2nd Symposium on System, Structure and Control, Oaxaca, 2004

POLYNOMIAL APPROACH FOR FILTERING AND IDENTIFICATION OF A CLASS OF UNCERTAIN SYSTEMS [1]

D. Di Martino * **A. Germani** ** **C. Manes** ** **P. Palumbo** **

* *Gran Sasso National Laboratory, LNGS-INFN (National Institute of Nuclear Physics), SS 17/bis, Km 18+910, 67010 Assergi (L'Aquila), Italy,*
domenico.dimartino@lngs.infn.it
** *Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", IASI-CNR (National Research Council of Italy), Viale Manzoni 30, 00185 Roma, Italy,*
germani@ing.univaq.it, manes@ing.univaq.it, palumbo@iasi.rm.cnr.it

Abstract: this paper considers the filtering and identification problems for a class of discrete-time uncertain stochastic systems that admit a finite number of linear working modes. It is shown here that this class of uncertain systems can be modeled by using a suitably defined extended system, whose state evolves according to a bilinear model. A polynomial filtering algorithm is derived for such extended system, which readily provides the polynomial estimates of both the original state and the working mode. Simulations show the effectiveness of the proposed approach and the improvements with respect to standard linear filtering algorithms.

Keywords: Polynomial filtering, Stochastic Systems, Bilinear Systems, Uncertain systems

1. INTRODUCTION

This work considers discrete-time stochastic linear systems described by equations of the type:

$$\begin{aligned} x(k+1) &= A(\mu)x(k)+B(\mu)u(k)+F(\mu)N(k), \\ y(k) &= C(\mu)x(k)+D(\mu)u(k)+G(\mu)N(k), \end{aligned} \quad (1)$$

$k \in \mathbb{Z}^+$, where $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^p$ is the known input, $y(k) \in \mathbb{R}^q$ is the measured output, $N(k) \in \mathbb{R}^b$ is the "noise", a sequence of zero-mean independent random vectors, not necessarily Gaussian. All system matrices in (1) depend on an unknown parameter μ taking values on a finite set $\mathcal{W}$. Without loss of generality, we assume here that μ takes values on the set of the first m integers, i.e. $\mathcal{W} = \{1, \dots, m\}$. Stated in other words, the system is characterized by m linear working modes. The problem considered in this paper is to estimate both the system state and the current working mode.

A great deal of literature treats the state estimation problem for systems of the type (1) when the parameter μ undergoes Markov transitions (Germani *et al.*, 2003; Costa, 1994; Blom and Bar-Shalom, 1988). In this work the unknown parameter is assumed constant (at least over "long" time intervals). A minimum variance polynomial filter is presented here to solve both the parameter and state estimation problems. The polynomial approach has led to important results in the field of suboptimal filtering of non Gaussian linear (Carravetta *et al.*, 1996) and bilinear systems

[1] Work partially supported by MIUR (Italian Ministry for Education and Research), INFN and by EU project HYBRIDGE.

(Carravetta *et al.*, 1997). Recently, polynomial filters have also been studied in the framework of descriptor (Germani *et al.*, 2004) and Markov switching systems (Germani *et al.*, 2003). The key-point in this paper is the construction of an extended system, whose state contains a suitable parameterization of the unknown variable μ and its Kronecker products with the Kronecker powers of the original state, up to a chosen degree ν. The extended system has the structure of a bilinear model, i.e. a linear system driven by multiplicative noise. The output of the extended system is made of the original output vector and of its Kronecker powers up to the chosen degree ν. The best polynomial filter for the original system is derived through the computation of the best linear filter for the extended system. Although the polynomial filter is derived here specifically for systems with unknown and constant working mode, simulation results show the good performances of the filter also when the system undergoes rare switching.

2. A MODEL FOR THE UNCERTAIN SYSTEM

This section presents an alternative representation for the uncertain system (1). For the model derivation it is useful to regard the constant parameter μ as governed by the trivial difference equation

$$\mu(k+1) = \mu(k), \quad \mu(0) = \mu_0. \tag{2}$$

Consider the natural basis in $\mathbb{R}^m$, denoted $\mathcal{E}_m = \{e_1, e_2, \cdots, e_m\}$. Defining the matrix

$$\widetilde{A} = \big[A(1)\ A(2)\ \cdots\ A(m)\big] \tag{3}$$

the following identities hold

$$A(i) = \widetilde{A}\big(e_i \otimes I_n\big), \quad i = 1, \ldots, m, \tag{4}$$

as it is easy verified from

$$\widetilde{A}\big(e_i \otimes I_n\big) = \big[A(1) \cdots A(i) \cdots A(m)\big] \begin{bmatrix} O_{(i-1)n\times n} \\ I_n \\ O_{(m-i)n\times n} \end{bmatrix}, \tag{5}$$

where the symbol $\otimes$ denotes the standard matrix Kronecker product. Throughout the paper superscripts in square brackets denote Kronecker powers, defined for a given matrix H by:

$$H^{[0]} = 1, \quad H^{[i]} = H \otimes H^{[i-1]}, \quad i \geq 1. \tag{6}$$

(see (Bellman, 1970) for more details, see (Carravetta *et al.*, 1997) for a quick survey on the Kronecker product and its properties).

