k) + T_j(k)\xi_j(k)
$$

Predicted covariance:

$$
\mathbf{P}_j(k + 1|k) = A_j(k)\mathbf{P}_j^0(k|k)A_j(k)^T + T_j(k)\mathbf{Q}(k)T_j(k)^T
$$

Measurement residual:

$$
\mathbf{v}_j = z(k + 1) - C_j(k)\hat{x}_j(k + 1|k) - c_j(k + 1) - \bar{\eta}_j(k + 1)
$$

Residual covariance:

$$
\mathbf{S}_j = C_j(k + 1)\mathbf{P}_j(k + 1|k)C_j(k + 1)^T + \mathbf{R}_j(k + 1)
$$

Filter gain:

$$
\mathbf{K}_j = \mathbf{P}_j(k + 1|k)C_j(k + 1)^T\mathbf{S}_j^{-1}
$$

Updated state:

$$
\hat{x}_j(k + 1|k + 1) = \hat{x}_j(k + 1|k) + \mathbf{K}_j\mathbf{v}_j
$$

Updated covariance:

$$
\mathbf{P}_j(k + 1|k + 1) = \mathbf{P}_j(k + 1|k) - \mathbf{K}_j\mathbf{S}_j\mathbf{K}_j^T
$$

(3) Mode probability update (for $j = 1, \ldots, N$)

Likelihood function:

$$
L_j(k + 1) = \frac{1}{\sqrt{2\pi|\mathbf{S}_j|}} \exp\left(-\frac{1}{2\mathbf{v}_j^T\mathbf{S}_j^{-1}\mathbf{v}_j}\right)
$$

Mode probability:

$$
\mu_j(k + 1) = \frac{\mu_j(k + 1|k + 1)\mathbf{L}_j(k)}{\sum_{i=1}^{N} \mu_i(k + 1|k + 1)\mathbf{L}_i(k)}
$$

(4) Combination of estimates

Overall state estimate:

$$
\hat{s}(k + 1|k + 1) = \sum_{i=1}^{N} \pi_i(k + 1|k + 1)\hat{x}_i(k + 1|k + 1)
$$

Overall covariance:

$$
\mathbf{P}(k + 1|k + 1) = \sum_{i=1}^{N} \pi_i(k + 1|k + 1)\mathbf{P}_i(k + 1|k + 1)
$$

+ \left[ \hat{s}(k + 1|k + 1) - \hat{x}_i(k + 1|k + 1) \right]^T \pi_i(k + 1|k + 1) \left[ \hat{s}(k + 1|k + 1) - \hat{x}_i(k + 1|k + 1) \right] \pi_i(k + 1|k + 1) \mathbf{P}_i(k + 1|k + 1) \pi_i(k + 1|k + 1)
the mode probabilities $\mu_i(k)$ with respect to $\pi$ is very high.

A recommended choice of the diagonal entries in the transition probability matrix is to match roughly the mean sojourn time of each mode:

$$\pi_{ii} = \max \left\{ l_i, 1 - T/\tau_i \right\},$$

where $\tau_i$ is the expected sojourn time of the $i$th mode; $T$ is the sampling interval; $l_i$ is a designed limit of the transition probability of the $i$th mode to itself. For example, the “normal-to-normal” transition probability can be obtained by $\pi_{11} = 1 - T/\tau_1$, where $\tau_1$ denotes the mean time between failures, which in practice, is significantly greater than $T$.

### 4. The MM-based GPC

In this section it will be shown how the GPC, adopted from Kinnaert (1989), can be extended to the case of systems with offset and combined with the IMM estimator to yield a technique for control of non-linear systems.

#### 4.1. The GPC for systems with offset

First, the optimal GPC for systems with offset will be derived. Consider the state-space model in the innovation form:

$$M: \begin{cases} \dot{x}(k + 1) = \tilde{A}x(k) + \tilde{a} + \tilde{B}\Delta u(k) + \tilde{K}e(k), \\ z(k) = \tilde{C}x(k) + \tilde{c} + e(k), \end{cases}$$

(6)

where $e(k)$ is an innovation sequence, $z(k) \in \mathbb{R}^m$, $\Delta u(k) \in \mathbb{R}^r$, $x(k) \in \mathbb{R}^n$, $\tilde{K}$ is the gain of the Kalman filter, and $\tilde{a}$ and $\tilde{c}$ are offsets in the system state and output, respectively.

