Proceedings of the 42nd IEEE
Conference on Decision and Control
Maui, Hawaii USA, December 2003

TuP12-3

Polynomial Filtering for Stochastic Systems with Markovian Switching Coefficients

A. Germani[1,2] C. Manes[1,2] P. Palumbo[2]

[1]Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, Poggio di Roio, 67040 L'Aquila - Italy, {germani,manes}@ing.univaq.it

[2]Istituto di Analisi dei Sistemi e Informatica del CNR "A. Ruberti", IASI-CNR Viale Manzoni 30, 00185 Roma - Italy, palumbo@iasi.rm.cnr.it

Abstract

In this paper the state estimation problem for discrete-time Markovian switching systems affected by additive noise (not necessarily Gaussian) is solved following a polynomial approach. The key point for the derivation of the optimal polynomial filter is the possibility to represent the Markov switching systems as bilinear systems (linear drift, multiplicative noise) by means of a suitable state augmentation. By construction, the optimal polynomial filter of a given degree ν provides the minimum error variance among all polynomial output transformations of the same degree. Obviously, for $\nu > 1$ better performances are obtained with respect to linear filters. Simulation results are reported as a validation of the theory.

1. Introduction

Many authors investigated the problem of state estimation for linear systems with switching coefficients modeled by a finite-state Markov Chain (see e.g. [1,3,6,7,8,10,12] for the discrete-time case and [11,13,14] for the continuous-time case). The problem was first formulated in [1], where the authors pointed out the complexity of the exact solution and proposed an approximate solution. In [8], assuming a partial observation of the switching process, an almost-recursive implementation of the minimum variance state estimator is derived, whose complexity grows geometrically with time. In [6] a linear filter is implemented based on a clever use of the characteristic function associated to the Markovian jump parameter. In [7] different approximate state estimators have been analyzed, without assuming observations on the switching process. All estimators proposed in [7] are iterative algorithms over a finite observation time, and do not allow a recursive implementation.

This paper proposes a recursive polynomial algorithm for the state estimation of discrete-time Markovian switching systems. The polynomial approach finds the minimum variance state estimator in the closed linear space of all polynomial output transformations of a chosen degree ν. This approach has led to important results in the field of suboptimal filtering of non Gaussian linear [4] and bilinear [5] systems. In [9] the authors presented the equations of the optimal linear filter for stochastic switching systems (a filter statistically equivalent to the one presented in [6]). The polynomial filter proposed in this paper improves the performances of linear filters, and this can be particularly appreciated in presence of highly asymmetric non-Gaussian noises. All proofs of Lemmas and Theorems have been omitted due to the lack of space.

This work is supported by *CNR* (Italian National Research Council) and by *ASI* (Italian Aerospace Agency).

2. Modeling of Switching Systems

The aim of this paper is to derive the optimal polynomial filter of a chosen degree ν for the class of systems:

$$\begin{aligned} x(k+1) &= A_{\mu(k)}x(k) + B_{\mu(k)}u(k) + F_{\mu(k)}N(k), \quad k \in Z^+ \\ y(k) &= C_{\mu(k)}x(k) + D_{\mu(k)}u(k) + G_{\mu(k)}N(k), \qquad (2.1) \end{aligned}$$

where $x(k)$ is a stochastic variable in $\mathbb{R}^n$, $u(k)$ is a deterministic known input in $\mathbb{R}^p$, $y(k)$ is the measured output in $\mathbb{R}^q$. All matrices in (2.1) take values on a finite set, depending on the jump parameter $\mu(k)$, which is a scalar Markov process taking values in $\mathcal{M} = \{1, \dots, m\}$ with known probability transition matrix $\Pi \in [0,1]^{m \times m}$ and initial distribution $p \in [0,1]^m$:

$$[\Pi]_{i,j} = P\Big(\mu(k+1) = i | \mu(k) = j\Big), \quad i,j \in \mathcal{M}, \quad (2.2)$$

$$p_i = P\big(\mu(0) = i\big), \quad i \in \mathcal{M}. \quad (2.3)$$

The noise $N(k) \in \mathbb{R}^b$ is a sequence of zero-mean independent random vectors with finite and available moments up to the $2\nu^{\text{th}}$ degree, named:

$$\mathbb{E}\big[N^{[j]}(k)\big] = \xi_j, \quad 0 \le j \le 2\nu, \quad (2.4)$$

where the superscript $^{[i]}$ denotes the Kronecker power, defined for a given matrix M by

$$M^{[0]} = 1, \qquad M^{[i]} = M \otimes M^{[i-1]}, \quad i \ge 1, \quad (2.5)$$

0-7803-7924-1/03/$17.00 ©2003 IEEE

with $\otimes$ the standard Kronecker product (for a quick survey on Kronecker products and their principal properties, see [5] and references therein). Note that, according to the noise statistics: $\xi_0 = 1$ and $\xi_1 = 0$. The initial state $x(0)$ is a random variable with finite and available moments up to the $2\nu^{\text{th}}$-degree, named:

$$\mathbb{E}\big[x_0^{[j]}\big] = \zeta_j, \qquad 1 \le j \le 2\nu. \tag{2.6}$$

It is assumed that the sequences $N(k), \mu(k)$ and the initial state $x(0)$ form a set of independent random variable. Throughout the paper the symbol I_n denotes the identity matrix in $\mathbb{R}^{n\times n}$. In case of ambiguity, a zero matrix in $\mathbb{R}^{p\times q}$ is denoted by $O_{p\times q}$, otherwise, no subscripts are adopted.