Instead of the integer $\mu \in \mathcal{W}$ a vector $\vartheta \in \mathcal{E}_m$ can be used to parameterize the m working modes of system (1), suitably exploiting the bijection $\mathcal{E}_m \leftrightarrow \mathcal{W}$. Defining $\vartheta = e_\mu$, all matrices in equation (1) can be written as follows

$$\begin{aligned} A(\mu) &= \widetilde{A}\big(\vartheta \otimes I_n\big), \quad B(\mu) = \widetilde{B}\big(\vartheta \otimes I_p\big), \\ C(\mu) &= \widetilde{C}\big(\vartheta \otimes I_n\big), \quad D(\mu) = \widetilde{D}\big(\vartheta \otimes I_p\big), \\ F(\mu) &= \widetilde{F}\big(\vartheta \otimes I_b\big), \quad G(\mu) = \widetilde{G}\big(\vartheta \otimes I_b\big), \end{aligned} \tag{7}$$

where all matrices $\widetilde{B}, \widetilde{C}, \widetilde{D}, \widetilde{F}, \widetilde{G}$ are defined as $\widetilde{A}$ in (3). From (2), the sequence $\vartheta(k) = e_{\mu(k)}$ is governed by equations

$$\vartheta(k+1) = \vartheta(k), \quad \vartheta(0) = \vartheta_0. \tag{8}$$

Since $\vartheta(k) \in \mathcal{E}_m$, it follows that

$$\vartheta^{[2]}(k) = E_2 \vartheta(k), \text{ with } E_2 = \big[e_1^{[2]} \cdots e_m^{[2]}\big]. \tag{9}$$

Proposition 1. System (1) admits the representation:

$$\begin{aligned} x(k+1) &= \widetilde{A}\big(\vartheta(k) \otimes x(k)\big) + \overline{B}(k)\vartheta(k) \\ &\quad + \widetilde{F}\big(\vartheta(k) \otimes N(k)\big), \\ \vartheta(k+1) &= \vartheta(k), \\ y(k) &= \widetilde{C}\big(\vartheta(k) \otimes x(k)\big) + \overline{D}(k)\vartheta(k) \\ &\quad + \widetilde{G}\big(\vartheta(k) \otimes N(k)\big), \end{aligned} \tag{10}$$

where the time-varying matrices $\overline{B}(k)$, $\overline{D}(k)$ depend on the known input $u(k)$ as follows:

$$\overline{B}(k) = \widetilde{B}\big(I_m \otimes u(k)\big), \quad \overline{D}(k) = \widetilde{D}\big(I_m \otimes u(k)\big). \tag{11}$$

PROOF. Using identities (7) the state and output equations of system (1) can be put in the form:

$$\begin{aligned} x(k+1) &= \widetilde{A}\big(\vartheta(k) \otimes I_n\big)x(k) + \widetilde{B}\big(\vartheta(k) \otimes I_p\big)u(k) \\ &\quad + \widetilde{F}\big(\vartheta(k) \otimes I_b\big)N(k), \\ y(k) &= \widetilde{C}\big(\vartheta(k) \otimes I_n\big)x(k) + \widetilde{D}\big(\vartheta(k) \otimes I_p\big)u(k) \\ &\quad + \widetilde{G}\big(\vartheta(k) \otimes I_b\big)N(k). \end{aligned} \tag{12}$$

According to the Kronecker product properties:

$$\begin{aligned} \big(\vartheta(k) \otimes I_n\big)x(k) &= \big(\vartheta(k) \otimes I_n\big) \cdot \big(1 \otimes x(k)\big) \\ &= \big(\vartheta(k) \cdot 1\big) \otimes \big(I_n \cdot x(k)\big) = \vartheta(k) \otimes x(k), \\ \big(\vartheta(k) \otimes I_p\big)u(k) &= \big(\vartheta(k) \otimes I_p\big) \cdot \big(1 \otimes u(k)\big) \\ &= \vartheta(k) \otimes u(k) = \big(I_m \cdot \vartheta(k)\big) \otimes \big(u(k) \cdot 1\big) \\ &= \big(I_m \otimes u(k)\big) \cdot \big(\vartheta(k) \otimes 1\big) = \big(I_m \otimes u(k)\big)\vartheta(k), \end{aligned} \tag{13}$$

so that (10) and (11) are easily obtained.

3. THE POLYNOMIAL FILTER

It is well known that the optimal solution to the minimum variance filtering problem is given by the expectation of the state conditioned to all the measurements up to the current time, that is the projection of the state onto the linear space of all the Borel functions of the measurements:

$$\hat{x}(k) = \mathbb{E}\big[x(k) | \sigma\big(Y_k\big)\big] = \Pi\big[x(k) | \mathcal{B}(Y_k)\big], \tag{14}$$

where $Y_k = \big[y^T(0)\ \cdots\ y^T(k)\big]^T$. In the Gaussian case the conditional expectation is a linear transformation of the measurements, recursively implemented by the Kalman filter. In the non Gaussian case, when the