Consider the filter $F$ given in state space by

$$F: \begin{cases} x_F(k + 1) = A_F x_F(k) + B_F z(k), \\ \psi(k) = C_F x_F(k) + D_F z(k), \end{cases}$$

(7)

where $\psi(k) \in \mathbb{R}^n$ is a vector of filtered output signals.

The augmented system is then obtained by combining (6) and (7):

$$S: \begin{cases} x(k + 1) = Ax(k) + a + B\Delta u(k) + Ke(k), \\ \psi(k) = Cx(k) + c + De(k), \end{cases}$$

(8)

where $x(k) = [\tilde{x}(k)\ x_F(k)]^T$, and

$$A = \begin{bmatrix} \tilde{A} & 0 \\ B_F \tilde{C} & A_F \end{bmatrix}, \quad a = \begin{bmatrix} \tilde{a} \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix},$$

$$C = [D_F \tilde{C} \ C_F], \quad c = D_F \tilde{c}, \quad D = D_F$$

and $K = \begin{bmatrix} \tilde{K} \\ 0 \end{bmatrix}$.

(9)

Now, if $x(k | \hat{x} - 1)$ is defined as the state estimate of $x(k)$ obtained from a Kalman filter, the following result holds.

**Theorem 4.1** ($j$-step ahead predictor for systems with offset). Consider the augmented system with offset (8). The best prediction (in terms of minimum covariance of the prediction error of $x(k+j)$) given the information $\{z(k), z(k-1), \ldots, u(k-1), u(k-2), \ldots\}$ is

$$\hat{x}(k + j | k) = \mathcal{A}^j\hat{x}(k | k - 1) + \sum_{i=0}^{j-1} \mathcal{A}^{j-i-1}a$$

$$+ \sum_{i=0}^{j-1} \mathcal{A}^{j-i-1}\mathcal{B}\Delta u(k + i) + \mathcal{A}^{j-1}\mathcal{K}e(k).$$

The best prediction of the filtered output signal is

$$\hat{\psi}(k + j | k) = \mathcal{C}\mathcal{A}^j\hat{x}(k | k - 1) + \sum_{i=0}^{j-1} \mathcal{C}\mathcal{A}^{j-i-1}a$$

$$+ \sum_{i=0}^{j-1} \mathcal{C}\mathcal{A}^{j-i-1}\mathcal{B}\Delta u(k + i) + \mathcal{C}\mathcal{A}^{j-1}\mathcal{K}e(k) + \mathcal{C}c.$$  (10)

For the proof see the appendix.

If the matrices are now formed

$$U(k) = [\Delta u^T(k), \ldots, \Delta u^T(k + NU - 1)]^T$$

and

$$\hat{\psi}(k) = [\hat{\psi}^T(k + N_1 | k), \ldots, \hat{\psi}^T(k + N_2 | k)]^T$$

the predictive model for the filtered output for $(N_2 - N_1 + 1)$ future time instants can be written as

$$\hat{\psi}(k) = HU(k) + I\hat{x}(k | k - 1) + WE(k) + T$$

with

$$H = \begin{bmatrix} \mathcal{C}\mathcal{A}^{N_1 - 1}\mathcal{B} & \mathcal{C}\mathcal{A}^{N_1 - 2}\mathcal{B} & \ldots & \mathcal{C}\mathcal{A}^{N_1 - NU}\mathcal{B} \\ \mathcal{C}\mathcal{A}^{N_2 - 1}\mathcal{B} & \mathcal{C}\mathcal{A}^{N_2 - 2}\mathcal{B} & \ldots & \mathcal{C}\mathcal{A}^{N_2 - NU}\mathcal{B} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}\mathcal{A}^{N_2 - 1}\mathcal{B} & \mathcal{C}\mathcal{A}^{N_2 - 2}\mathcal{B} & \ldots & \mathcal{C}\mathcal{A}^{N_2 - NU}\mathcal{B} \end{bmatrix},$$

