Int. J. Appl. Math. and Comp. Sci., 1999, Vol.9, No.1, 175–191

ADAPTIVE PREDICTIVE CONTROLLER USING ORTHONORMAL SERIES FUNCTIONS

GUSTAVO H.C. OLIVEIRA*, WAGNER C. AMARAL**
GÉRARD FAVIER***

A constrained adaptive predictive control method that uses uncertain process modelling based on orthonormal series functions is considered. Such unstructured modelling is described as a weighted sum of orthonormal functions using approximate information about the time constant of the process. The orthonormal series functions model can thus be used to derive a j-step-ahead output prediction according to the constrained adaptive predictive control law. In relation to predictive controllers based on structured models, this approach presents the advantage of not requiring prior knowledge of the order or time delay, which decrease prediction errors and lead to a better closed loop performance when these parameters are not well known. Stability issues of the proposed control scheme are discussed and, finally, a simulation example is given to show the performance of the algorithm.

Keywords: model-based predictive control, adaptive control, uncertain process, orthonormal series functions.

1. Introduction

Model Based Predictive Controllers (MBPC) are, by definition, based on predicted behavior of the process. The principle of the MBPC control law consists in calculating the control input by the minimization of a cost function over a future time horizon under certain constraints of the process. The cost function is defined in terms of the tracking error, i.e., the difference between the predicted output and desired set-point. Output predictions are made by using a model of the process; hence, the closed loop performance depends on the choice of an appropriate model for prediction. Using this scheme, many different MBPC algorithms have been proposed in the literature (Clarke, 1994). When model parameters are unknown or time varying, they can be estimated by an identification method, with the estimated model used to establish

* LAS/CCET/PUC PR, R. Imaculada Conceição, 1155, 80215–901, Curitiba/PR, Brazil, e-mail: oliv@rla01.pucpr.br.

** DCA/FEEC/Unicamp, PO Box 6101, 13085–970, Campinas/SP, Brazil, e-mail: wagner@dca.fee.unicamp.br.

*** Lab. I3S/CNRS, 2000, Route des Lucioles, Sophia Antipolis, F–06410, Biot, France, e-mail: favier@i3s.unice.fr.

the control law, following the standard indirect adaptive control procedure. This procedure can easily be applied in MBPC controllers.

Adaptive MBPC based on structured models such as, for instance, the Generalized Predictive Controller (GPC) (Clarke and Mohtadi, 1989) which uses a CARIMA model, provides reasonable performance when the model structure is well selected, i.e., when the model order and time delay are known. When these parameters are not well known, however, the closed loop performance can be deteriorated or even become unstable (Rohrs *et al.*, 1982). In fact, this kind of MBPC is not robust to uncertainties in the process structure. MBPC algorithms based on unstructured models, such as the finite impulse response (FIR) model, can overcome this problem, i.e., when exact information about these parameters is unavailable. However, due to the intrinsically infinite dimensional characteristics of this representation, the number of parameters to estimate on-line can become very large, even in the case of a simple process.

In the present paper, the use of an unstructured model representation based on orthonormal series functions in the MBPC algorithm is considered. As an unstructured model, there is no need of exact specification of the order or time delay of the process and, moreover, the degree of freedom given by the choice of the basis functions can increase the rate of convergence of the series coefficients, reducing the model parameters. The FIR model can be viewed as a special case of this kind of modelling. The characteristics of orthonormal functions models in the context of system identification has already been analyzed by several authors (Dumont, 1998; Gunnarsson and Wahlberg, 1991; Ninness and Gustafsson, 1995; Olivier, 1994; Wahlberg, 1991a; 1991b; Wahlberg and Makila, 1996), and some works describing its application in both predictive control and robust predictive control algorithms can be found in (Dumont, 1998; Elshafei *et al.*, 1994; Finn *et al.*, 1993; Oliveira *et al.*, 1996a; 1996b; 1997; Zervos and Dumont, 1988). Here, a review of these works including the case of Kautz functions, input/output signals constraints and infinite norm cost functions is presented.

The paper is organized as follows. Section 2 presents a description of orthonormal series function modelling, whereas Section 3 describes a constrained adaptive predictive controller using this kind of modelling. Section 4 discusses the stability of the closed loop system and Section 5 illustrates the performance of the algorithm with a simulation example; finally, Section 6 presents the conclusions.

2. Orthonormal Series Function Modelling

A stable linear system can be characterized by its impulse response $h(k)$ and, supposing a causal system with $h(k)$ in the Lebesgue space $L_2[0,\infty[$, the signal $h(k)$ can be modeled by an orthonormal basis function expansion, as follows:

$$h(k) = \sum_{i=1}^{\infty} c_i \phi_i(k) \tag{1}$$

where $\{\phi_i(k)\}_{i=1}^{\infty}$ are orthonormal basis functions and c_i represents the set of the parameters associated with this orthonormal basis expansion.

Various orthonormal basis functions can be used to model such a stable system and a basis constructed using the knowledge of the system poles, such as the one presented in (Broome, 1965; Heuberger *et al.*, 1995; Ninness and Gustafsson, 1995), is used in this paper. The $\mathcal{Z}$-transform of such a basis is as follows:

$$\Phi_i(z) = z^{-1}\frac{\sqrt{1-|p_i|^2}}{1-p_i z^{-1}} \prod_{k=1}^{i-1}\left(\frac{z^{-1}-\overline{p}_k}{1-p_k z^{-1}}\right), \quad i = 1,\dots \tag{2}$$

where p_i represents the set of system poles and $\overline{p}_i$ is the complex conjugate pole of p_i. It has been shown that, in the modelling of the process, if an orthonormal series has an infinite number of functions, with its poles strictly inside the unit circle, the basis (2) is complete in the Lebesgue space (Ninness and Gustafsson, 1995). This means that any stable system can be modeled using this approach.

Moreover, there are orthonormal bases which use functions constructed with a single pole for the development of a series of $h(k)$, i.e., the Laguerre and Kautz functions (Lindskog, 1996; Wahlberg and Makila, 1996). These are special cases of the base presented in (2), as will be discussed below.

