ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
"Antonio Ruberti"
CONSIGLIO NAZIONALE DELLE RICERCHE

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

POLYNOMIAL APPROACH FOR FILTERING AND IDENTIFICATION OF A CLASS OF UNCERTAIN SYSTEMS

R. 603 Dicembre 2003

Domenico Di Martino – Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, 67040 Monteluco (L'Aquila), Italy. Email: `domenico@ing.univaq.it`.

Alfredo Germani – Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, 67040 Monteluco (L'Aquila), Italy and Istituto di Analisi dei Sistemi ed Informatica del CNR, Viale Manzoni 30, 00185 Roma, Italy. Email: `germani@ing.univaq.it`.

Costanzo Manes – Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, 67040 Monteluco (L'Aquila), Italy and Istituto di Analisi dei Sistemi ed Informatica del CNR, Viale Manzoni 30, 00185 Roma, Italy. Email: `manes@ing.univaq.it`.

Pasquale Palumbo – Istituto di Analisi dei Sistemi ed Informatica del CNR, Viale Manzoni 30, 00185 Roma, Italy. Email: `palumbo@iasi.rm.cnr.it`.

ISSN: 1128–3378

Collana dei Rapporti dell'Istituto di Analisi dei Sistemi ed Informatica, CNR
viale Manzoni 30, 00185 ROMA, Italy

tel. ++39-06-77161
fax ++39-06-7716461
email: iasi@iasi.rm.cnr.it
URL: http://www.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.

Key words: Polynomial filtering, Stochastic Systems, Bilinear Systems, Uncertain systems

1. Introduction

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

$$\begin{aligned} x(k+1) &= A(\mu)x(k)+B(\mu)u(k)+F(\mu)N(k), \\ y(k) &= C(\mu)x(k)+D(\mu)u(k)+G(\mu)N(k), \end{aligned} \tag{1.1}$$

$k \in \mathbb{Z}^+$, 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 (1.1) 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 (1.1) when the parameter μ undergoes Markov transitions (see [5]–[11]). 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 [3] and bilinear systems [4]. Recently, polynomial filters have also been studied in the framework of descriptor [12] and Markov switching systems [5]. 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 (1.1). For the model derivation it is useful to regard the constant parameter μ as governed by the trivial difference equation

$$\mu(k+1) = \mu(k), \quad \mu(0) = \mu_0. \tag{2.1}$$

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

$$\widetilde{A} = \begin{bmatrix} A(1) & A(2) & \cdots & A(m) \end{bmatrix} \tag{2.2}$$

the following identities hold

$$A(i) = \widetilde{A}(e_i \otimes I_n), \quad i = 1, \dots, m, \tag{2.3}$$

as it is easy verified from

$$\widetilde{A}(e_i \otimes I_n) = \begin{bmatrix} A(1) \cdots A(i) \cdots A(m) \end{bmatrix} \begin{bmatrix} O_{(i-1)n \times n} \\ I_n \\ O_{(m-i)n \times n} \end{bmatrix}, \tag{2.4}$$

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:

$$H^{[0]} = 1, \quad H^{[i]} = H \otimes H^{[i-1]}, \quad i \geq 1. \tag{2.5}$$

(see Appendix A and/or [1] for more details, see [4] 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 (1.1), suitably exploiting the bijection $\mathcal{E}_m \leftrightarrow \mathcal{W}$. Defining $\vartheta = e_\mu$, all matrices in equation (1.1) can be written as follows

$$\begin{aligned} A(\mu) &= \widetilde{A}\big(\vartheta \otimes I_n\big), \quad B(\mu) = \widetilde{B}\big(\vartheta \otimes I_p\big), \\ C(\mu) &= \widetilde{C}\big(\vartheta \otimes I_n\big), \quad D(\mu) = \widetilde{D}\big(\vartheta \otimes I_p\big), \\ F(\mu) &= \widetilde{F}\big(\vartheta \otimes I_b\big), \quad G(\mu) = \widetilde{G}\big(\vartheta \otimes I_b\big), \end{aligned} \tag{2.6}$$

where all matrices $\widetilde{B}, \widetilde{C}, \widetilde{D}, \widetilde{F}, \widetilde{G}$ are defined as $\widetilde{A}$ in (2.2). From (2.1), the sequence $\vartheta(k) = e_{\mu(k)}$ is governed by equations

$$\vartheta(k+1) = \vartheta(k), \quad \vartheta(0) = \vartheta_0. \tag{2.7}$$

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

$$\vartheta^{[2]}(k) = E_2 \vartheta(k), \text{ with } E_2 = \big[e_1^{[2]} \cdots e_m^{[2]}\big]. \tag{2.8}$$

Proposition 2.1. *System* (1.1) *admits the representation:*

$$\begin{aligned} x(k+1) &= \widetilde{A}\big(\vartheta(k) \otimes x(k)\big) + \overline{B}(k)\vartheta(k) + \widetilde{F}\big(\vartheta(k) \otimes N(k)\big), \\ \vartheta(k+1) &= \vartheta(k), \\ y(k) &= \widetilde{C}\big(\vartheta(k) \otimes x(k)\big) + \overline{D}(k)\vartheta(k) + \widetilde{G}\big(\vartheta(k) \otimes N(k)\big), \end{aligned} \tag{2.9}$$