$$\Gamma = \begin{bmatrix} \mathcal{C}\mathcal{A}^{N_1} \\ \mathcal{C}\mathcal{A}^{N_1 + 1} \\ \vdots \\ \mathcal{C}\mathcal{A}^{N_2} \end{bmatrix}, \quad W = \begin{bmatrix} \mathcal{C}\mathcal{A}^{N_1} & \mathcal{C}\mathcal{A}^{N_2} & \ldots & \mathcal{C}\mathcal{A}^{N_2 - 1}\mathcal{K} \end{bmatrix},$$

$$T = \begin{bmatrix} c + \sum_{j=0}^{N_1 - 1} \mathcal{C}a \\ c + \sum_{j=0}^{N_1 - 1} \mathcal{C}a \\ \vdots \\ c + \sum_{j=0}^{N_1 - 1} \mathcal{C}a \end{bmatrix}.$$  (11)
Theorem 4.2 (The GPC control law for systems with offset). Consider the controlled system (8) and the following cost function:

\[ J = E \left\{ \sum_{j=N_1}^{N_2} \| y(k+j) - \omega(k+j) \|^2 \right\} + \sum_{j=1}^{NU} \| \Delta u(k+j-1) \|^2 \]  

where \( N_1 \) is the minimum costing horizon, \( N_2 \) is the maximum costing horizon, \( NU \) is the control horizon, \( x_i = x_i(k) \), \( I_n \) is the \( (n_x \times n_x) \) identity matrix, \( r \) is an \( (n_x \times n_u) \) diagonal matrix \( r = \text{diag}(r_1) \), and \( \omega(k) \) is the vector of the references for each output. The optimal control law that minimizes this cost function is

\[ U(k) = -(H^T H + R)^{-1} H^T (f \tilde{x}(k) | k-1) + W e(k) + T - \Omega(k), \]

where

\[ R = \text{diag}(r) \]  

and

\[ \Omega(k) = [\omega^T(k+N_1), \ldots, \omega^T(k+N_2)]^T. \]

For the proof see the appendix.

4.2. The combination of the GPC with the IMM estimator

Next, the MM-based GPC for systems with offset will be derived. It is based on a combination of the IMM estimator and the GPC, both for systems with offset.

Consider the model set of augmented systems \( \mathcal{S} = \{ S_1, S_2, \ldots, S_N \} \), where each \( S_i \) is represented in form (8) by its own set of matrices \((A_{i}, A_{t}, B_{i}, C_{i}, D_{i}, K_{i})\).

Let also

\[ \hat{y}_{i}(k) = H_{i} U(k) + \Gamma_{i} \tilde{x}_{i}(k | k-1) + W_{i} e_{i}(k) + T_{i} \]

be the corresponding predictive model of the filtered output.

The results here also hold for the case of time varying systems. However, to simplify the expressions, the time \( k \) will be omitted in the matrices \( A_{i}, B_{i}, A_{t}, C_{i}, D_{i} \) and \( K_{i} \).

Remark 4.1. Since the innovation sequences \( e_{i}(t) \) are not measured, they have to be reconstructed from the knowledge of the process input and output signals. This can be achieved using the state-space models \( S_i \) and the estimates from the Kalman filters:

\[ e_{i}(k) = z(k) - C_{i} \tilde{x}_{i}(k | k-1) - c_{i}. \]

The state combination in the IMM estimator (see Table 1) allows the “true system” model to be presented as a probabilistically weighted sum (convex combination) of the models in the model set, i.e.

\[ S = \sum_{i=1}^{N} \mu_{i} S_{i}, \quad \mu_{i} \in \mathcal{R} \]

with

\[ \sum_{i=1}^{N} \mu_{i} = 1. \]

Then the matrices in the state-space description of \( S \) are

\[ \begin{bmatrix} A_{1} & 0 & \cdots & 0 \\ 0 & A_{2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{N} \end{bmatrix}, \quad \begin{bmatrix} a_{1} \\ a_{2} \\ \vdots \\ a_{N} \end{bmatrix}, \quad \begin{bmatrix} B_{1} \\ B_{2} \\ \vdots \\ B_{N} \end{bmatrix}, \]

\[ C = [\mu_{1} C_{1}, \mu_{2} C_{2}, \ldots, \mu_{N} C_{N}], \quad c = \sum_{i=1}^{N} \mu_{i} c_{i}, \]

\[ D = [\mu_{1} D_{1}, \mu_{2} D_{2}, \ldots, \mu_{N} D_{N}], \]

\[ \begin{bmatrix} K_{1} & 0 & \cdots & 0 \\ 0 & K_{2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & K_{N} \end{bmatrix} \]