As a first step, it is useful to introduce a state space realization for the Markov process. In the following, let e_j be the j-th column of I_m, and let $\mathcal{E}_m$ denote the natural basis in $\mathbb{R}^m$, i.e. $\mathcal{E}_m = \{e_j,\ j = 1, \ldots, m\}$.

Lemma 2.1. *Let $\{\theta(k) \in \mathcal{E}_m,\ k \in Z^+\}$ be a stochastic sequence obeying the recursive equation:*

$$\theta(0) = \theta_0, \quad \theta(k+1) = V(k)\theta(k), \quad k \in Z^+, \tag{2.7}$$

where θ_0 is a random variable in $\mathcal{E}_m$ with

$$P(\theta_0 = e_i) = p_i, \quad i \in \mathcal{M}, \tag{2.8}$$

and $V(k)$ is a sequence of random matrices, whose j-th column sequence $V_j(k)$ takes values in $\mathcal{E}_m$ with probabilities

$$P\Big(V_j(k) = e_i\Big) = [\Pi]_{i,j}, \quad i,j \in \mathcal{M}. \tag{2.9}$$

The matrix Π and the vector p are the same as defined in (2.2) and (2.3). Moreover, the columns of $V(k)$, together with θ_0, forms a set of independent random vectors. Then, defining the matrix $\widetilde{A} = [A_1 \cdots A_m] \in \mathbb{R}^{n\times n\cdot m}$, the sequence $A_{\mu(k)}$ can be represented as

$$A_{\mu(k)} = \widetilde{A}\cdot\big(\theta(k)\otimes I_n\big). \tag{2.10}$$

Similar representations are valid for all the system matrices defined in (2.1), e.g. $C_{\mu(k)} = \widetilde{C}\cdot\big(\theta(k)\otimes I_n\big)$ with $\widetilde{C} = [C_1 \cdots C_m] \in \mathbb{R}^{q\times n\cdot m}$, and so on.

Since $\theta(k) \in \mathcal{E}_m$, we have:

$$\theta^{[2]}(k) = E_2\theta(k), \quad \text{where } E_2 = \big[e_1^{[2]} \cdots e_m^{[2]}\big]. \tag{2.11}$$

Lemma 2.2. *The random variables $V(k)$ and $\theta(k)$ (at a given k) are independent. Moreover, $\mathbb{E}\big[V(k)\big] = \Pi$ and the zero-mean random sequence $\mathcal{V}(k) = V(k) - \Pi$ is such that:*

$$\mathbb{E}\big[\mathcal{V}_i(k)\otimes\mathcal{V}_j(k)\big] = \begin{cases} E_2\cdot\Pi_i - \Pi_i^{[2]}, & i = j, \\ O, & i \ne j, \end{cases} \tag{2.12}$$

with $\mathcal{V}_i(k)$ and Π_i the i-th columns of the matrices $\mathcal{V}(k)$ and Π, respectively. Moreover,

$$\begin{aligned} &\mathbb{E}\big[\mathcal{V}^{[2]}(k)\big] = \bar{\mathcal{V}}_2(k)E_2^T, \\ &\text{where } \bar{\mathcal{V}}_2(k) = \Big[\mathbb{E}\big[\mathcal{V}_1^{[2]}(k)\big] \cdots \mathbb{E}\big[\mathcal{V}_m^{[2]}(k)\big]\Big]. \end{aligned} \tag{2.13}$$

Proposition 2.3. *The switching system (2.1) admits the representation:*

$$\begin{aligned} x(k+1) &= \widetilde{A}\big(\theta(k)\otimes x(k)\big) + \bar{B}(k)\theta(k) + \widetilde{F}\big(\theta(k)\otimes N(k)\big), \\ \theta(k+1) &= \Pi\theta(k) + \mathcal{V}(k)\theta(k), \qquad k \ge 0, \\ y(k) &= \widetilde{C}\big(\theta(k)\otimes x(k)\big) + \bar{D}(k)\theta(k) + \widetilde{G}\big(\theta(k)\otimes N(k)\big), \end{aligned} \tag{2.14}$$

where the matrices $\bar{B}(k)$, $\bar{D}(k)$ are given by:

$$\bar{B}(k) = \widetilde{B}\big(I_m\otimes u(k)\big), \quad \bar{D}(k) = \widetilde{D}\big(I_m\otimes u(k)\big). \tag{2.15}$$

where $\widetilde{B} = [B_1 \cdots B_m]$ and $\widetilde{D} = [D_1 \cdots D_m]$.