When $p_i = p$, $\forall\ i$, with $p \in \mathbb{R}$, is set, the expression (2) is reduced to the Laguerre function:

$$\Phi_{\text{lagu},i}(z) = \sqrt{1-p^2}\ \frac{z^{-1}(z^{-1}-p)^{i-1}}{(1-pz^{-1})^i}, \quad i = 1,\dots \tag{3}$$

Also, when $p = 0$, Φ_i is given by:

$$\Phi_{ir,i}(z) = z^{-i}, \qquad i = 1,\dots \tag{4}$$

This special type of orthonormal functions results in an impulse response model, so this model is a particular case of the orthonormal series functions model. The Kautz functions constitute an orthonormal basis defined by the use of a pair of complex poles $(p\,,p^*)$ (Ninness and Gustafsson, 1994; Wahlberg, 1991b) and are given by

$$\Phi_{\text{kautz},i}(z^{-1}) = \begin{cases} \dfrac{z^{-2}\sqrt{(1-\alpha^2)(1-\gamma^2)}}{1-\alpha(\gamma+1)z^{-1}+\gamma z^{-2}}\left(\dfrac{\gamma-\alpha(\gamma+1)z^{-1}+z^{-2}}{1-\alpha(\gamma+1)z^{-1}+\gamma z^{-2}}\right)^{(i-1)/2} & \text{for } i \text{ odd} \\[2ex] \dfrac{z^{-1}\sqrt{(1-\gamma^2)}(1-\alpha z^{-1})}{1-\alpha(\gamma+1)z^{-1}+\gamma z^{-2}}\left(\dfrac{\gamma-\alpha(\gamma+1)z^{-1}+z^{-2}}{1-\alpha(\gamma+1)z^{-1}+\gamma z^{-2}}\right)^{i/2} & \text{for } i \text{ even} \end{cases} \tag{5}$$

and $i = 1,\dots$, where $|\alpha| < 1$ and $|\gamma| < 1$ are defined such as p and p^* are the roots of $z^2 - \alpha(\gamma+1)z + \gamma$. Due to these characteristics, the Kautz functions should be used to model systems with resonant dynamics.

In the modelling of an actual process, only a finite number of functions ϕ_i can be used to approximate $h(k)$. Hence the series (1) has to be truncated to n terms and is given by

$$\tilde{h}(k) = \sum_{i=1}^{n} c_i\phi_i(k) \tag{6}$$

where $\tilde{h}(k)$ is an n-th order approximation of $h(k)$ and e is the truncation error, computed $e = \sum_{k=0}^{\infty} |h(k) - \tilde{h}(k)|$.

In this way, the input and output signals of a stable process are related by means of the following orthonormal series of functions:

$$y(k) = \sum_{i=1}^{n} c_i \Phi_i(q) u(k) = \sum_{i=1}^{n} c_i l_i(k) \tag{7}$$

where q is the shift operator and $l_i(k)$ is the output of the i-th function $\Phi_i(q)$. This can be expressed more compactly as follows:

$$y(k) = \boldsymbol{c}^T \boldsymbol{l}(k) \tag{8}$$

The vector $\boldsymbol{c} = [\, c_1 \ \cdots \ c_n \,]^T$ gives the series coefficients and the vector $\boldsymbol{l}(k) = [\, l_1(k) \ \cdots \ l_n(k) \,]^T$ is the basis functions state vector.

The Laguerre and Kautz functions, which are typical orthonormal functions derived from (2), are also recursive functions, i.e., the i-th function can be written using the $(i-1)$-th one. Thus, it is possible to describe the basis-function state vector $\boldsymbol{l}(k)$ using the following equations:

$$\begin{aligned} \boldsymbol{l}(k+1) &= \boldsymbol{A}\boldsymbol{l}(k) + \boldsymbol{b}u(k) \\ y(k) &= \boldsymbol{c}^T \boldsymbol{l}(k) \end{aligned} \tag{9}$$

Here the matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$ depend on the pole p and on the number n of functions used in the series expansion. In the case of Laguerre functions, they are given by

$$\boldsymbol{A} = \begin{bmatrix} p & 0 & 0 & \cdots & 0 \\ 1-p^2 & p & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ (-p)^{n-2}(1-p^2) & (-p)^{n-3}(1-p^2) & & \cdots & p \end{bmatrix} \tag{10}$$

$$\boldsymbol{b} = \sqrt{1-p^2} \begin{bmatrix} 1 & -p & (-p)^2 & \cdots & (-p)^{n-1} \end{bmatrix}^T \tag{11}$$

This orthonormal series approach is able to model a stable process with an impulse response in the Lebesgue space. However, processes with integral action can also be modeled as follows:

$$\begin{aligned} \boldsymbol{l}(k+1) &= \boldsymbol{A}\boldsymbol{l}(k) + \boldsymbol{b}u(k) \\ \Delta y(k) &= \boldsymbol{c}^T \boldsymbol{l}(k) \end{aligned} \tag{12}$$

where $\Delta = 1 - q^{-1}$.

In this way, the model is constructed by selecting the pole p (a real pole or a pair of complex ones) and the number n of functions. The c_i coefficients characterize the process dynamics. Due to the property of completeness, a linear stable process can be modeled by using any selection of p. Although this selection is not crucial, it is important for ensuring that the coefficients of higher order functions in the series moves quickly towards zero. This means that some of the higher order functions can be eliminated, thus reducing the number of model parameters. In this way, p is usually selected by using *a-priori* knowledge of the dominant dynamics of the process (Ninness and Gustafsson, 1995). Various algorithms for determining the best selection of p have been proposed in the literature, e.g., (Fu and Dumont, 1993), which proposes an algorithm for finding the best p to minimize the number of series functions n and (Masnadi-Shirazi and Ahmed, 1991), which proposes an algorithm for finding the best p to minimize the series truncation error e. These algorithms use process impulse response coefficients to compute the optimal value for p.

The ideal number of functions in a series is such that it makes the truncation error tend to zero. In practice, however, the selection of n depends on the process complexity, which means that the more complex the process is, the more functions will be necessary for process modelling. In the case of undamped systems, the use of Kautz functions based models results in a model with a smaller number of parameters than the equivalent Laguerre functions based model. However, generally the selection of 10 functions is sufficient to make the truncation error approximately equal to zero. The computation of the matrix $\boldsymbol{c}$ is discussed below.

The coefficients c_i, $i = 1, \dots, n$, can be computed by using the process impulse response $h(k)$, as follows:

$$c_i = \sum_{k=0}^{\infty} h(k)\Phi_i(q)\delta(k) \tag{13}$$

where $\delta(k)$ is the unit impulse function. For a model with integral action (12), the coefficients of matrix $\boldsymbol{c}$ are computed by filtering the process impulse signal with the use of Δ.

However, when the impulse response of the process is not available or it is time varying, an indirect adaptive control scheme can be used. If the model (8) is expressed as an ordinary linear regression, then the classical RLS algorithm (Ljung, 1987) or some of its variations (Latawiec, 1998; Shook *et al.*, 1991; Yoon and Clarke, 1994), can be used to estimate on-line the matrix $\boldsymbol{c}$ parameters.