The state vector for the system \( S \) is \( x(k) = [x_{1}^{T}(k), \ldots, x_{N}^{T}(k)]^{T} \), and the innovation sequence \( e(k) = [e_{1}^{T}(k), \ldots, e_{N}^{T}(k)]^{T} \). This is a fictitious model and it will not be used for state estimation and control.

Theorem 4.3 (The MM-based GPC for systems with offset). Consider the system \( S \) with state-space matrices given by (15), and with state reconstructed by the IMM estimator (given in Table 1). Then

(a) the predictive model for \( S \) is

\[ \hat{y}(k) = \sum_{i=1}^{N} \mu_{i}(k) \hat{y}_{i}(k). \]

(b) assuming that the mode probabilities do not change over the maximum costing horizon, i.e. \( \mu_{i}(k+j) = \mu_{i}(k), \forall j \leq N_2 \), the cost function for the system \( S \)

\[ J = E \left\{ \sum_{j=N_1}^{N_2} \| y(k+j) - \omega(k+j) \|^2 \right\} + \sum_{j=1}^{NU} \| \Delta u(k+j-1) \|^2 \]  

is minimized by

\[ U(k) = -(H^T H + R)^{-1} H^T (f \tilde{x}(k | k-1) + W e(k) + T - \Omega(k)), \]
Remark 4.2. Notice that although the global predictive model \( \bar{\Phi}(k) \) for the system \( S \) is a convex combination of the global predictive models \( \overline{\Phi}_i(k) \) for the systems \( S_i \), this is not the case with the optimal control law \( U(k) \), i.e., it cannot be represented as a convex combination of the optimal control laws \( U_i(k) \) obtained by the GPC controllers corresponding to the systems \( S_i \).

Since it is not necessary to compute the future control actions at the current time instant, a matrix \( L \) is formed of the first \( n_u \) rows of the matrix \( (H^TH + R)^{-1} \) and calculate only the current control action as (Table 2)

\[
\Delta u(k) = -LH^T(I\hat{x}(k|k-1) + We(k) + T - \Omega(k)).
\]

5. Simulation results

In this section two case studies will be presented. In the first one a linear model of one joint of a space robot manipulator (SRM) is used and the case is considered when sensor failures occur. The second example deals with a non-linear model of the inverted pendulum on a cart. The model set in this experiment is obtained by performing a piecewise linear approximation over the non-linear equations of the system.

5.1. Experiment with the SRM

The case study that will be presented in this section is simple but also a very illustrative one. It considers the linear model of one joint of a space robot manipulator (SRM) system. A schematic representation of the SRM is given in Fig. 1.

This experiment considers the case when a faulty model occurs which is not in the model set. This faulty model will be represented as a convex combination of the models in \( \mathcal{M} \).

The system parameters are given in Table 3.

The equations of motion of the SRM are as follows:

1. \( N^2I_m\ddot{\Omega} + I_{son}(\ddot{\Omega} + \dddot{\varphi}) + \beta(\dot{\Omega} + \dddot{\varphi}) = T_{eff}, \) \( (16) \)
2. \( I_{son}(\ddot{\Omega} + \dddot{\varphi}) + \beta(\dot{\Omega} + \dddot{\varphi}) = T_{def}. \) \( (17) \)

The actuator model of the motor plus the gearbox is

\( T_{eff} = NT_m, \) \( T_m = Ki, \) \( (18) \)

and the deformation torque \( T_{def} \) is described as

\( T_{def} = ce. \) \( (19) \)

Denote \( x_p = [\Omega, \dot{\Omega}, e, \dot{e}]^T, \)

\[
y_p = \begin{bmatrix} \Omega + e \\ N\dot{\Omega} \end{bmatrix}.
\]