3. The Polynomial Filter

It is well known that the state expectation conditioned to all measurements up to the current time provides the minimum error variance state estimate, and coincides with the projection of the state onto the linear space of all the Borel functions of the measurements:

$$\hat{x}(k) = \Pi\big[x(k)|\mathcal{B}(Y_k)\big], \quad Y_k = \begin{bmatrix} y(0) \\ \vdots \\ y(k) \end{bmatrix}. \tag{3.1}$$

In the Gaussian case the optimal estimator is a linear transformation of the measurements, implemented by the Kalman filter. In the non-Gaussian case the conditional expectation can be extremely difficult to compute, and it is convenient to consider suboptimal polynomial estimates, obtained by projecting the state onto subspaces of polynomial transformations of the measurements [4, 5]. To this aim, consider the Hilbert space of all polynomial transformations of the measurements of a given chosen degree ν:

$$L(Y_k^\nu) = \text{span}\{1, Y^\nu(0), \cdots, Y^\nu(k)\} \subset \mathcal{B}(Y_k), \tag{3.2}$$

$$\text{where } Y_k^\nu = \begin{bmatrix} Y^\nu(0) \\ \vdots \\ Y^\nu(k) \end{bmatrix}, \quad Y^\nu(h) = \begin{bmatrix} y(h) \\ y^{[2]}(h) \\ \vdots \\ y^{[\nu]}(h) \end{bmatrix}, \tag{3.3}$$

(the extra-assumption that $\mathbb{E}\{|y^{[i]}(h)|^2\} < \infty$, for $i = 1, \ldots, \nu$ is needed). Then, the optimal polynomial state estimate of degree ν is

$$\hat{x}_\nu(k) = \Pi\big[x(k)|L(Y_k^\nu)\big]. \tag{3.4}$$

Theorem 3.1. *The optimal ν-th degree polynomial estimate of the state $x(k)$ of the switching system* (2.1) *is given by:*

$$\hat{x}_\nu(k) = \Sigma_n \widehat{X}^\nu(k) = \Sigma_n \Pi\big[X^\nu(k)|L(Y_k^\nu)\big], \qquad (3.5)$$

where

$$\begin{aligned}\Sigma_n &= \big[O_{n\times m}\ \Sigma\ O_{n\times m(n^2+\cdots+n^\nu)}\big],\\ \Sigma &= \big[I_n \cdots I_n\big] \in \mathbb{R}^{n\times mn}\\ X^\nu(k) &= \begin{bmatrix} X_0(k)\\ \vdots \\ X_\nu(k)\end{bmatrix}, \quad X_i(k) = \theta(k)\otimes x^{[i]}(k).\end{aligned} \qquad (3.6)$$

Proof. According to the measurement equation in (2.14), all the Kronecker powers of the output depend on the vectors X_i defined in (3.6). Moreover, considering that $\theta(k) \in \mathcal{E}_m$, then $x(k) = \Sigma\big(\theta(k)\otimes x(k)\big) = \Sigma X_1(k) = \Sigma_n X^\nu(k)$ so that the polynomial minimum variance estimate in (3.2) is:

$$\begin{aligned}\hat{x}_\nu(k) &= \Pi\big[x(k)|L(Y_k^\nu)\big] = \Pi\big[\Sigma_n X^\nu(k)|L(Y_k^\nu)\big]\\ &= \Sigma_n \Pi\big[X^\nu(k)|L(Y_k^\nu)\big] = \Sigma_n \widehat{X}^\nu(k),\end{aligned} \qquad (3.7)$$

∎

The remaining of the paper is devoted to the computation of the projection in equation (3.7). The first step is to show that the sequences $\{X^\nu(k)\}$ and $\{Y^\nu(k)\}$ obey difference equations of the type

$$\begin{aligned}X^\nu(k+1) &= \mathbf{A}^\nu(k)X^\nu(k) + \mathcal{F}(k),\\ Y^\nu(k) &= \mathbf{C}^\nu(k)X^\nu(k) + \mathcal{G}(k),\end{aligned} \quad k \geq 0 \qquad (3.8)$$

with $\mathbf{A}^\nu(k)$ and $\mathbf{C}^\nu(k)$ suitably defined deterministic matrices and

$$\begin{aligned}\mathcal{F}(k) &= \widetilde{\mathcal{F}}\big(k, u(k), X^\nu(k), N(k)\big),\\ \mathcal{G}(k) &= \widetilde{\mathcal{G}}\big(k, u(k), X^\nu(k), N(k)\big),\end{aligned} \qquad (3.9)$$

with $\widetilde{\mathcal{F}}$ and $\widetilde{\mathcal{G}}$ suitably defined function where $X^\nu(k)$ multiplies the noise $N(k)$ and its powers up to order ν, in a way that $\mathcal{F}(k)$ and $\mathcal{G}(k)$ result to be white sequences.

The importance of the representation (3.8) is that, once $\mathbf{A}^\nu(k)$ and $\mathbf{C}^\nu(k)$ are known, together with the covariance matrices of the white sequences $\mathcal{F}(k)$ and $\mathcal{G}(k)$, the minimum variance filter for such a kind of bilinear system (see [5]) can be used to estimate the extended state $X^\nu(k)$, from which the state $x(k)$ is estimated.

In the sequel some Lemmas are reported showing the construction of the matrices $\mathbf{A}^\nu(k)$, $\mathbf{C}^\nu(k)$ and the computations of the statistics of the noises of the extended system (3.8). Before the statement of the Lemmas some notations must be introduced.