conditional expectation is difficult to compute, a suboptimal estimation approach can be followed. By definition, suboptimal polynomial estimates are optimal in the Hilbert space of all polynomial transformations of measurements (Carravetta *et al.*, 1996; Carravetta *et al.*, 1997). Choosing an integer ν and assuming that, for all $h \in \mathbb{Z}^+$,

$$\mathbb{E}\big[\|y^{[i]}(h)\|^2\big] < \infty,\ i = 1, \ldots, 2\nu, \tag{15}$$

the Hilbert space of ν-degree polynomial transformations of the output sequence can be defined as follows:

$$L(Y_k^\nu) = \text{span}\big\{1,\ Y^\nu(0),\ \cdots,\ Y^\nu(k)\big\}, \tag{16}$$

$$\text{with}\ \ Y_k^\nu = \begin{bmatrix} Y^\nu(0) \\ \vdots \\ Y^\nu(k) \end{bmatrix},\ \ Y^\nu(h) = \begin{bmatrix} y(h) \\ \vdots \\ y^{[\nu]}(h) \end{bmatrix}. \tag{17}$$

The optimal (min. error variance) state and parameter estimates in $L(Y_k^\nu)$ are given by the projections:

$$\begin{aligned} \hat{x}_\nu(k) &= \Pi\big[x(k)|L(Y_k^\nu)\big], \\ \hat{\vartheta}_\nu(k) &= \Pi\big[\vartheta(k)|L(Y_k^\nu)\big]. \end{aligned} \tag{18}$$

In order to ensure that all the moments in (15) are finite, the following assumptions are needed:
1) the noise variable $N(k)$ has finite moments up to degree 2ν

$$\mathbb{E}\big[N^{[j]}(k)\big] = \xi_j < \infty, \quad 1 \le j \le 2\nu, \tag{19}$$

(note that, being $N(k)$ white, it is $\xi_1 = 0$).
2) The initial state $x(0) = x_0$, independent of the noise sequence, has finite moments up to degree 2ν:

$$\mathbb{E}\big[x_0^{[j]}\big] = \zeta_j < \infty, \quad 1 \le j \le 2\nu. \tag{20}$$

Assumptions 1) and 2) guarantee that the *polynomial extended output sequence* $Y^\nu(k)$, defined in (17), has bounded mean and covariance. Consider now the extended state sequence $X^\nu(k)$ defined as

$$X^\nu(k) = \begin{bmatrix} X_0(k) \\ \vdots \\ X_\nu(k) \end{bmatrix}, \quad X_i(k) = \vartheta(k) \otimes x^{[i]}(k), \tag{21}$$

(note that $X_0(k) = \vartheta(k)$). In the following it will be shown that $X^\nu(k)$ and $Y^\nu(k)$ admit a stochastic bilinear generation model of the type

$$\begin{aligned} X^\nu(k+1) &= \mathbf{A}^\nu(k)X^\nu(k) + \mathbf{F}^\nu(N(k), X^\nu(k)), \\ Y^\nu(k) &= \mathbf{C}^\nu(k)X^\nu(k) + \mathbf{G}^\nu(N(k), X^\nu(k)), \end{aligned} \tag{22}$$

where $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ are suitably defined deterministic matrices, while $\mathbf{F}^\nu(N(k), X^\nu(k))$ and $\mathbf{G}^\nu(N(k), X^\nu(k))$ are terms in which noise terms multiplies the extended state. The structure of matrices $\mathbf{A}^\nu$ and $\mathbf{C}^\nu$ and the properties of the noise sequences $\mathcal{F}(k) = \mathbf{F}^\nu(N(k), X^\nu(k))$ and $\mathcal{G}(k) = \mathbf{G}^\nu(N(k), X^\nu(k))$ will be presented in lemmas 5, 6 and 8.

The best linear estimate of $X^\nu(k)$ is the projection $\widehat{X}^\nu(k) = \Pi\big[X^\nu(k)|L(Y_k^\nu)\big]$. Since $\mathcal{F}(k)$ and $\mathcal{G}(k)$ are sequences of zero-mean, uncorrelated random vectors, $\widehat{X}^\nu(k)$ can be recursively computed using the Kalman filter applied to system (22).

Theorem 2. The optimal ν-degree polynomial estimate of the state $x(k)$ of system (1) and of the unknown vector $\vartheta(k)$ are given by:

$$\begin{aligned} \hat{x}_\nu(k) &= \mathcal{M}_n\widehat{X}^\nu(k) = \mathcal{M}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big], \\ \hat{\vartheta}_\nu(k) &= \mathcal{T}_n\widehat{X}^\nu(k) = \mathcal{T}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big], \end{aligned} \tag{23}$$

where:

$$\begin{aligned} \mathcal{M}_n &= \big[O_{n\times m} \ \ \mathcal{M} \ \ O_{n\times m(n^2+\cdots+n^\nu)}\big], \\ \mathcal{T}_n &= \big[I_m \ \ O_{m\times m(n+\cdots+n^\nu)}\big], \end{aligned} \tag{24}$$

with $\mathcal{M} = \big[I_n \cdots I_n\big] \in \mathbb{R}^{n\times mn}$.