Fig. 1. Structure of one joint of the SRM.
and $u_p = i_c$ as the input. Then the state-space model of the system is given by

$$\dot{x}_p = A_p x_p + B_p u_p$$

$$= \begin{bmatrix}
0 & 1 & 0 & 0 \\
0 & 0 & \frac{c}{N^2 I_m} & 0 \\
0 & 0 & 0 & 1 \\
0 & -\beta I_{son} & -\left(\frac{c}{N^2 I_m} + \frac{c}{I_{son}}\right) & -\beta I_{son}
\end{bmatrix} x_p$$

$$+ \begin{bmatrix}
0 \\
\frac{1}{N^2 I_m} NK_i i_c \\
0 \\
-\frac{1}{N^2 I_m}
\end{bmatrix}$$

This model is discretized with sampling period $T_s = 0.1$ (s).

The following two models comprise the model set in this experiment:

- $M_1$: the nominal (no faults) model (see Eq. (20) and Table 3).
- $M_2$: a faulty model, representing 10% (partial) failure of sensor no. 1.

Note, that each model representing a partial failure of sensor no. 1 in the interval (10%, 100%) can be written as a convex combination of the two models in $\mathcal{M}$, i.e. $\mathbb{P}(M_1, M_2)$.

The scenario for this experiment is the following:

- The system is in its normal mode of operation (model $M_1$ is active) in the time interval (0, 99).
- At $k = 100$ a 75% (partial) failure of sensor no. 1 occurs. It corresponds to the following convex combination of the models in the model set:

$$M_{REAL} = 0.7222 M_1 + 0.2778 M_2.$$

The following choice of the predictive control parameters is made: $N_1 = 1$, $N_2 = 15$, $NU = 8$, and $R = 0.02 I_8$, where $I_8$ is the $8 \times 8$ identity matrix. The transition probability matrix is selected as $\pi = \begin{bmatrix} 0.55 & 0.45 \\ 0.55 & 0.45 \end{bmatrix}$.

As a reference $\omega(k)$ a (low-pass) filtered step signal $\tilde{\omega}(k)$ from 0 to 1 at $k = 0$, and from 1 to 0 at $k = 80$ is selected. The filter used is the following:

$$x_F(k+1) = 0.8187 x_F(k) + 0.0906 \tilde{\omega}(k),$$

$$\omega(k) = 2 x_F(k). \quad (21)$$

Figs. 2 and 3 present the results from this experiment. Since in the initial experiments there were very big fluctuations in the mode probabilities, the following low-pass filter was introduced

$$\hat{\mu}_i(k) = 0.98 \hat{\mu}_i(k-1) + 0.02 \mu_i(k - 1).$$

---

Table 3

The relevant values of the parameters in the linear model of one joint of the SRM

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gearbox ratio</td>
<td>$N$</td>
<td>260.6</td>
</tr>
<tr>
<td>Joint angle of inertial axis</td>
<td>$\Omega$</td>
<td>Variable</td>
</tr>
<tr>
<td>Effective joint input torque</td>
<td>$T_0^{eff}$</td>
<td>Variable</td>
</tr>
<tr>
<td>Motor torque constant</td>
<td>$K_t$</td>
<td>0.6</td>
</tr>
<tr>
<td>The damping coefficient</td>
<td>$\beta$</td>
<td>0.4</td>
</tr>
<tr>
<td>Deformation torque of the gearbox</td>
<td>$T_{def}$</td>
<td>Variable</td>
</tr>
<tr>
<td>Inertia of the input axis</td>
<td>$I_m$</td>
<td>0.0011</td>
</tr>
<tr>
<td>Inertia of the output system</td>
<td>$I_{son}$</td>
<td>400</td>
</tr>
<tr>
<td>Joint angle of the output axis</td>
<td>$\theta$</td>
<td>Variable</td>
</tr>
<tr>
<td>Motor current</td>
<td>$i_c$</td>
<td>Variable</td>
</tr>
<tr>
<td>Spring constant</td>
<td>$c$</td>
<td>130,000</td>
</tr>
</tbody>
</table>

---

![Mode Probabilities](image-url)
It can be seen from Fig. 2 that during the normal operation of the system ($k < 100$), the probability that corresponds to model $M_1$ is equal to 1. Then, when at $k = 100$ a 75% failure occurs, the two mode probabilities change accordingly. To be more precise, their means during the second-half of the simulation time are

- $\bar{\mu}_1(k) = 0.7213$, for $k = 100, \ldots, 200$, and
- $\bar{\mu}_2(k) = 0.2787$, for $k = 100, \ldots, 200$

and as a result a model representing 74.92% failure of sensor no. 1 is detected.