Given a pair of integers (a, b), the symbol $C_{a,b}$ denotes a *commutation matrix*, that is a matrix in $\{0,1\}^{ab\times ab}$ such that, given any two matrices $A \in \mathbb{R}^{r_a\times c_a}$ and $B \in \mathbb{R}^{r_b\times c_b}$ (see [5])

$$B\otimes A = C^T_{r_a,r_b}(A\otimes B)C_{c_a,c_b}. \qquad (3.10)$$

The Kronecker products and powers of vectors X_j, defined in (3.6), satisfy

$$X_j^{[h]} = \Theta_n^{h,j}X_{jh}, \quad X_i\otimes X_j = \Xi_{i,j}X_{i+j}, \quad \forall i,j,h \in Z^+, \qquad (3.11)$$

where $\Theta_n^{0,j} = [1\cdots 1]$, and for $h > 0$

$$\Theta_n^{h+1,j} = \big(\Theta_n^{h,j}\otimes I_{mn^j}\big)\big(I_m\otimes C^T_{mn^j,n^{jh}}\big)\big(E_2\otimes I_{n^{j(h+1)}}\big), \qquad (3.12)$$

and $\Xi_{i,j} = (I_m\otimes C^T_{mn^j,n^i})(E_2\otimes I_{n^{i+j}})$.

Note that:

$$\Theta_n^{1,j} = I_{mn^j}, \quad \Theta_n^{2,0} = \Xi_{0,0} = E_2. \qquad (3.13)$$

Recall that the stack of a matrix $A \in \mathbb{R}^{r\times c}$ is the vector in $\mathbb{R}^{r\cdot c}$ that piles up all the columns of matrix A, and is denoted $\mathrm{st}(A)$. The inverse operation is denoted $\mathrm{st}^{-1}_{r,c}(\cdot)$, and transforms a vector of size $r\cdot c$ in an $r\times c$ matrix. When written without any subscript, the inverse stack operator should be intended to generate a square matrix, so that if A is a square matrix then $\mathrm{st}^{-1}\big(\mathrm{st}(A)\big) = A$.

Lemma 3.2. *The iterative equation of the component $X_j(k)$ of the extended state $X^\nu(k)$ as defined in* (3.6), *that is the j-th row-block of the state equation* (3.8) *for $j = 0, 1, \ldots \nu$, can be put in the form*

$$\begin{aligned}X_j(k+1) &= \sum_{t_1=0}^{j} \mathbf{A}_{j,t_1}(k)X_{t_1}(k) + \mathcal{F}_j(k),\\ \text{where } \mathcal{F}_j(k) &= \sum_{t_1=0}^{j} S^j_{t_1}(k)X_{t_1}(k),\end{aligned} \qquad (3.14)$$

with $\mathbf{A}_{j,t_1}(k)$, $S^j_{t_1}(k)$ sequences of deterministic and random matrices, respectively. The sequence of random vectors $\mathcal{F}(k) = \big(\mathcal{F}_0(k)^T\ \cdots\ \mathcal{F}_\nu(k)^T\big)^T$ is white, i.e. zero-mean and such that $\mathbb{E}\big[\mathcal{F}(k)\mathcal{F}(h)^T\big] = 0$, $\forall k \neq h$.

The expressions of matrices $\mathbf{A}_{j,t_1}(k)$, $S^j_{t_1}(k)$ and of the covariances $\Psi^{\mathcal{F}}_{j,i}(k) = \mathbb{E}\big[\mathcal{F}_j(k)\mathcal{F}_i(k)^T\big]$ are quite complex and are reported in the Appendix.

Lemma 3.3. *The measurements equations for j-th Kronecker power of the output, that is the j-th row-block of the output equation* (3.8) *for $j = 1, 2, \ldots \nu$, can be put in the form*

$$\begin{aligned}y^{[j]}(k) &= \sum_{t_1=0}^{j} \mathbf{C}_{j,t_1}(k)X_{t_1}(k) + \mathcal{G}_j(k),\\ \text{with } \mathcal{G}_j(k) &= \sum_{t_1=0}^{j} T^j_{t_1}(k)X_{t_1}(k)\end{aligned} \qquad (3.15)$$

with $\mathbf{C}_{j,t_1}(k)$, $\mathcal{T}^j_{t_1}(k)$ sequences of deterministic and random matrices, respectively. The sequence of random vectors $\mathcal{G}(k) = \big(\mathcal{G}_1(k)^T \ \cdots \ \mathcal{G}_\nu(k)^T\big)^T$ is white, i.e. zero-mean and such that $\mathbb{E}\big[\mathcal{G}(k)\mathcal{G}(h)^T\big] = 0, \ \forall k \neq h$.

The expressions of matrices $\mathbf{C}_{j,t_1}(k)$, $\mathcal{T}^j_{t_1}(k)$ and of the covariances $\Psi^{\mathcal{G}}_{j,i}(k) = \mathbb{E}\big[\mathcal{G}_j(k)\mathcal{G}_i(k)^T\big]$ are quite complex and are reported in the Appendix.