The identification properties of orthonormal series functions modelling have been studied by various authors (Dumont, 1998; Gunnarsson and Wahlberg, 1991; Lindskog and Wahlberg, 1993; Ninness and Gustafsson, 1995; Olivier, 1994; Oliveira *et al.*, 1998; Van den Hof *et al.*, 1995; Wahlberg and Makila, 1996; Zervos and Dumont, 1988) who delineate some nice characteristics, such as the fact that there is no need for knowledge about the order or time delay of the process to use the identification algorithm. This represents an advantage over structured model based MBPC when exact information about these parameters in unavailable. Other properties of this OSF modelling are summarized as follows.

If the number n of parameters in the model is changed, the coefficients of low order in the orthonormal series remain almost constant (Zervos and Dumont, 1988), so it is easy to adjust the number of parameter on-line during the identification phase, in contrast to CARIMA models, where a change in the model order leads to the change in almost all model parameters.

By the use of approximate *a-priori* knowledge of the dominant time constant of the process in the selection of p, the OSF modelling is able to represent a model with fewer parameters c_i than impulse response modelling, leading to better quality of estimation, i.e., a smaller estimator variance. The OSF model also reduces the mean square error (MSE) of the estimation, in relation to impulse response modelling (Gunnarsson and Wahlberg, 1991).

To apply the OSF model in the development of predictive controllers, it is necessary to compute the j-step ahead output predictions using the model (9), as follows:

$$\hat{y}(k+j/k) = \hat{y}(k+j-1/k) + \boldsymbol{c}^T \Delta \boldsymbol{l}(k+j) \tag{14}$$

By successively substituting $\Delta \boldsymbol{l}(k+j)$, $\hat{y}(k+j-1/k)$ and assuming $\Delta u(k+j/k) = 0 \;\; \forall\, j \geq N_u$, with N_u being the control horizon, which is standard in MBPC strategy, we have

$$\hat{y}(k+j/k) = y(k) + \boldsymbol{c}^T(\boldsymbol{K}_j - \boldsymbol{I})\Delta \boldsymbol{l}(k) + \boldsymbol{c}^T \sum_{m=1}^{N_u} \boldsymbol{K}_{j-m} \boldsymbol{b} \Delta u(k+m-1/k) \tag{15}$$

where $\boldsymbol{K}_j = \sum_{i=0}^{j} \boldsymbol{A}^i$, with $\boldsymbol{A}^i = 0$ for $i < 0$ and $\boldsymbol{K}_j = 0$ for $j < 0$, $\boldsymbol{l}(k) = 0$ for $k \leq 0$. $\boldsymbol{I}$ is the n-th order identity matrix and $\Delta u(k+j-1/k)$ is the incremental control signal $k+j-1$, computed at k.

For processes with integral action, the output predictions are given by model (12). Hence

$$\hat{y}(k+j/k) = \hat{y}(k+j-1/k) + \boldsymbol{c}^T \boldsymbol{l}(k+j) \tag{16}$$

By successively substituting $\Delta \boldsymbol{l}(k+j)$ and $\hat{y}(k+j-1/k)$, the following equation is obtained:

$$\begin{aligned} \hat{y}(k+j/k) &= y(k) + \boldsymbol{c}^T(\boldsymbol{K}_j - \boldsymbol{I})\boldsymbol{l}(k) + \boldsymbol{c}^T \boldsymbol{\mathcal{K}}_{j-1} \boldsymbol{b} u(k-1) \\ &\quad + \boldsymbol{c}^T \sum_{m=1}^{N_u} \boldsymbol{\mathcal{K}}_{j-m} \boldsymbol{b} \Delta u(k+m-1/k) \end{aligned} \tag{17}$$

where $\boldsymbol{\mathcal{K}}_j = \sum_{i=0}^{j} \boldsymbol{K}_i$, with $\boldsymbol{\mathcal{K}}_j = 0$ for $j < 0$.

The set of output predictions for $j = N_1, \ldots, N_y$ can be expressed as

$$\hat{\boldsymbol{y}} = \boldsymbol{G} \Delta \boldsymbol{u} + \hat{\boldsymbol{y}}_l \tag{18}$$

where $\hat{\boldsymbol{y}} = [\hat{y}(k+N_1/k) \;\cdots\; \hat{y}(k+N_y/k)]^T$; $\Delta \boldsymbol{u} = [\Delta u(k/k) \;\cdots\; \Delta u(k+N_u-1/k)]^T$; $\hat{\boldsymbol{y}}_l = [\hat{y}_l(k+N_1/k) \;\cdots\; \hat{y}_l(k+N_y/k)]^T$. Here $\hat{y}_l(\cdot)$ is the part of the output predictions (15) and (17) that depends only on the process past behavior. The matrix $\boldsymbol{G}$ is given by

$$\boldsymbol{G} = \begin{bmatrix} g_{N_1} & \cdots & g_{N_1-N_u+1} \\ \vdots & \vdots & \vdots \\ g_{N_y} & \cdots & g_{N_y-N_u+1} \end{bmatrix} \tag{19}$$

where g_i is equal to $\boldsymbol{c}^T\boldsymbol{K}_{i-1}\boldsymbol{b}$ in the case of eqn. (15), or equal to $\boldsymbol{c}^T\mathcal{K}_{i-1}\boldsymbol{b}$ in the case of eqn. (17) (processes with integral action).

3. Constrained Adaptive Predictive Controller

An adaptive predictive controller is described by using a model (the parameters of which are obtained through an identification algorithm) to compute the predicted process output. Also, a cost function related to the closed loop performance of the system is defined, and a control signal is obtained by the minimization of the cost function. Finally, the first of these signals is applied in the process (a receding horizon strategy).

The predicted output j-step ahead $\hat{y}(k+j/k)$ is calculated as described above and used to derive the control law, resulting in an Adaptive Predictive Control based on the OSF modelling (PC-OSF).