Fig. 3 gives a plot of the system output, its prediction and the (filtered) reference signal.

As it was argued in the introduction, the existing MMAC algorithms (Maybeck & Stevens, 1991; Athans et al., 1977), based on a bank of LQG controllers, can be inefficient when an unanticipated failure occurs. This is because the optimal LQG controller for the model of such a failure is not a convex combination of the optimal LQG controllers, each based on a given model from the model set. To illustrate this an additional experiment is presented, in which two optimal LQG controllers were designed: one for the nominal system (model $M_1$), and one for the faulty model $M_2$, representing a 10% failure of sensor no. 1. In this scenario the unanticipated 75% failure of sensor no. 1 is in effect throughout the whole simulation. The reference signal was selected in the same way as in the simulation with the MM-based GPC (see Eq. (21)). Fig. 4 depicts the results from this experiment. It can be seen that in this experiment such a convex combination of the control actions led to an unstable closed-loop system.

5.2. Experiment with the inverted pendulum on a cart

The next experiment illustrates the application of the MM-based GPC to the control of non-linear systems. For this purpose a non-linear model of the inverted pendulum on a cart is considered (Khalil, 1992) as a control system. First, a piecewise linear approximation of the non-linear model around five different operating points is performed, leading to a model set of five models. Afterwards the MM-based GPC is applied to the non-linear model of the pendulum.
The dynamic equations of the system (see Fig. 5) are

\[
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= \frac{g}{l^2} \sin(x_1) - \frac{a}{l^2} \cos(x_1),
\end{align*}
\]

(22)

where \(x_1 = \Theta (\text{rad})\) is the angle between the pendulum and the vertical axis, \(x_2 = \dot{\Theta} (\text{rad/s})\) is the angular velocity of the pendulum, \(g = 9.81 (\text{m/s}^2)\) is gravity acceleration, \(l (\text{m})\) is the length of the pendulum, and \(a (\text{m/s}^2)\) is the acceleration of the cart.

**Remark 5.1.** A similar problem is encountered during a rocket launch, when the rocket boosters have to be fired in a controlled manner so as to maintain the upright position of the rocket.

Let the system be linearized around the point \(x^* = [x_1^*, 0]^T\). Defining

\[ f_2(x_1, x_2, a) = \frac{g}{l^2} \sin(x_1) - \frac{a}{l^2} \cos(x_1) \]

the linearized systems can be derived the following way (the first equation is linear and will not be considered):

\[
\begin{align*}
\dot{x}_2 &= f_2(x_1^*, x_2^*, 0) + \frac{\partial f_2}{\partial x_1} \bigg|_{x=x^*} (x_1 - x_1^*) \\
&\quad + \frac{\partial f_2}{\partial x_2} \bigg|_{x=x^*} (x_2 - x_2^*) + \frac{\partial f_2}{\partial a} \bigg|_{x=x^*} a \\
&= \frac{g}{l^2} \sin(x_1^*) + \frac{g}{l^2} \cos(x_1^*)(x_1 - x_1^*) + \frac{1}{l^2} \cos(x_1^*)a \\
&= \left( \frac{g}{l^2} \cos(x_1^*) \right)x_1 - \frac{g}{l^2} (\sin(x_1^*) + \cos(x_1^*))x_1^* \\
&\quad + \left( \frac{1}{l^2} \cos(x_1^*) \right)a.
\end{align*}
\]

Therefore, the system can be written in form (5) with

\[
\begin{align*}
A_1 &= \begin{bmatrix} g/l^2 \cos(x_1^*) & 0 \\ 0 & 1 \end{bmatrix}, & a_i &= \begin{bmatrix} 0 \\ -g/l^2 (\sin(x_1^*) + \cos(x_1^*))x_1^* \end{bmatrix},
\end{align*}
\]

where \(x_1^i\) is the \(i\)th linearization point. The model set considered for this simulation corresponds to the linearization points \(x_1^i = [0, \pm 10, \pm 20] (\text{deg})\).