Lemma 3.4. *The sequences $\mathcal{F}(k)$ and $\mathcal{G}(k)$ are such that:*

$$\begin{aligned} &\mathbb{E}\big[\mathcal{F}_j(k)\mathcal{G}_i^T(h)\big] = 0, \quad k \neq h, \quad 0 \le j \le \nu, \\ &\mathbb{E}\big[\mathcal{F}_j(k)\mathcal{G}_i^T(k)\big] = Q_{j,i}(k), \qquad 0 < i \le \nu, \end{aligned} \quad \forall\, k,h \in Z^+. \tag{3.16}$$

with

$$Q_{j,i}(k) = \sum_{t_1=0}^{j}\sum_{r_1=0}^{i} \mathrm{st}^{-1}_{mn^j,q^i}\Big(\mathcal{Q}^{i,j}_{r_1,t_1}(k)\Xi_{r_1,t_1}\mathbb{E}\big[X_{r_1+t_1}(k)\big]\Big) \tag{3.17}$$

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

Now, having shown that the extended polynomial state and output of system (2.1), and of its bilinear representation (2.14), obeys the equations (3.8), following Theorem 3.1, the polynomial filtering algorithm can be constructed as follows:

Proposition 3.5. *The ν-th degree polynomial filtering algorithm is the following:*

$$\begin{aligned} \widehat{X}^\nu(k+1|k) &= \mathbf{A}^\nu(k)\widehat{X}^\nu(k|k-1) && (3.18a)\\ &\quad + \mathcal{H}(k)\big(Y^\nu(k) - \mathbf{C}^\nu(k)\widehat{X}^\nu(k|k-1)\big), \\ \widehat{X}^\nu(k) &= \widehat{X}^\nu(k|k-1) && (3.18b)\\ &\quad + \mathcal{K}(k)\Big(Y^\nu(k) - \mathbf{C}^\nu(k)\widehat{X}^\nu(k|k-1)\Big), \\ \hat{x}_\nu(k) &= \Sigma_n\widehat{X}^\nu(k), && (3.18c) \end{aligned}$$

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

$$\begin{aligned} \mathcal{Z}(k) &= Q(k)\Big(\mathbf{C}^\nu(k)\mathcal{P}_P(k)\mathbf{C}^{\nu T}(k) + \Psi^{\mathcal{G}}(k)\Big)^\dagger && (3.19a)\\ \mathcal{P}_P(k+1) &= \mathbf{A}^\nu(k)\mathcal{P}(k)\mathbf{A}^{\nu T}(k) + \Psi^{\mathcal{F}}(k) \\ &\quad + \mathcal{Z}(k)Q^T(k) - \mathcal{H}(k)Q^T(k) - Q(k)\mathcal{H}^T(k) && (3.19b)\\ \mathcal{P}(k) &= \mathcal{P}_P(k) - \mathcal{K}(k)\mathbf{C}^\nu(k)\mathcal{P}_P(k) && (3.19c)\\ \mathcal{K}(k) &= \mathcal{P}_P(k)\mathbf{C}^{\nu T}(k) \\ &\quad \cdot\Big(\mathbf{C}^\nu(k)\mathcal{P}_P(k)\mathbf{C}^{\nu T}(k) + \Psi^{\mathcal{G}}(k)\Big)^\dagger && (3.19d)\\ \mathcal{H}(k) &= \mathbf{A}^\nu(k)\mathcal{K}(k) + \mathcal{Z}(k) && (3.19e) \end{aligned}$$

(in (3.19a and (3.19d) the Moore-Penrose pseudoinverse has been used). Matrices $\Psi^{\mathcal{F}}(k)$, $\Psi^{\mathcal{G}}(k)$ are the extended state and measurements noise covariance matrices and are reported in Appendix. $Q(k)$ is the covariance matrix between $\mathcal{F}(k)$ and $\mathcal{G}(k)$ sequences at the same instants, given by equation (3.17) of lemma 3.4.

Proof. The filter equations are those of the classical Kalman filter [2] for the case of correlated state and output noises, applied to the system (3.8), that has a multiplicative noise structure (see equations (3.15), (3.14), describing the components of $\mathcal{F}$ and of $\mathcal{G}$). The use of the Kalman algorithm on an extended system to achieve optimal polynomial filtering of system with multiplicative noise has been demonstrated in [5]. ∎

Note that the Riccati equations (3.19) employ the covariance matrices $\Psi^{\mathcal{F}}(k)$, $\Psi^{\mathcal{G}}(k)$, whose expressions are reported in the Appendix, and the (mutual) covariance $Q(k)$, given by (3.17). The computation of these matrices requires the sequence $\mathbb{E}\big[X^{2\nu}(k)\big]$ of the moments of the state up to order 2ν. Such sequence can be obtained computing the evolution of the following (deterministic) system:

$$\mathbb{E}\big[X^{2\nu}(k+1)\big] = \mathbf{A}^{2\nu}(k)\mathbb{E}\big[X^{2\nu}(k)\big], \tag{3.20}$$

suitably initialized (note that $\mathbb{E}\big[X_j(0)\big] = \mathbb{E}\big[\theta(0)\big]\otimes\mathbb{E}\big[x^{[j]}(0)\big] = p\otimes\zeta_j$).