PROOF. The proof is easily obtained noting that $x(k)$ and $\vartheta(k)$ are both linear transformations of the extended state $X^\nu(k)$:

$$\begin{aligned} x(k) &= \mathcal{M}\big(\vartheta(k) \otimes x(k)\big) = \mathcal{M}X_1(k) = \mathcal{M}_nX^\nu(k) \\ \vartheta(k) &= \mathcal{T}_nX^\nu(k), \end{aligned} \tag{25}$$

so that the polynomial minimum variance state estimates in (23) are:

$$\begin{aligned} \hat{x}_\nu(k) &= \Pi\big[x(k)|L(Y_k^\nu)\big] = \Pi\big[\mathcal{M}_nX^\nu(k)|L(Y_k^\nu)\big] \\ &= \mathcal{M}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big] = \mathcal{M}_n\widehat{X}^\nu(k), \\ \hat{\vartheta}_\nu(k) &= \Pi\big[\vartheta(k)|L(Y_k^\nu)\big] = \Pi\big[\mathcal{T}_nX^\nu(k)|L(Y_k^\nu)\big] \\ &= \mathcal{T}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big] = \mathcal{T}_n\widehat{X}^\nu(k). \end{aligned} \tag{26}$$

Remark 3. The covariance of the estimation error $x(k) - \hat{x}_\nu(k)$ can be extracted from the covariance of the estimation error of the extended state as follows:

$$\text{Cov}\big(x(k) - \hat{x}_\nu(k)\big) = \mathcal{M}_n\text{Cov}\big(X^\nu(k) - \widehat{X}^\nu(k)\big)\mathcal{M}_n^T. \tag{27}$$

Remark 4. Since in general $\widehat{\theta}_\nu(k) \notin \mathcal{E}_m$, a strategy for the estimation of the mode $\mu(k)$ is to choose among the elements of $\mathcal{E}_m$ the closest one to the estimate $\widehat{\theta}_\nu(k)$, according to the L_∞-norm:

$$\hat{\mu}(k):\ \|e_{\hat{\mu}(k)} - \widehat{\theta}_\nu(k)\|_\infty \le \|e_j - \widehat{\theta}_\nu(k)\|_\infty; \tag{28}$$

for $j = 1, \ldots, m$. The motivation for this strategy is that the choice (28), when applied to the conditional expectation of $\vartheta(k)$, provides the Maximum Likelihood Estimate of $\mu(k)$. This happens because the components of $\widehat{\vartheta}(k) = \mathbb{E}\{\vartheta(k)|\sigma(Y_k)\}$ coincide with the conditional distribution of $\vartheta(k)$.

The following lemmas give some insights into the structure and properties of the model (22). All the results presented exploit the fact that, according to definition (21) and to identity (9), $\forall i, j, h \in \mathbb{Z}^+$:

$$X_j^{[h]} = \Theta_n^{h,j}X_{jh}, \quad X_i \otimes X_j = \Xi_{i,j}X_{i+j}, \tag{29}$$

where $\Theta_n^{h,j}$ and $\Xi_{i,j}$ are the matrices defined by:

$$\begin{aligned}\Theta_n^{h+1,j} &= \big(\Theta_n^{h,j} \otimes I_{mn^j}\big)\big(I_m \otimes C^T_{mn^j,n^{jh}}\big) \\ &\quad \cdot \big(E_2 \otimes I_{n^{j(h+1)}}\big), \\ \Theta_n^{0,j} &= [1 \cdots 1] \in I\!R^{1\times m}, \\ \Xi_{i,j} &= \big(I_m \otimes C^T_{mn^j,n^i}\big)(E_2 \otimes I_{n^{i+j}}),\end{aligned} \tag{30}$$

with $C_{a,b}$ suitably dimensioned commutation matrices for the Kronecker product (Carravetta *et al.*, 1997), and E_2 as in (9) (see (Germani *et al.*, 2003) for more details).

Lemma 5. The iterative equation of the component $X_j(k)$ defined in (21) is:

$$\begin{aligned}X_j(k+1) &= \sum_{t_1=0}^{j} \mathbf{A}_{j,t_1}(k)X_{t_1}(k) + \mathcal{F}_j(k), \\ \mathcal{F}_j(k) &= \sum_{t_1=0}^{j} S^j_{t_1}(k)X_{t_1}(k),\end{aligned} \tag{31}$$

where $\mathbf{A}_{j,t_1}(k)$, $S^j_{t_1}(k)$ are the following sequences of deterministic and random matrices:

$$\mathbf{A}_{j,t_1}(k) = \big(I_m \otimes J^j_{t_1}(k)\big)\Xi_{0,t_1}, \tag{32}$$

$$S^j_{t_1}(k) = \big(I_m \otimes \mathcal{L}^j_{t_1}(k)\big)\Xi_{0,t_1}, \tag{33}$$

with:

$$J^j_{t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} L^j_t(k)\Big(I_{mn^{t_1}} \otimes \xi_{t_3}(k)\Big), \tag{34}$$

$$\mathcal{L}^j_{t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} L^j_t(k)\Big(I_{mn^{t_1}} \otimes \big(N^{[t_3]}(k) - \xi_{t_3}(k)\big)\Big), \tag{35}$$

$$L^j_t(k) = M^j_t\Big(\widetilde{A}^{[t_1]} \otimes \overline{B}^{[t_2]}(k) \otimes \widetilde{F}^{[t_3]}\Big)K^j_t, \tag{36}$$

$$K^j_t = \big(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\big)\big(I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}\big). \tag{37}$$