The cost function of the Adaptive PC-OSF is defined by using the output prediction error, relative to the system set-point, and the weighted control signal, which can lead to a quadratic cost function as follows:

$$J_q(\Delta\boldsymbol{u}) = \sum_{j=N_1}^{N_y} \Big(\hat{y}(k+j/k) - w(k+j)\Big)^2 + \sum_{j=1}^{N_u} \lambda\Big(\Delta u(k+j-1/k)\Big)^2 \tag{20}$$

where N_1 and N_y are the prediction horizons, N_u is the control horizon, $w(\cdot)$ is the set-point and λ is a weighting factor; or can lead to an infinite norm cost function, as follows:

$$J_\infty(\Delta\mathbf{u}) = \max_j \left|\hat{y}(k+j/k) - w(k+j)\right|, \quad j = N1, \ldots, NY \tag{21}$$

Taking the classical approach for the minimization of the cost function (as in the case of GPC and DMC controllers), the Adaptive PC-OSF control law is obtained as

$$\Delta\boldsymbol{u} = \arg\min J(\Delta\boldsymbol{u})$$

subject to

$$\begin{aligned} &\Delta u(k+j/k) = 0 && \forall\, j \geq N_u \\ &u_{\min} \leq u(k+j-1/k) \leq u_{\max} && \forall\, j = 1,\ldots,N_u \\ &\Delta u_{\min} \leq \Delta u(k+j-1/k) \leq \Delta u_{\max} && \forall\, j = 1,\ldots,N_u \\ &y_{\min} \leq \hat{y}(k+j/k) \leq y_{\max} && \forall\, j = N_1,\ldots,N_y \end{aligned} \tag{22}$$

where $J(\cdot)$ is the cost function and $\Delta \boldsymbol{u}$ is the optimal value of the control signal, computed at k. The signal $u(k)$, applied in the process, is obtained from the optimal $\Delta \boldsymbol{u}$ vector as $u(k){=}u(k-1)+\Delta u(k/k)$.

When the quadratic cost function is used, the optimization problem (22) can be rewritten as follows:

$$\min_{\Delta \boldsymbol{u}} \ \Delta \boldsymbol{u}^T \boldsymbol{Q} \Delta \boldsymbol{u} + \boldsymbol{f}^T \Delta \boldsymbol{u}$$

subject to

$$\boldsymbol{\mathcal{A}} \Delta \boldsymbol{u} \leq \boldsymbol{v} \tag{23}$$

where

$$\boldsymbol{Q} = \boldsymbol{G}^T \boldsymbol{G} + \lambda \boldsymbol{I} \tag{24}$$

$$\boldsymbol{f} = 2\boldsymbol{G}^T \left(\hat{\boldsymbol{y}}_l - \boldsymbol{w} \right) \tag{25}$$

with $\boldsymbol{w} = [\, w(k+N_1) \ \ \cdots \ \ w(k+N_y)\,]^T$. $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{v}$ are built using the information about the process constraints.

The infinite norm does not attempt to minimize errors at all future time horizons as in the case presented above, but only at the instant where the error is maximum. So, using the infinite norm cost function, the control law (22) can be rewritten as follows:

$$\min_{\Delta \boldsymbol{u}, \mu} \ \mu$$

subject to

$$\begin{array}{ll} -\mu \leq \hat{y}(k+j/k) - w(k+j) \leq \mu & \forall\, j = N_1, \ldots, N_y \\ u_{\min} \leq u(k+j-1/k) \leq u_{\max} & \forall\, j = 1, \ldots, N_u \\ \Delta u_{\min} \leq \Delta u(k+j-1/k) \leq \Delta u_{\max} & \forall\, j = 1, \ldots, N_u \\ y_{\min} \leq \hat{y}(k+j/k) \leq y_{\max} & \forall\, j = N_1, \ldots, N_y \end{array} \tag{26}$$

or:

$$\min_{\Delta \boldsymbol{u}, \mu} \ \mu$$

subject to

$$\begin{array}{l} \boldsymbol{G} \Delta \boldsymbol{u} - \boldsymbol{\mu} \leq -\hat{\boldsymbol{y}}_l + \boldsymbol{w} \\ -\boldsymbol{G} \Delta \boldsymbol{u} - \boldsymbol{\mu} \leq \hat{\boldsymbol{y}}_l - \boldsymbol{w} \\ \boldsymbol{\mathcal{A}} \, \Delta \boldsymbol{u} \leq \boldsymbol{v} \end{array} \tag{27}$$

where $\boldsymbol{\mu} = [\, \mu \ \ \cdots \ \ \mu \,]^T \in \mathbb{R}^{(N_y - N_1 + 1)}$.

Thus, for the optimization problem based on the quadratic norm, the solution is delivered using a QP (*Quadratic Programming*) algorithm (Bazaraa and Shetty, 1979) or analytically in the unconstrained case. It must be remarked that the numerical solution used to solve the control law is not of the main interest in this paper. Thus, the control law was written as a QP problem in such a way that several methods can be used to solve it. There are examples of the method that reduces the Kuhn-Tucker conditions of the QP to a linear complementary problem: the dual programming and the gradient projection method of Rosen (Bazaraa and Shetty, 1979). Moreover, the work (Soeterboek, 1990) discusses an application of the latter in the context of MBPC algorithms. Rewriting the control law as a constrained least-squares problem (Golub and Van Loan, 1985) may result in a more robust solution to finding the optimal control law, which merits future research, especially if a particular structure of the matrices involved in the control law is explored.

For the infinite norm problem, a solution is obtained using an LP (*Linear Programming*) algorithm (Luenberger, 1984).

4. Stability Results

In this section, the stability of the closed loop system for the unconstrained PC-OSF with quadratic cost function and control horizon N_u equal to 1 is investigated. Assuming the unconstrained case, the solution to the minimization problem (23) in relation to $\Delta \boldsymbol{u}$ is

$$\Delta \boldsymbol{u} = -\frac{1}{2}\boldsymbol{Q}^{-1}\boldsymbol{f} \tag{28}$$

When N_u is equal to 1, $\boldsymbol{G}$ is the column vector, which is represented by $\boldsymbol{g}$, and $\Delta \boldsymbol{u}$ is the scalar $\Delta u(k/k)$. In this way, the control signal $\Delta u(k/k)$ is given by

$$\Delta u(k/k) = \left(\boldsymbol{g}^T\boldsymbol{g} + \lambda\right)^{-1}\boldsymbol{g}^T\left(\boldsymbol{w} - \hat{\boldsymbol{y}}_l\right) \tag{29}$$