Remark 3.6. The covariance of the estimation error $x(k) - \hat{x}_\nu(k)$ can be extracted from $\mathcal{P}(k)$, the error covariance of the extended state used in the algorithm of Proposition 3.5, as follows.

$$\mathbb{E}\Big[\big(x(k)-\hat{x}_\nu(k)\big)\big(x(k)-\hat{x}_\nu(k)\big)^T\Big] = \Sigma_n\mathcal{P}(k)\Sigma_n^T \tag{3.21}$$

∎

4. Numerical simulations

This section reports simulation results referred to a system of the type (2.1), characterized by the following data:

$$\begin{gathered} x(k) \in \mathbb{R}^2, \quad u(k), y(k) \in \mathbb{R}, \quad \mathcal{M} = \{1,2,3\}; \\ A_1 = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.50 & 0.25 \\ -1.75 & 0.50 \end{bmatrix}, \\ A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \\ B_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}; \\ C_1 = [1\ 1], \quad C_2 = [1\ 0], \quad C_3 = [2\ 1]; \\ D_1 = 1, \quad D_2 = 0.5, \quad D_3 = 0; \\ F_1 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0.1 & 0 & 0 \end{bmatrix}, \quad F_2 = \begin{bmatrix} 0 & 0.1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \\ F_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.1 & 0 \end{bmatrix}; \\ G_1 = [0\ 0\ 0.1], \quad G_2 = [0\ 0\ 0.2], \quad G_3 = [0\ 0\ 0.04]; \end{gathered}$$

The deterministic input used in the simulation is $u(k) \equiv 1, \quad k \geq 0$. The noise $N(k) = (N_1(k), N_2(k), N_3(k))^T$ has independent components, following discrete asymmetric distributions:

$$\begin{aligned} &P(N_1(k) = -1/2) = 0.8, \quad &&P(N_2(k) = -1/3) = 0.9, \\ &P(N_1(k) = 2)0.2, &&P(N_2(k) = 3) = 0.1, \end{aligned} \tag{4.1}$$

(the distribution of N_3 is identical to the one of N_1;

The transition probability matrix of the Markovian parameter is

$$\Pi = \begin{bmatrix} 0.3 & 0.6 & 0.2 \\ 0.2 & 0.3 & 0.5 \\ 0.5 & 0.1 & 0.3 \end{bmatrix}$$

Figures 4.1 and 4.2 display the state estimates obtained with a first order filter ($\nu = 1$) and a second order filter ($\nu = 2$). The sampling error variances (over a simulation of 1000 steps) for the linear and quadratic filters are

$$\begin{aligned} \sigma_1^2|_{\nu=1} &= 0.0622, \quad &\sigma_1^2|_{\nu=2} &= 0.0468, \\ \sigma_2^2|_{\nu=1} &= 0.3554, &\sigma_2^2|_{\nu=2} &= 0.2676. \end{aligned} \tag{4.2}$$

The improvement of the quadratic filter over the linear one is evident.

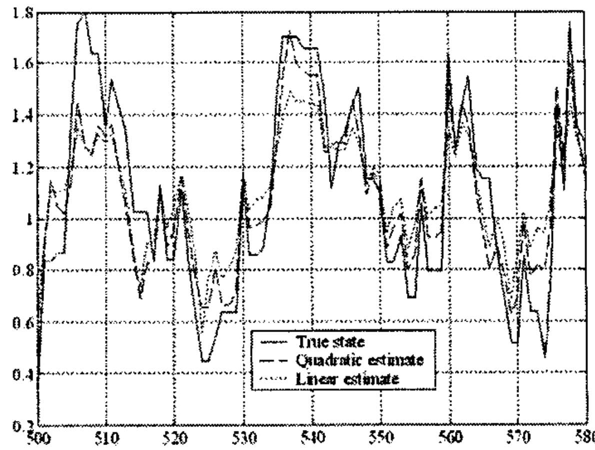

Fig. 4.1 True and estimated states: the first component.

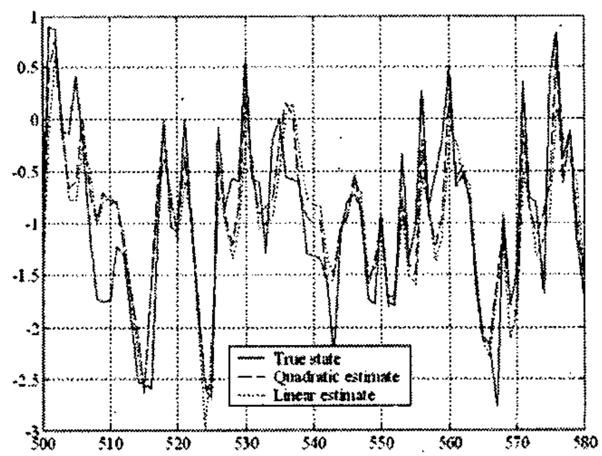

Fig. 4.2 True and estimated states: the second component.