M^j_t in (36) are the matrix coefficients for the Kronecker power expansion (Carravetta *et al.*, 1997), $t = \big(t_1, t_2, t_3\big)^T$ is a multi-index in $(\mathbb{Z}^+)^3$ and $\mathcal{R}_j = \big\{t \in (\mathbb{Z}^+)^3 : \ t_1 + t_2 + t_3 = j\big\}$. Moreover $\mathcal{F}(k) = \big[\mathcal{F}_0(k)^T \ \cdots \ \mathcal{F}_\nu(k)^T\big]^T$ is a sequence of zero-mean uncorrelated random vectors, whose covariance matrices $\Psi^{\mathcal{F}}_{j,i}(k) = I\!E\big[\mathcal{F}_j(k)\mathcal{F}_i(k)^T\big]$ are given by:

$$\begin{aligned}\Psi^{\mathcal{F}}_{j,i}(k) = \sum_{j_1=0}^{j}\sum_{i_1=0}^{i} \mathrm{st}^{-1}_{mn^{j_1},mn^{i_1}}\Big(&\Phi^{S,i,j}_{r_1,t_1}(k) \\ &\cdot \Xi_{r_1,t_1} I\!E\big[X_{r_1+t_1}(k)\big]\Big),\end{aligned} \tag{38}$$

with st^{-1} the inverse of the stack operator (Carravetta *et al.*, 1997) and:

$$\begin{aligned}\Phi^{S,i,j}_{r_1,t_1}(k) &= I\!E\big[S^i_{r_1}(k) \otimes S^j_{t_1}(k)\big] \\ &= \big(I_m \otimes C^T_{mn^j,n^i}\big)\big(I_{m^2} \otimes \Phi^{\mathcal{L},j,i}_{t_1,r_1}(k)\big) \\ &\quad \cdot \big(I_m \otimes C_{m^2n^{t_1},mn^{r_1}}\big)\big(\Xi_{0,r_1} \otimes \Xi_{0,t_1}\big),\end{aligned} \tag{39}$$

where:

$$\begin{aligned}\Phi^{\mathcal{L},j,i}_{t_1,r_1}(k) &= I\!E\big[\mathcal{L}^j_{t_1}(k) \otimes \mathcal{L}^i_{r_1}(k)\big] \\ &= \sum_{t_2,t_3}^{t\in\mathcal{R}_j}\sum_{r_2,r_3}^{r\in\mathcal{R}_i} \big(L^j_t(k) \otimes L^i_r(k)\big)\big(I_{mn^{t_1}} \otimes C^T_{mn^{r_1}b^{r_3},b^{t_3}}\big) \\ &\quad \cdot \Big(I_{m^2n^{t_1+r_1}} \otimes \big(\xi_{r_3+t_3}(k) - \xi_{r_3}(k) \otimes \xi_{t_3}(k)\big)\Big) \\ &\quad \cdot \big(I_{mn^{t_1}} \otimes C_{mn^{r_1},1}\big).\end{aligned} \tag{40}$$

PROOF. The proof is a straightforward consequence of Lemma 3.2 in (Germani *et al.*, 2003). In that framework system (1) is a switching system, and $\mu(k)$ is a Markov chain with known transition probability matrix. The iterative equation (31) easily comes by taking into account that in the present case, the unknown parameter does not switch so that, by consequence, the transition probability matrix is necessarily the identity matrix. The fact that the extended noise $\big\{\mathcal{F}(k)\big\}$ is a sequence of zero-mean uncorrelated random vectors, comes taking into account that $S^j_{t_1}(k)$ and $S^i_{s_1}(h)$ are zero-mean and uncorrelated for any j, i, t_1, s_1 and for any $k \neq h$ and, moreover, $S^j_{t_1}(k)$ is independent of $X_{s_1}(k)$.

Lemma 6. The equations for the Kronecker powers of the measurements defined in (17) are:

$$\begin{aligned}y^{[j]}(k) &= \sum_{t_1=0}^{j} \mathbf{C}_{j,t_1}(k)X_{t_1}(k) + \mathcal{G}_j(k), \\ \mathcal{G}_j(k) &= \sum_{t_1=0}^{j} \mathcal{T}^j_{t_1}(k)X_{t_1}(k),\end{aligned} \tag{41}$$

where $\mathbf{C}_{j,t_1}(k)$, $\mathcal{T}^j_{t_1}(k)$ are the following sequences of deterministic and random matrices:

$$\mathbf{C}_{j,t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} T^j_t(k)\Big(I_{mn^{t_1}} \otimes \xi_{t_3}(k)\Big), \tag{42}$$

$$\mathcal{T}^j_{t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} T^j_t(k)\Big(I_{mn^{t_1}} \otimes \big(N^{[t_3]}(k) - \xi_{t_3}(k)\big)\Big), \tag{43}$$