From eqn. (15), it follows that the vector $\hat{\boldsymbol{y}}_l$ can be written as

$$\hat{\boldsymbol{y}}_l = \boldsymbol{y}_m + \left[\boldsymbol{c}^T(\boldsymbol{K}_j - \boldsymbol{I})\right]_{j=N_1,\dots,N_y}\Delta \boldsymbol{l}(k) \tag{30}$$

where $\boldsymbol{y}_m = [\,y_m(k)\ \cdots\ y_m(k)\,]^T$, and $y_m(k)$ is the measured output at the instant k;

$$\left[\boldsymbol{c}^T(\boldsymbol{K}_j - \boldsymbol{I})\right]_{j=N_1,\dots,N_y} = \begin{bmatrix} \boldsymbol{c}^T(\boldsymbol{K}_{N_1} - \boldsymbol{I}) \\ \boldsymbol{c}^T(\boldsymbol{K}_{N_1+1} - \boldsymbol{I}) \\ \vdots \\ \boldsymbol{c}^T(\boldsymbol{K}_{N_y} - \boldsymbol{I}) \end{bmatrix} \tag{31}$$

Substitution of (30) in (29) gives

$$\left(\boldsymbol{g}^T\boldsymbol{g} + \lambda\right)\Delta u(k/k) = \boldsymbol{g}^T\left(\boldsymbol{w} - \boldsymbol{y}_m - \left[\boldsymbol{c}^T(\boldsymbol{K}_j - \boldsymbol{I})\right]_{j=N_1,\dots,N_y}\Delta \boldsymbol{l}(k)\right) \tag{32}$$

or

$$\left(\boldsymbol{g}^T\boldsymbol{g}+\lambda\right)\Delta u(k/k)=\boldsymbol{g}^T\left(\boldsymbol{w}-\boldsymbol{y}_m-\left[\boldsymbol{c}^T(\boldsymbol{A}^j-\boldsymbol{I})\right]_{j=N_1,\ldots,N_y}\boldsymbol{l}(k)-\boldsymbol{g}u(k-1)\right) \quad (33)$$

Hence $u(k)$ is given by

$$u(k)=\alpha\left(\boldsymbol{g}^T\left(\boldsymbol{w}-\boldsymbol{y}_m-\left[\boldsymbol{c}^T(\boldsymbol{A}^j-\boldsymbol{I})\right]_{j=N_1,\ldots,N_y}\boldsymbol{l}(k)\right)+\lambda u(k-1)\right) \quad (34)$$

where

$$\alpha=\left(\boldsymbol{g}^T\boldsymbol{g}+\lambda\right)^{-1} \quad (35)$$

This equation can be rewritten as follows:

$$\begin{aligned} u(k) &= \alpha\left(-\boldsymbol{g}^T\left[\boldsymbol{c}^T\boldsymbol{A}^j\right]_{j=N_1,\ldots,N_y}\boldsymbol{l}(k)+\lambda u(k-1)\right) \\ &\quad +\alpha\boldsymbol{g}^T\boldsymbol{w}+\alpha\sum_{i=N_1}^{N_y} g_i\left(\boldsymbol{c}^T\boldsymbol{l}(k)-y_m(k)\right) \end{aligned} \quad (36)$$

and the closed-loop system is given by

$$\begin{cases} \boldsymbol{l}(k)= & \boldsymbol{A}\boldsymbol{l}(k-1)+\boldsymbol{b}u(k-1) \\ u(k)= & -\alpha\boldsymbol{g}^T\left[\boldsymbol{c}^T\boldsymbol{A}^{j+1}\right]_{j=N_1,\ldots,N_y}\boldsymbol{l}(k-1) \\ & -\alpha\left(\lambda-\boldsymbol{g}^T\left[\boldsymbol{c}^T\boldsymbol{A}^j\right]_{j=N_1,\ldots,N_y}\boldsymbol{b}\right) \\ & \times u(k-1)+\alpha\boldsymbol{g}^T\boldsymbol{w}+\alpha\sum_{i=N_1}^{N_y} g_i\left(\boldsymbol{c}^T\boldsymbol{l}(k)-y_m(k)\right) \end{cases} \quad (37)$$

In the case of a plant-model match, the measured output signal y_m is given by

$$y_m(k)=\boldsymbol{c}^T\boldsymbol{l}(k)+\xi(k) \quad (38)$$

where $\xi(k)$ is a measurement noise.

Thus, the closed loop equations (37) are given by

$$\begin{bmatrix} \boldsymbol{l}(k) \\ u(k/k) \end{bmatrix}=[\boldsymbol{\Phi}+\alpha\boldsymbol{\Gamma}]\begin{bmatrix} \boldsymbol{l}(k-1) \\ u(k-1) \end{bmatrix}+\begin{bmatrix} \boldsymbol{0} \\ \alpha \end{bmatrix}\left(\boldsymbol{g}^T\boldsymbol{w}+\sum_{i=N_1}^{N_y} g_i\xi(k)\right) \quad (39)$$

where

$$\boldsymbol{\Phi}=\begin{bmatrix} \boldsymbol{A} & \boldsymbol{b} \\ 0 & 0 \end{bmatrix} \quad (40)$$

and

$$\boldsymbol{\Gamma}=\begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ -\boldsymbol{g}^T\left[\boldsymbol{c}^T\boldsymbol{A}^{j+1}\right]_{j=N_1,\ldots,N_y} & \left(\lambda-\boldsymbol{g}^T\left[\boldsymbol{c}^T\boldsymbol{A}^j\right]_{j=N_1,\ldots,N_y}\boldsymbol{b}\right) \end{bmatrix} \quad (41)$$

Since the process model is open-loop stable, the matrix $\boldsymbol{\Phi}$ has its eigenvalues inside the unit circle. In (Elshafei *et al.*, 1994), it is shown that the stability of a closed-loop system having the same structure of system (39) can be assured by selecting a sufficiently small term $\alpha\boldsymbol{\Gamma}$. This result is given by the following theorem:

Theorem 1. (Elshafei *et al.*, 1994) *Let us consider the system*

$$z(k+1) = \boldsymbol{\Phi} z(k) \tag{42}$$

where

$$\exists\, \boldsymbol{P} > 0 \quad \forall\, \boldsymbol{Q} > 0 \tag{43}$$

such that

$$\boldsymbol{\Phi}^T \boldsymbol{P} \boldsymbol{\Phi} - \boldsymbol{P} = -\boldsymbol{Q} \tag{44}$$

The system given by

$$z(k+1) = (\boldsymbol{\Phi} + \delta) z(k) \tag{45}$$

is stable if

$$0 \leq ||\delta|| < -||\boldsymbol{\Phi}|| + \sqrt{||\boldsymbol{\Phi}||^2 + \frac{\lambda_{\min}(\boldsymbol{Q})}{||\boldsymbol{P}||}} \tag{46}$$

From eqn. (35) it follows that by increasing the value of N_y and/or λ, it is possible to decrease the value of the scalar α, in such a way that Theorem 1 is satisfied. Thus the stability of the closed loop system can be guaranteed by an adequate selection of controller parameters N_y and λ.