All these models are discretized with sampling time \(T_s = 0.1 (\text{s})\). It was decided that models corresponding to angles \(|\Theta| > 20 (\text{deg})\) are unnecessary since the problem will be to maintain an angle of 10 (deg) between the pendulum and its upright position.

The following parameters were chosen:

- Length of the pendulum: \(l = 1 (\text{m})\).
- Minimum costing horizon: \(N_1 = 1\).
- Maximum costing horizon: \(N_2 = 5\).
- Control horizon: \(NU = 4\).
- Weights on the control action: \(R = \text{diag}[0, 0, 0, 0, 0]\).
- Reference signal \(w(k) = 10^{\pi}[1, 1, 1]^T (\text{rad})\). This corresponds to a setpoint of 10 (deg).
- Transition probability matrix:

\[
\pi = \begin{bmatrix}
0.95 & 0.0125 & 0.0125 & 0.0125 & 0.0125 \\
0.0125 & 0.95 & 0.0125 & 0.0125 & 0.0125 \\
0.0125 & 0.0125 & 0.95 & 0.0125 & 0.0125 \\
0.0125 & 0.0125 & 0.0125 & 0.95 & 0.0125 \\
0.0125 & 0.0125 & 0.0125 & 0.0125 & 0.95
\end{bmatrix}.
\]

- Means of the noises: \(\xi = \hat{n} = 0\).
- Covariances of the noises: \(Q = 10^{-6} I_2\) and \(R = 10^{-6}\).

The results of this simulation are depicted in Figs. 6 and 7. The simulations are made with the non-linear model of the inverted pendulum. Fig. 7 shows both the angle \(\Theta\) and its prediction \(\hat{\Theta}\) by the IMM estimator, which seem to overlap. A deviation between the two curves, of course, exists. Note, that when the system output gets close to 10, which corresponds to model \(M_2\) from the model set, the corresponding to this model mode probability \(\mu_2\) (see Fig. 6) gets close to one, i.e. \(\mu_2 \approx 1\). Since the pendulum never goes to negative degrees, the mode probabilities \(\mu_3\) and \(\mu_5\) stay at zero during the simulation.

### 6. Conclusions

This paper presented an algorithm for the control of non-linear systems represented by hybrid dynamic models or piecewise linear systems. The method consists of two parts: identification and controller reconfiguration. The first part is essentially the IMM estimator whose purpose is to give a state estimate and a mode probability for each model in the model set. The controller reconfiguration part utilizes this information to derive a GPC
action, assuming that the mode probabilities are constant over the maximum costing horizon.

The performance of the IMM estimator is strongly dependent on the choice of the transition probability matrix, as well as on the models in the model set \( \mathcal{M} \). A model set consisting of models close to one another results in a deterioration of the performance of the IMM estimator, which in turn affects the performance of the MM-based GPC.

Another very important issue is the choice of the transition probability matrix \( \pi \). It should be pointed out that serious difficulties regarding the selection of this matrix were encountered, since the IMM estimator turns out to be extremely sensitive to this design parameter. In addition, the entries in the transition probability matrix represent the probabilities for switches from one expected mode to another expected mode. However, they do not reflect the probabilities for jumps to unexpected modes.

**Acknowledgements**

The authors would like to say thanks to the Dutch National Aerospace Laboratory (NLR), and especially to Dr. J.F.T. Bos, for providing us with the data for the space robot manipulator.