Appendix

This Appendix is aimed to give the expression of some matrices appearing in the paper, whose structure is quite complex. In this Appendix the symbol t denotes a multi-index $t = (t_1, t_2, t_3) \in Z^{+^3}$. The modulus of a multiindex, denoted $|t|$ is the sum of its entries, i.e. $|t| = t_1 + t_2 + t_3\}$. The symbol M_t^s denotes the matrix coefficients of the following polynomial Kronecker power expansion

$$(a_1 + a_2 + a_3)^{[j]} = \sum_{|t|=j} M_t^s \left(a_1^{[t_1]} \otimes a_2^{[t_2]} \otimes a_3^{[t_3]}\right), \quad (A.3)$$

where $a_i \in I\!R^s$ and $M_t^s \in I\!R^{s^{|t|} \times s^{|t|}}$

The expressions of the matrices $\mathbf{A}_{j,t_1}(k)$, $S_{t_1}^j(k)$ appearing in Theorem 3.2 are the following:

$$\mathbf{A}_{j,t_1}(k) = (\Pi \otimes J_{t_1}^j(k))\Xi_{0,t_1}, \tag{A.4a}$$

$$\begin{aligned} S_{t_1}^j(k) = \Big(&\Pi \otimes \mathcal{L}_{t_1}^j(k) + \mathcal{V}(k) \otimes J_{t_1}^j(k) \\ &+ \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k)\Big)\Xi_{0,t_1}, \end{aligned} \tag{A.4b}$$

in which

$$J_{t_1}^j(k) = \sum_{t_2,t_3}^{t \in \mathcal{R}_j} L_t^j(k)\left(I_{mn^{t_1}} \otimes \xi_{t_3}\right), \tag{A.5a}$$

$$\mathcal{L}_{t_1}^j(k) = \sum_{t_2,t_3}^{t \in \mathcal{R}_j} L_t^j(k)\left(I_{mn^{t_1}} \otimes (N^{[t_3]}(k) - \xi_{t_3})\right), \tag{A.5b}$$

$$L_t^j(k) = M_t^n\left(\widetilde{A}^{[t_1]} \otimes \bar{B}^{[t_2]}(k) \otimes \widetilde{F}^{[t_3]}\right) K_t^j, \tag{A.5c}$$

$$\begin{aligned} K_t^j = &\left(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\right) \\ &\cdot \left(I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}\right)\left(\Xi_{t_1,0} \otimes I_{b^{t_3}}\right). \end{aligned} \tag{A.5d}$$

The expression of the (mutual) covariance matrices $\Psi_{j,i}^{\mathcal{F}}(k) = I\!E\left[\mathcal{F}_j(k)\mathcal{F}_i(k)^T\right]$ of the random sequences appearing in (3.14) is:

$$\Psi_{j,i}^{\mathcal{F}}(k) = \sum_{j_1=0}^{j} \sum_{i_1=0}^{i} \mathrm{st}_{mn^{j_1}, mn^{i_1}}^{-1}\left(\Phi_{r_1,t_1}^{S,i,j}(k) \cdot \Xi_{r_1,t_1} I\!E\left[X_{r_1+t_1}(k)\right]\right), \tag{A.6}$$

where $\Phi^{S,i,j}_{r_1,t_1}(k) = \mathbb{E}\big[S^i_{r_1}(k)\otimes S^j_{t_1}(k)\big]$ *is given by:*

$$\begin{aligned}\Phi^{S,i,j}_{r_1,t_1}(k) &= \big(I_m\otimes C^T_{mn^j,n^i}\big)\Big(\Pi^{[2]}\otimes\Phi^{\mathcal{L},j,i}_{t_1,r_1}(k)\\ &\quad+\big(\bar{V}_2(k)E_2^T\big)\otimes J^j_{t_1}(k)\otimes J^i_{r_1}(k)\\ &\quad+\big(\bar{V}_2(k)E_2^T\big)\otimes\Phi^{\mathcal{L},j,i}_{t_1,r_1}(k)\Big)\\ &\quad\big(I_m\otimes C_{m^2n^{t_1},mn^{r_1}}\big)\big(\Xi_{0,r_1}\otimes\Xi_{0,t_1}\big),\end{aligned} \tag{A.7}$$

with

$$\begin{aligned}\Phi^{\mathcal{L},j,i}_{t_1,r_1}(k) &= \mathbb{E}\big[\mathcal{L}^j_{t_1}(k)\otimes\mathcal{L}^i_{r_1}(k)\big]\\ &= \sum_{t_2,t_3}^{t\in\mathcal{R}_j}\sum_{r_2,r_3}^{r\in\mathcal{R}_i}\big(L^j_t(k)\otimes L^i_r(k)\big)\big(I_{mn^{t_1}}\otimes C^T_{mn^{r_1}b^{r_3},b^{t_3}}\big)\\ &\quad\cdot\Big(I_{m^2n^{t_1+r_1}}\otimes\big(\xi_{r_3+t_3}-\xi_{r_3}\otimes\xi_{t_3}\big)\Big)\\ &\quad\cdot\big(I_{mn^{t_1}}\otimes C_{mn^{r_1},1}\big).\end{aligned} \tag{A.8}$$

The expressions of the matrices $\mathbf{C}_{j,t_1}(k)$, $\mathcal{T}^j_{t_1}(k)$ appearing in Theorem 3.3 are the following

$$\mathbf{C}_{j,t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} T^j_t(k)\Big(I_{mn^{t_1}}\otimes\xi_{t_3}\Big), \tag{4.9a}$$

$$\mathcal{T}^j_{t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} T^j_t(k)\Big(I_{mn^{t_1}}\otimes\big(N^{[t_3]}(k)-\xi_{t_3}\big)\Big), \tag{4.9b}$$

$$T^j_t(k) = M^q_t\Big(\tilde{C}^{[t_1]}\otimes\bar{D}^{[t_2]}(k)\otimes\tilde{G}^{[t_3]}\Big)K^j_t, \tag{4.9c}$$

where matrices K^j_t are defined in (A.5d).