5. Simulation Example

In this section, a simulation example is presented to illustrate the behavior of the proposed Adaptive PC-OSF in a time-varying process. Its performance is compared to a predictive algorithm based on a structured model, the Adaptive GPC. The constraints are not taken into account in this example to highlight the performance of the control schemes.

The process is defined by the following transfer function:

$$y(s) = G(s)e^{-ds}u(s) \tag{47}$$

where

$$G(s) = \frac{1}{10s+1} + k\left\{\frac{1}{10s+1}\,\frac{-2s+1}{2s+1} - \frac{1}{10s+1}\right\} \tag{48}$$

The process dynamics change as follows: for $t < 60$, $d = 0$ and $k = 0$; for $60 \leq t < 120$, $d = 2$ and $k = 0$; and for $t \geq 120$, $d = 2$ and $k = 1$. The sampling time is 1 s.

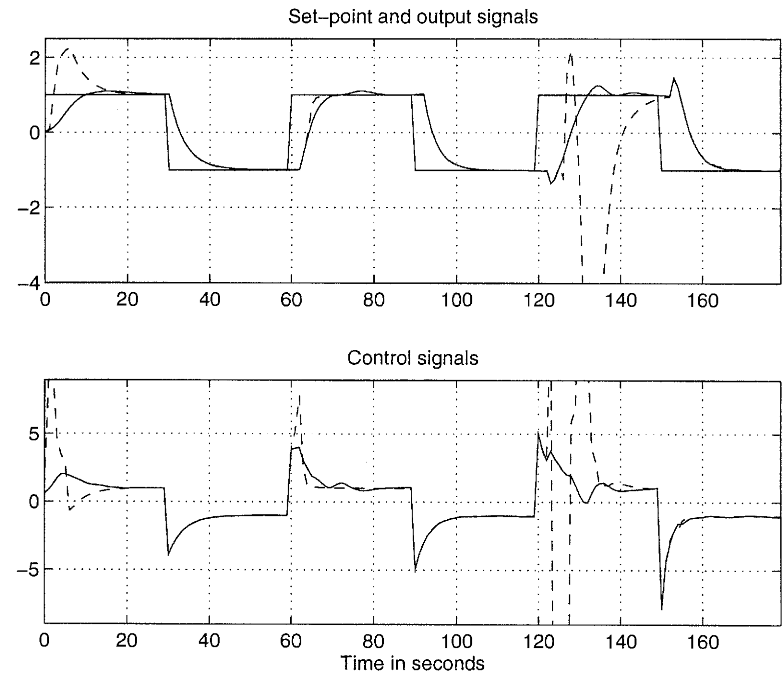

Fig. 1. Closed-loop performance for a time-varying process (dashed line: GPC, solid line: PC-OSF).

The orthonormal series model is defined as follows. The pole p is set at 0.6 ($p = 0.6$) to approximate the dynamics of the process and the number of functions is set at 10 ($n = 10$). The matrix $\boldsymbol{A}$ and the vector $\boldsymbol{b}$ are computed by using these two parameters and the matrix $\hat{\boldsymbol{c}}$ is initiated with $c_1 = 1$ and $c_i = 0$ for $i = 2, \dots, 10$.

The GPC controller of CARIMA model is given by

$$A(q^{-1})y(k) = B(q^{-1})u(k-1) + T(q^{-1})\xi(k)/\Delta \tag{49}$$

and is set by selecting the orders 2 and 3 for the polynomials $A(q^{-1})$ and $B(q^{-1})$, respectively, to allow for the identification of all the process dynamics; the polynomials $A(q^{-1})$ and $B(q^{-1})$ are initiated as $A(q^{-1}) = 1$ and $B(q^{-1}) = 1$. $T(q^{-1})$ is set at 1.

The OSF and CARIMA model parameters are identified by using a recursive least-squares estimator, with a forgetting factor of 0.97 and an initial covariance matrix of $100\boldsymbol{I}$. At each process change, i.e., at $t = 60$ and $t = 120$ seconds, a covariance matrix reset strategy is applied in the estimator algorithm, and the covariance matrix becomes $5\boldsymbol{I}$. The tuning parameters are $N_1 = 1$, $N_y = 10$, $N_u = 1$ and $\lambda = 0$, and the approach used for the control law is that shown in eqn. (22).

Figure 1 shows the input/output signals for the Adaptive GPC and Adaptive PC-OSF for this time-varying process. From these closed-loop performances, it can

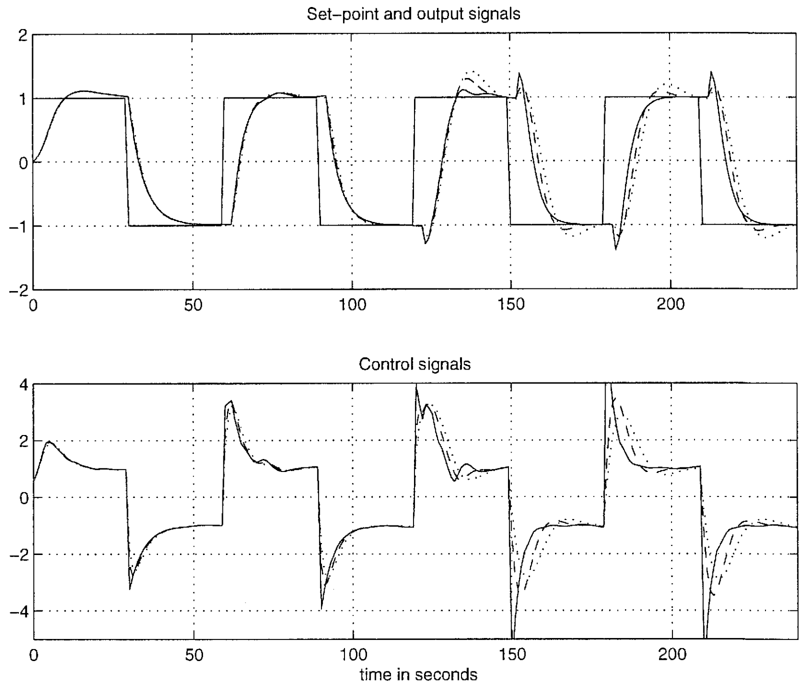

Fig. 2. Closed-loop performance for different values of the control weighting factor (solid line: $\lambda = 0$, dashed line: $\lambda = 0.5$, dotted line: $\lambda = 1$).

be seen that during the identification phase, the Adaptive PC-OSF provided more accurate set-point tracking than did the Adaptive GPC controller, although both the adaptive schemes resulted in a similar performance during the steady-state phase. An accurate selection of the order and time delay of the model is necessary to maintain closed-loop system stability for the GPC controller, whereas, for the CP-OSF controller, approximate knowledge of the process dynamics is sufficient. Thus, in the present example, orthonormal series function modelling requires less information about the process than would CARIMA modelling, and it results in a predictive controller with a more accurate closed-loop performance.