$$T^j_t(k) = M^j_t\Big(\widetilde{C}^{[t_1]} \otimes \overline{D}^{[t_2]}(k) \otimes \widetilde{G}^{[t_3]}\Big)K^j_t, \tag{44}$$

and K^j_t as in (37). $\mathcal{G}(k) = \big[\mathcal{G}_1(k)^T \ldots \ \mathcal{G}_\nu(k)^T\big]^T$ is a sequence of zero-mean uncorrelated random vectors, whose covariance matrices $\Psi^{\mathcal{G}}_{j,i}(k) = I\!E\big[\mathcal{G}_j(k)\mathcal{G}_i(k)^T\big]$ are given by:

$$\begin{aligned}\Psi^{\mathcal{G}}_{j,i}(k) = \sum_{t_1=0}^{j}\sum_{r_1=0}^{i} \mathrm{st}^{-1}_{q^j,q^i}\Big(&\Phi^{\mathcal{T},i,j}_{r_1,t_1}(k) \\ &\cdot \Xi_{r_1,t_1} I\!E\big[X_{r_1+t_1}(k)\big]\Big),\end{aligned} \tag{45}$$

with:

$$\begin{aligned}\Phi^{\mathcal{T},i,j}_{r_1,t_1}(k) &= I\!E\big[\mathcal{T}^i_{r_1}(k)\otimes\mathcal{T}^j_{t_1}(k)\big]\\ &= \sum_{r_2,r_3}^{r\in\mathcal{R}_i}\sum_{t_2,t_3}^{t\in\mathcal{R}_j}\big(T^i_r(k)\otimes T^j_t(k)\big)\big(I_{mn^{r_1}}\otimes C^T_{mn^{t_1}b^{t_3},b^{r_3}}\big)\\ &\quad\cdot\Big(I_{m^2n^{r_1+t_1}}\otimes\big(\xi_{t_3+r_3}(k)-\xi_{t_3}(k)\otimes\xi_{r_3}(k)\big)\Big)\\ &\quad\cdot\big(I_{mn^{r_1}}\otimes C_{mn^{t_1},1}\big).\end{aligned}\tag{46}$$

PROOF. The proof is a straightforward consequence of Lemma 3.3 in (Germani *et al.*, 2003), according to the same remarks considered in the proof of Lemma 5.

Remark 7. It has to be stressed that, according to Lemmas 5 and 6, system (22) provides an *exact* generation model for the sequences $X^\nu(k)$ and $Y^\nu(k)$ (i.e. no approximation has been introduced).

Lemma 8. The noise sequences $\{\mathcal{F}(k)\}$ and $\{\mathcal{G}(k)\}$ are such that, for $1\le i,j\le\nu$ and $\forall k,h\in\mathbb{Z}^+$:

$$\begin{aligned}&I\!E\big[\mathcal{F}_j(k)\mathcal{G}^T_i(h)\big]=0, &&\forall k\neq h\\ &I\!E\big[\mathcal{F}_j(k)\mathcal{G}^T_i(k)\big]=Q_{j,i}(k).\end{aligned}\tag{47}$$

with:

$$\begin{aligned}Q_{j,i}(k) = \sum_{t_1=0}^{j}\sum_{r_1=0}^{i}\mathrm{st}^{-1}_{mn^j,q^i}\Big(&\mathcal{Q}^{i,j}_{r_1,t_1}(k)\Xi_{r_1,t_1}\\ &\cdot I\!E\big[X_{r_1+t_1}(k)\big]\Big),\end{aligned}\tag{48}$$

where $\mathcal{Q}^{i,j}_{r_1,t_1}(k)=I\!E\big[\mathcal{T}^i_{r_1}(k)\otimes S^j_{t_1}(k)\big]$.

PROOF. The proof is a straightforward consequence of Lemma 3.4 in (Germani *et al.*, 2003), according to the same remarks considered in the proof of Lemma 5.

According to Lemma 8, the extended noises $\mathcal{F}(k)$ and $\mathcal{G}(k)$ are correlated at the same instant k, so that the Kalman Filter equations for correlated noises have been adopted for the computation of $\widehat{X}^\nu(k)$, i.e. the best linear filter for system (22) (Balakhrishnan, 1984); the straightforward algorithm is the following:

$$\begin{aligned}&\widehat{X}^\nu(0|-1)=I\!E\big[X^\nu(0)\big],\\ &\widehat{X}^\nu(k)=\widehat{X}^\nu(k|k-1)+\mathcal{K}(k)\\ &\qquad\cdot\big(Y^\nu(k)-\mathbf{C}^\nu(k)\widehat{X}^\nu(k|k-1)\big),\\ &\widehat{X}^\nu(k+1|k)=\mathbf{A}^\nu(k)\widehat{X}^\nu(k)+\mathcal{Z}(k)\\ &\qquad\cdot\big(Y^\nu(k)-\mathbf{C}^\nu(k)\widehat{X}^\nu(k|k-1)\big).\end{aligned}\tag{49}$$