**Appendix**

**Proof of Theorem 4.1.** It can be written that

\[
\dot{x}(k+j|k) = A \dot{x}(k+j-1|k) + a \\
+ B \Delta u(k+j-1) K e(k+j-1) \\
= A(A \dot{x}(k+j-2|k) + a + B \Delta u(k+j-2) \\
+ K e(k+j-2)) + a \\
+ B \Delta u(k+j-1) + K e(k+j-1) \\
= A(A(\ldots(A(A \dot{x}|k-1) + a + B \Delta u(k) \\
+ K e(k)) + a + B \Delta u(k+1) \\
+ K e(k+1)) + \cdots) + a + B \Delta u(k+j-1) \\
+ K e(k+j-1).
\]

Since the innovation \( e(k+j) \) is not known for \( j > 0 \), but is white noise, the best prediction of the state \( \dot{x}(k+j) \) is
obtained by taking \( e(k+j) = 0 \) for \( j > 0 \), i.e.
\[
\dot{x}(k+j | k) = A^j \dot{x}(k | k-1) + \sum_{i=0}^{j-1} A^{j-i-1} b a_i + \sum_{i=0}^{j-1} A^{j-i-1} B \Delta u(k+i) + A^{j-1} K e(k).
\]

Eqs. (10) then follows directly by observing that
\[
\dot{\hat{y}}(k+j | k) = C \dot{x}(k+j | k) + c. \quad \square
\]

**Proof of Theorem 4.2.** The cost function can be rewritten as
\[
J = (H u(k) + \Gamma \dot{x}(k | k-1) + W e(k) + T - \Omega(k))^T (H u(k) + \Gamma \dot{x}(k | k-1) + W e(k) + T - \Omega(k)) + U(k)^T R(U(k)).
\]

For the sake of simplicity the following substitution is made
\[
Q = \Gamma \dot{x}(k | k-1) + W e(k) + T - \Omega(k).
\]
Notice that \( Q \) is independent on \( U(k) \). Therefore,
\[
J = (H u(k) + Q)^T (H u(k) + Q) + U(k)^T R(U(k))
\]
\[
= U(k)^T H^T H u(k) + U(k)^T H^T Q + Q^T H u(k) + Q^T Q + U(k)^T R(U(k))
\]
\[
= U(k)^T (H^T H + R) U(k) + U(k)^T H^T Q + Q^T Q.
\]

Taking the partial derivative of \( J \) with respect to \( U(k) \) yields
\[
\frac{\partial J}{\partial U(k)} = 2 (H^T H + R) U(k) + H^T Q.
\]

Therefore, the control action that minimizes the cost function (12) is
\[
U(k) = -(H^T H + R)^{-1} H^T (\Gamma \dot{x}(k | k-1) + W e(k) + T - \Omega(k))
\]
which completes the proof. \( \square \)

**Proof of Theorem 4.3.** (a) The prediction of the filtered output for the system \( S \) can be written as (see Theorem 4.1)
\[
\hat{y}(k+j | k)
\]
\[
= C A^j \hat{x}(k | k-1) + \sum_{i=0}^{j-1} C A^{j-i-1} B a_i + \sum_{i=0}^{j-1} C A^{j-i-1} B \Delta u(k+i) + C A^{j-1} K e(k) + c
\]
\[
= \sum_{i=1}^{N} \mu_i C A^i \hat{x}_i(k | k-1) + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} A p a_i + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} B p \Delta u(k+p) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + c
\]
\[
= \sum_{i=1}^{N} \mu_i \hat{y}_i(k+j | k) + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} A p a_i + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} B p \Delta u(k+p) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + c
\]
\[
= \sum_{i=1}^{N} \mu_i \hat{y}_i(k+j | k) + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} A p a_i + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} B p \Delta u(k+p) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + c
\]
\[
(\text{b) With } \hat{y}(k+j | k) \text{ given by part (a) of this theorem, it can be written}
\]
\[
\hat{y}(k) = \sum_{i=1}^{N} \mu_i \hat{y}_i(k) + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} A p a_i + \sum_{i=1}^{N} \sum_{p=0}^{j-1} \mu_i C A^{i-1} B p \Delta u(k+p) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + \sum_{i=1}^{N} \mu_i C A^{i-1} K e_i(k) + c
\]
\[
\Rightarrow \hat{y}(k) = \sum_{i=1}^{N} \mu_i \hat{y}_i(k).
\]

Application of Theorem 4.2 to this predictive model yields the optimal control law
\[
U(k) = -(H^T H + R)^{-1} H^T (\Gamma \dot{x}(k | k-1) + W e(k) + T - \Omega(k))
\]
\[
+ W e(k) + T - \Omega(k)). \quad \square
\]

**References**