The influence of the control weighting factor on the closed-loop performance is shown in Fig. 2. It follows that an increase in the weighting factor λ tends to decrease the amplitude of the control signal, leading to a slower closed-loop behavior.

In the simulation presented in Fig. 1, the covariance matrix reset has been made just after the process changes, since the main objective was not to compare identification methods but the modelling strategies in the context of adaptive controllers. Thus, Fig. 3 illustrates the system behavior, in comparison with the case presented in Fig. 1, when: (a) the reset is made 20 s after the process changes; (b) there is no reset in the covariance matrix; and finally, (c) without covariance matrix reset and with

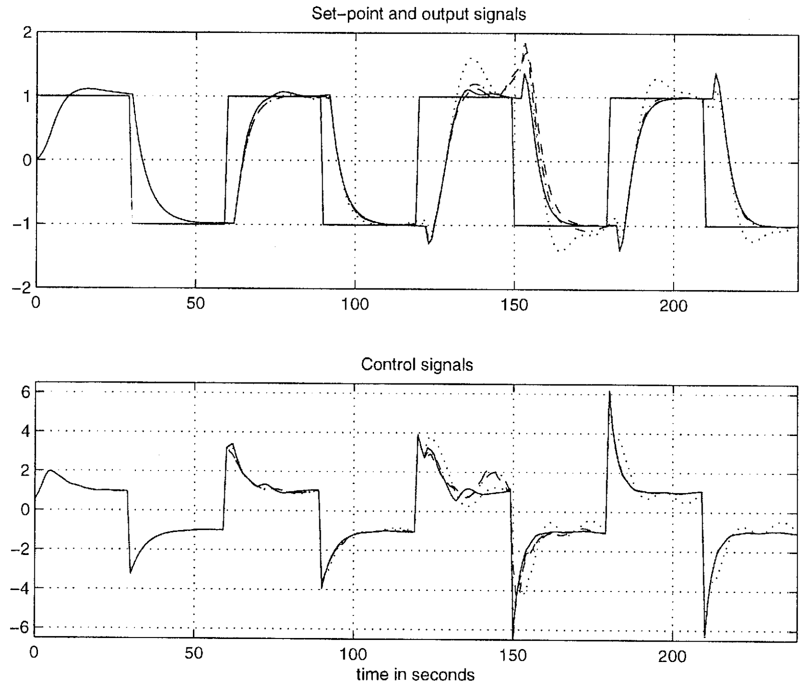

Fig. 3. Analysis of the covariance matrix reset (solid line: (a), dashed line: (b), dashdot line: (c), dotted line: (d)).

the forgetting factor equal to 1. It can be noticed that the performance slightly deteriorated in the cases where the covariance matrix was not made just after the process changes. However, cases (a) and (b) present a quite similar behavior. Moreover, the behavior of the system with an increase, in relation to situation (b), in the forgetting factor value, i.e. case (c), represents the worst performance of the RLS algorithm.

Now, the use of an infinite norm in the cost function instead of the quadratic cost function is illustrated. In Fig. 4, a comparison between infinite and quadratic cost functions is presented, with the use of process (47) when $d = 2,\ k = 0$ and assuming the plant-model match case. The infinite norm does not attempt to minimize the errors at all future time horizons, but only at the instant where the error is maximum and, in this example, this characteristic results in a predictive controller with a slightly faster time response to set-point changes when compared with the one based on a quadratic cost function.

6. Conclusion

An approach to a constrained adaptive predictive controller, based on an unstructured representation of a process using orthonormal series functions, has been presented.

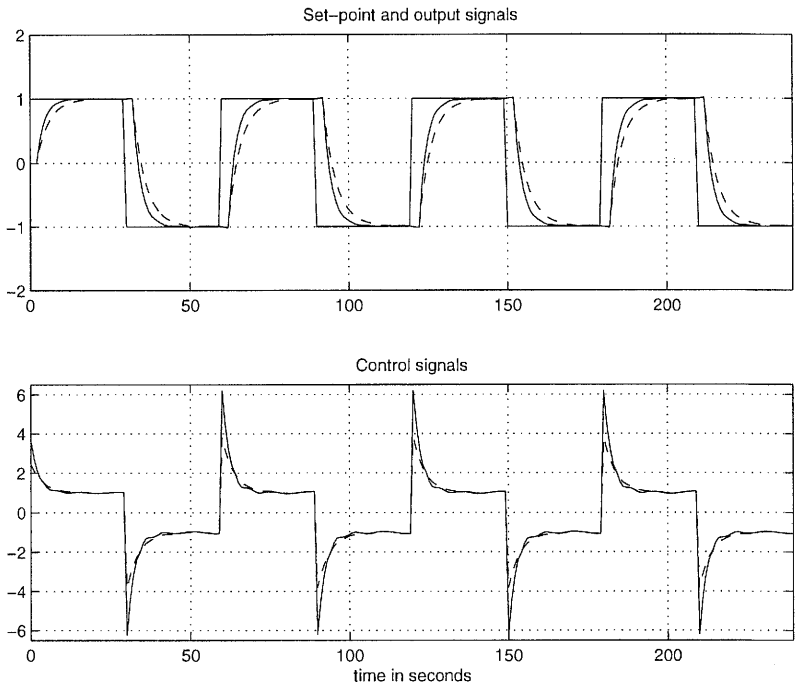

Fig. 4. Comparison between infinite and quadratic cost functions (dashed line—infinite norm cost function, solid line—quadratic cost function).

Two types of norms to derive the cost function were taken into account: the 2-norm (quadratic cost function) and infinite-norm. In such a predictive controller, some approximative information about the open-loop dynamics of the process are assumed for the on-line parameter estimation. The control law can be derived for the process both with and without integral action.

When compared with other adaptive predictive control strategies, the algorithm described here has some advantages. First, fewer unknown parameters are involved in the identification than in an impulse response model, which improves the quality of estimation. Second, there is no need for exact specification of the order and time delay in the model, which constitutes an advantage in relation to predictive controllers with structured models when exact information about these parameters is unavailable. Moreover, the closed-loop system is stable by an appropriate selection of the prediction horizon and/or the control signal weighting.