The gain matrices $\mathcal{K}(k)$ and $\mathcal{Z}(k)$ are recursively computed through the following Riccati equations:

$$\begin{aligned}&P_P(0)=\mathrm{Cov}\big(X^\nu(0)\big),\\ &\mathcal{Z}(k)=Q(k)\big(\mathbf{C}^\nu(k)P_P(k)\mathbf{C}^{\nu T}(k)+\Psi^{\mathcal{G}}(k)\big)^\dagger,\\ &\mathcal{K}(k)=P_P(k)\mathbf{C}^{\nu T}(k)\\ &\qquad\cdot\big(\mathbf{C}^\nu(k)P_P(k)\mathbf{C}^{\nu T}(k)+\Psi^{\mathcal{G}}(k)\big)^\dagger,\\ &P(k)=P_P(k)-\mathcal{K}(k)\mathbf{C}^\nu(k)P_P(k),\\ &P_P(k+1)=\mathbf{A}^\nu(k)P(k)\mathbf{A}^{\nu T}(k)+\Psi^{\mathcal{F}}(k)\\ &\qquad-\mathcal{Z}(k)Q^T(k)-\mathbf{A}^\nu(k)\mathcal{K}(k)Q^T(k)\\ &\qquad-Q(k)\mathcal{K}^T(k)\mathbf{A}^{\nu T}(k),\end{aligned}\tag{50}$$

where in (50) the Moore-Penrose pseudoinverse has been used.

Remark 9. The algorithm initialization (i.e. $\widehat{X}^\nu(0|-1)$ and $P_P(0)$) requires the knowledge of the initial state statistics up to 2ν degree, which are finite and available according to (20).

Remark 10. Note that the recursive computation of $\Psi^{\mathcal{F}}(k)$, $\Psi^{\mathcal{G}}(k)$ and $Q(k)$ requires the computation of the expectations $I\!E\big[X_i(k)\big]$, $i=1,\dots,2\nu$ (see (38), (45) and (48)). These are the components of $I\!E\big[X^\nu(k)\big]$, and are recursively computed as

$$I\!E\big[X^\nu(k+1)\big]=\mathbf{A}^{2\nu}(k)I\!E\big[X^\nu(k)\big].\tag{51}$$

4. SIMULATION RESULTS

This section reports simulation results referred to a system of the type (1), characterized by the following data:

- $x(k)\in I\!R^3$, $u(k)\in I\!R$, $y(k)\in I\!R^2$, $\mathcal{W}=\{1,2\}$;
- $A_1=\begin{bmatrix}0.5 & 0 & 0.2\\ -1.75 & 0.5 & 0\\ 0 & 0 & 0.1\end{bmatrix}$, $A_2=\begin{bmatrix}0.3 & 0.25 & 0.1\\ -1.75 & 0.5 & 0\\ 0 & 1.2 & 1\end{bmatrix}$;
- $B_1=\begin{bmatrix}0\\ -1\\ 1\end{bmatrix}$, $C_1=\begin{bmatrix}1 & -1 & 0\\ 2 & 0 & 1\end{bmatrix}$, $D_1=\begin{bmatrix}1\\ 2\end{bmatrix}$,

  $B_2=\begin{bmatrix}1\\ 2\\ 0\end{bmatrix}$, $C_2=\begin{bmatrix}-1 & 0 & 1\\ 0 & 0 & 1\end{bmatrix}$, $D_2=\begin{bmatrix}0.5\\ -0.2\end{bmatrix}$;
- $F_1=\begin{bmatrix}0.1 & 0 & 0\\ 0.1 & 0 & 0\\ 0 & 0.3 & 0\end{bmatrix}$, $F_2=\begin{bmatrix}0 & 0.1 & 0\\ 0.1 & 0.1 & 0\\ -0.1 & 0 & 0\end{bmatrix}$;
- $G_1=\begin{bmatrix}0 & 0 & 0.1\\ 0 & 0 & -0.3\end{bmatrix}$, $G_2=\begin{bmatrix}0 & 0 & 0.2\\ 0 & 0 & 0.1\end{bmatrix}$;
- the noise $N(k)\in I\!R^3$ has independent components, with distributions:

$$\begin{aligned}&P\big(N_1(k)=-1/2\big)=0.8,\ P\big(N_2(k)=-1/3\big)=0.9\\ &P\big(N_1(k)=2\big)=0.2,\ P\big(N_2(k)=3\big)=0.1,\end{aligned}\tag{52}$$

The distribution of N_3 is identical to that of N_1.

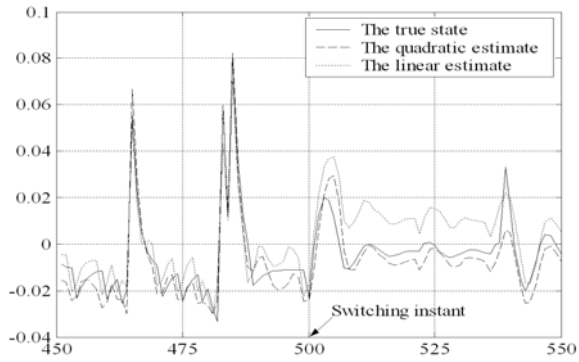

Fig. 1. True and estimated $x_1(k)$.