The expression of the (mutual) covariance matrices $\Psi^{\mathcal{G}}_{j,i}(k) = \mathbb{E}\big[\mathcal{G}_j(k)\mathcal{G}_i(k)^T\big]$ of the random sequences appearing in (3.15) is:

$$\Psi^{\mathcal{G}}_{j,i}(k) = \sum_{t_1=0}^{j}\sum_{r_1=0}^{i}\mathrm{st}^{-1}_{q^j,q^i}\left(\Phi^{\mathcal{T},i,j}_{r_1,t_1}(k)\Xi_{r_1,t_1}\mathbb{E}\big[X_{r_1+t_1}(k)\big]\right), \tag{A.10}$$

with

$$\begin{aligned}\Phi^{\mathcal{T},i,j}_{r_1,t_1}(k) &= \mathbb{E}\Big[\mathcal{T}^i_{r_1}(k)\otimes\mathcal{T}^j_{t_1}(k)\Big]\\ &= \sum_{r_2,r_3}^{r\in\mathcal{R}_i}\sum_{t_2,t_3}^{t\in\mathcal{R}_j}\Big(T^i_r(k)\otimes T^j_t(k)\Big)\big(I_{mn^{r_1}}\otimes C^T_{mn^{t_1}b^{t_3},b^{r_3}}\big)\\ &\quad\cdot\Big(I_{m^2n^{r_1+t_1}}\otimes\Big(\xi_{t_3+r_3}-\xi_{t_3}\otimes\xi_{r_3}\Big)\Big)\\ &\quad\cdot\big(I_{mn^{r_1}}\otimes C_{mn^{t_1},1}\big).\end{aligned} \tag{A.11}$$

References

[1] G. A. Ackerson, K. S. Fu, "On State Estimation in Switching Environments," *IEEE Trans. Automat. Contr.*, Vol. 15, No. 1, pp. 10–17, 1970.

[2] A. V. Balakrishnan, *Kalman Filtering Theory*, New York: Optimization Software, 1984.

[3] H.A.P. Blom and Y. Bar-Shalom, "The interactive multiple model algorithm for systems with Markovian switching coefficients," *IEEE Trans. Automat. Contr.*, Vol. 33, No. 8, pp. 780–783, 1988.

[4] F. Carravetta, A. Germani and M. Raimondi, "Polynomial filtering for linear discrete-time non-Gaussian systems," *SIAM J. Control and Optim.*, Vol. 34, No. 5, pp. 1666–1690, 1996.

[5] F. Carravetta, A. Germani and M. Raimondi, "Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise," *IEEE Trans. Automat. Contr.*, Vol. 42, No. 8, pp. 1106–1126, 1997.

[6] O. L. V. Costa, "Linear Minimum Mean Square Error Estimation for Discrete-Time Markovian Jump Linear Systems," *IEEE Trans. Automat. Contr.*, Vol. 39, No. 8, pp. 1685–1689, 1994.

[7] A. Doucet, A. Logothetis, V. Krishnamurthy, "Stochastic Sampling Algorithms for State Estimation of Jump Markov Linear Systems," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 1, pp. 188–202, 2000.

[8] R. J. Elliot, F. Dufour, D. D. Sworder, "Exact Hybrid Filters in Discrete Time," *IEEE Trans. Automat. Contr.*, Vol. 41, No. 12, pp. 1807–1810, 1996.

[9] A. Germani, C. Manes and P. Palumbo, "State Estimation for a Class of Stochastic Variable Structure Systems," proc. of 41^{th} *Conf. on Decision and Control*, pp. 2514–2519, Las Vegas, Nevada, 2002.

[10] R.E. Helmick, W.D. Blair, S.A. Hoffman, "Fixed-interval smoothing for Markovian switching systems," *IEEE Trans. on Inf. Theory*, Vol. 41, No. 6, pp. 1845-1855, 1995.

[11] J. L. Hibey, C. D. Charalambous, "Conditional Densities for Contiuous-Time Nonlinear Hybrid Systems with Applications to Fault Detection," *IEEE Trans. Automat. Contr.*, Vol. 44, No. 11, pp. 2164–2169, 1999.

[12] A. Logothetis, V. Krishnamurthy, "Expectation maximization algorithms for MAP estimation of Jump Markov Linear Systems," *IEEE Trans. n Signal Processing*, Vol. 47, No. 8, pp. 2139–2156, 1999.

[13] B. M. Miller, W. J. Runggaldier, "Kalman Filtering for Linear Systems with Coefficients Driven by a Hidden Markov Jump Process," *Syst. and Contr. Lett.*, No. 31, pp. 93–102, 1997.

[14] Q. Zhang, "Hybrid Filtering for Linear Systems with Non-Gaussian Disturbances," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 1, pp. 50–61, 2000.