Acknowledgments

This research was supported the by CNPq and FAPESP (Brazil) and Ministère des Affaires Étrangères (France).

References

Bazaraa M.S. and Shetty C.M. (1979): *Nonlinear Programming.* — New York: J. Wiley.

Broome P.W. (1965): *Discrete orthonormal sequences.* — J. ACM, Vol.12, No.2, pp.151–168.

Clarke D.W. (1994): *Advances in Model Based Predictive Control.* — Oxford, England: Oxford University.

Clarke D.W. and Mohtadi C. (1989): *Proprieties of generalized predictive control.* — Automatica, Vol.25, No.6, pp.859–875.

Dumont G. (1998): *From lust to brainwave: Fifteen years of persistent excitation.* — Proc. IFAC Workshop *Adaptive Systems in Control and Signal Processing*, Glasgow (UK), pp.472–479.

Elshafei A., Dumont G.A. and Elnaggar A. (1994): *Adaptive GPC based on Laguerre-filters modelling.* — Automatica, Vol.30, No.12, pp.1913–1920.

Finn C., Wahlberg B. and Ydstie B.E. (1993): *Constrained predictive control using orthogonal expansions.* — AIChE J., Vol.39, No.11, pp.1810–1826.

Fu Y. and Dumont G.A. (1993): *An optimum time scale for discrete Laguerre network.* — IEEE Trans. Automat. Contr., Vol.38, No.6, pp.934–938.

Golub G.H. and Van Loan C.F. (1985): *Matrix Computation.* — baltimore, Maryland: Johns Hopkins University Press.

Gunnarsson S. and Wahlberg B. (1991): *Some asymptotic results in recursive identification using Laguerre models.* — Int. J. Adapt. Contr. Signal Process., Vol.5, pp.313–333.

Heuberger P.S.C., Van den Hof P.M.J. and Bosgra O.H. (1995): *A generalized orthonormal basis for linear dynamical systems.* — IEEE Trans. Automat. Contr., Vol.40, No.3, pp.451–465.

Latawiec K.J. (1998): *Towards robust adaptive least-square parameter estimation with internal feedback.* — Proc. IFAC Workshop *Adaptive Systems in Control and Signal Processing*, Glasgow (UK), pp.442–447.

Lindskog P. (1996): *Methods, Algorithms and Tools for System Identification Based on Prior Knowledge.* — Ph.D. Thesis, Linkoping University, Sweden.

Lindskog P. and Wahlberg B. (1993): *Applications of Kautz model in system identification.* — Prep. 12th IFAC World Congress, Vol.5, pp.197–205, Sydney, Australia.

Ljung L. (1987): *System Identification: Theory for the User.* — New Jersey, USA: Prenctice Hall Inc.

Luenberger D.G. (1984): *Indroduction to Linear and Nonlinear Programming, 2nd Ed..* — Reading, Mass.: Addison-Wesley.

Masnadi-Shirazi M.A. and Ahmed N. (1991): *Optimum Laguerre networks for a class of discrete-time systems.* — IEEE Trans. Signal Process., Vol.39, No.9, pp.2104–2108.

Ninness B. and Gustafsson F. (1994): *A unifying construction of orthonormal bases for system identification.* — Tech. Rep. No.EE9433, DECE, University of Newcastle, Australia.

Ninness B. and Gustafsson F. (1995): *Orthonormal bases for system identification.* — Proc. 3rd *European Control Conference*, Vol.1, Roma (Italy), pp.13–18.

Oliveira G.H.C., Favier G., Dumont G. and Amaral W.C. (1996a): *Robust predictive controller based on Laguerre filters modelling.* — Proc. 13th IFAC World Congress, San Francisco (USA), Vol.G, pp.375–380,

Oliveira G.H.C., Lopez J.F., Favier G. and Amaral W.C. (1996b): *Review of predictive control methods using Laguerre functions.* — Proc. IEEE/IMACS Multiconf. *Computaional Engineering in Systems Applications CESA*, Symp. *Control Optimization and Supervision*, Lille (France), Vol.1, pp.304–309.

Oliveira G.H.C., Amaral W.C. and Favier G. (1997): *Constrained adaptive MBPC using orthonormal series functions.* — Proc. 4th Int. Symp. *Methods and Models in Automation and Robotics*, Międzyzdroje (Poland), pp.477–482.

Oliveira G.H.C., Favier G., Dumont G. and Amaral W.C. (1998): *Uncertainties identification using Laguerre models with application to paper machine headbox.* — Proc. IEEE/IMACS Multiconf. *Computational Engineering in Systems Applications CESA*, Symp. *Control Optimization and Supervision*, Hammamet (Tunisia), Vol.1, pp.329–334.

Olivier P.D. (1994): *On-line system identification using Laguerre series.* — IEE Proc. D, Vol.141, No.4, pp.249–254.

Rohrs C.E., Valavani L., Athans M. and Stein G. (1982): *Robustness of adaptive control algorithms in the presence of unmodeled dynamics.* — Proc. 21th Conf. *Decision and Control*, Orlando, Florida, pp.3–11.

Shook D.S., Mohtadi C. and Shah S.L. (1991): *Identification for long-range predictive control.* — IEE Proc. D, Vol.138, No.1, pp.75–84.

Soeterboek R. (1990): *Predictive Control: A Unified Approach.* — Ph.D. Thesis, Technische Universiteit Delft, Holland.

Van den Hof P.M.J., Heuberger P.S.C. and Bokor J. (1995): *System identification with generalized orthonormal basis functions.* — Automatica, Vol.31, No.12, pp.1821–1834.

Wahlberg B. (1991a): *System identification using Laguerre models.* — IEEE Trans. Automat. Contr., Vol.36, No.5, pp.551–562.

Wahlberg B. (1991b): *Identification of resonant systems using Kautz filters.* — Proc. 30th Conf. *Decision and Control*, Brighton, England, pp.2005–2010.

Wahlberg B. and Makila P.M. (1996): *Approximation of stable linear dynamical systems using Laguerre and Kautz functions.* — Automatica, Vol.32, No.5, pp.693–708.

Yoon T. and Clarke D.W. (1994): *Adaptive predictive control of the benchmark plant.* — Automatica, Vol.30, No.4, pp.621–628.

Zervos C.C. and Dumont G.A. (1988): *Deterministic adaptive control based on Laguerre series representation.* — Int. J. Contr., Vol.48, No.6, pp.2333–2359.