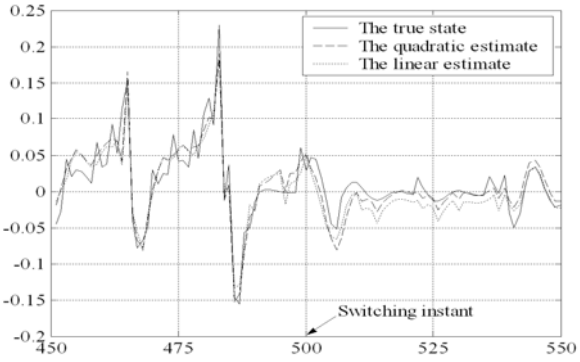

Fig. 2. True and estimated state $x_2(k)$.

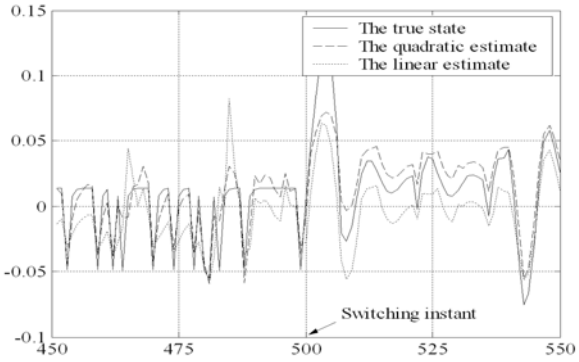

Fig. 3. True and estimated state $x_3(k)$.

In the simulation presented $u(k) \equiv 1,\ k \geq 0$. The initialization of the state estimate is made considering x_0 a gaussian variable, while the initial estimate of ϑ_0 is the mean of the components of the base vectors ($m = 2$).

As announced in the introduction, although the derivation of the polynomial filter has been made under the assumption of a constant parameter μ, the simulations here reported consider one switch of the parameter (i.e. a change of the system working mode) during the system evolution. In particular, the numerical data here reported refer to a simulation over a 1.000 steps interval, in which one switch of the parameter occurrs at time $k = 500$.

The figures report components of the true state and of the state estimates obtained with a first order ($\nu = 1$) and a second order ($\nu = 2$) filter. The sampling variances of the estimation errors of the linear and quadratic filters before and after the switching instant are reported below. The 500 steps before the switching ($\mu(k) = 1$) give the following error variances of the 3 state components:

$$\sigma_1^2|_{\nu=1} = 1.46 \cdot 10^{-4}, \qquad \sigma_1^2|_{\nu=2} = 8.01 \cdot 10^{-5},$$
$$\sigma_2^2|_{\nu=1} = 8.54 \cdot 10^{-4}, \qquad \sigma_2^2|_{\nu=2} = 6.67 \cdot 10^{-4},$$
$$\sigma_3^2|_{\nu=1} = 9.80 \cdot 10^{-4}, \qquad \sigma_3^2|_{\nu=2} = 3.94 \cdot 10^{-4}.$$

The 500 steps after the switching ($\mu(k) = 2$) give:

$$\sigma_1^2|_{\nu=1} = 1.94 \cdot 10^{-4}, \qquad \sigma_1^2|_{\nu=2} = 1.13 \cdot 10^{-4},$$
$$\sigma_2^2|_{\nu=1} = 6.57 \cdot 10^{-4}, \qquad \sigma_2^2|_{\nu=2} = 6.16 \cdot 10^{-4},$$
$$\sigma_3^2|_{\nu=1} = 8.25 \cdot 10^{-4}, \qquad \sigma_3^2|_{\nu=2} = 2.61 \cdot 10^{-4}.$$

The improvement of the quadratic filter over the linear one is evident: for some state components the reduction of the error variance is about 60%.

5. CONCLUSIONS

The problem of the simultaneous state and parameters estimation for a class of uncertain stochastic systems has been investigated in this paper, and the equations of the best polynomial filter are derived. Simulation results show the improvement of the second order filter with respect to the first order one.

REFERENCES

Balakhrishnan, A. V. (1984). *Kalman Filtering Theory*. Optimization Software. New York.

Bellman, R. (1970). *Introduction to matrix analysis*. McGraw-Hill. New york.

Blom, H.A.P. and Y. Bar-Shalom (1988). The interactive multiple model algorithm for systems with markovian switching coefficients. *IEEE Trans. on Automatic Control* **33**(8), 780–783.

Carravetta, F., A. Germani and M. Raimondi (1996). Polynomial filtering for linear discrete-time non-gaussian systems. *SIAM Journal on Control and Optimization* **34**(5), 1666–1690.

Carravetta, F., A. Germani and M. Raimondi (1997). Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise. *IEEE Trans. on Automatic Control* **42**(8), 1106–1126.

Costa, O.L.V. (1994). Linear minimum mean square error estimation for discrete-time markovian jump linear systems. *IEEE Trans. on Automatic Control* **39**(8), 1685–1689.

Germani, A., C. Manes and P. Palumbo (2003). Polynomial filtering for stochastic systems with markovian switching coefficients. In: *Proc. of 42^{th} IEEE Conf. on Decision and Control (CDC'03), Maui, Hawaii, USA*.

Germani, A., C. Manes and P. Palumbo (2004). Polynomial filtering for stochastic non-gaussian descriptor systems. *IEEE Trans. on Circuits and Systems II*.