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

$$\overline{B}(k) = \widetilde{B}\big(I_m \otimes u(k)\big), \quad \overline{D}(k) = \widetilde{D}\big(I_m \otimes u(k)\big). \tag{2.10}$$

Proof. Using identities (2.6) the state and output equations of system (1.1) can be put in the form:

$$\begin{aligned} x(k+1) &= \widetilde{A}\big(\vartheta(k) \otimes I_n\big)x(k) + \widetilde{B}\big(\vartheta(k) \otimes I_p\big)u(k) + \widetilde{F}\big(\vartheta(k) \otimes I_b\big)N(k), \\ y(k) &= \widetilde{C}\big(\vartheta(k) \otimes I_n\big)x(k) + \widetilde{D}\big(\vartheta(k) \otimes I_p\big)u(k) + \widetilde{G}\big(\vartheta(k) \otimes I_b\big)N(k). \end{aligned} \tag{2.11}$$

According to the Kronecker product properties:

$$\begin{aligned} \big(\vartheta(k) \otimes I_n\big)x(k) &= \big(\vartheta(k) \otimes I_n\big) \cdot \big(1 \otimes x(k)\big) = \big(\vartheta(k) \cdot 1\big) \otimes \big(I_n \cdot x(k)\big) = \vartheta(k) \otimes x(k), \\ \big(\vartheta(k) \otimes I_p\big)u(k) &= \big(\vartheta(k) \otimes I_p\big) \cdot \big(1 \otimes u(k)\big) = \vartheta(k) \otimes u(k) = \big(I_m \cdot \vartheta(k)\big) \otimes \big(u(k) \cdot 1\big) \\ &= \big(I_m \otimes u(k)\big) \cdot \big(\vartheta(k) \otimes 1\big) = \big(I_m \otimes u(k)\big)\vartheta(k), \end{aligned} \tag{2.12}$$

so that (2.9) and (2.10) 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:

$$\hat{x}(k) = \mathbb{E}\big[x(k)|\sigma(Y_k)\big] = \Pi\big[x(k)|\mathcal{B}(Y_k)\big], \tag{3.1}$$

where $Y_k = \big[y^T(0) \quad \cdots \quad y^T(k)\big]^T$. 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 [3, 4]. Choosing an integer ν and assuming that, for all $h \in \mathbb{Z}^+$,

$$\mathbb{E}\big[\|y^{[i]}(h)\|^2\big] < \infty, \; i = 1, \ldots, 2\nu, \tag{3.2}$$

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

$$L(Y_k^\nu) = \text{span}\big\{1, \; Y^\nu(0), \; \cdots, \; Y^\nu(k)\big\}, \tag{3.3}$$

$$\text{with} \;\; Y_k^\nu = \begin{bmatrix} Y^\nu(0) \\ \vdots \\ Y^\nu(k) \end{bmatrix}, \quad Y^\nu(h) = \begin{bmatrix} y(h) \\ \vdots \\ y^{[\nu]}(h) \end{bmatrix}. \tag{3.4}$$

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

$$\begin{aligned} \hat{x}_\nu(k) &= \Pi\big[x(k)|L(Y_k^\nu)\big], \\ \hat{\vartheta}_\nu(k) &= \Pi\big[\vartheta(k)|L(Y_k^\nu)\big]. \end{aligned} \tag{3.5}$$

In order to ensure that all the moments in (3.2) are finite, the following assumptions are needed:
1) the noise variable $N(k)$ has finite moments up to degree 2ν

$$\mathbb{E}\big[N^{[j]}(k)\big] = \xi_j < \infty, \quad 1 \le j \le 2\nu, \tag{3.6}$$

(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:

$$\mathbb{E}\big[x_0^{[j]}\big] = \zeta_j < \infty, \quad 1 \le j \le 2\nu. \tag{3.7}$$

Assumptions 1) and 2) guarantee that the *polynomial extended output sequence* $Y^\nu(k)$, defined in (3.4), has bounded mean and covariance. Consider now the extended state sequence $X^\nu(k)$ defined as

$$X^\nu(k) = \begin{bmatrix} X_0(k) \\ \vdots \\ X_\nu(k) \end{bmatrix}, \quad X_i(k) = \vartheta(k) \otimes x^{[i]}(k), \tag{3.8}$$

(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

$$\begin{aligned} X^\nu(k+1) &= \mathbf{A}^\nu(k)X^\nu(k) + \mathbf{F}^\nu(N(k), X^\nu(k)), \\ Y^\nu(k) &= \mathbf{C}^\nu(k)X^\nu(k) + \mathbf{G}^\nu(N(k), X^\nu(k)), \end{aligned} \tag{3.9}$$

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 3.4, 3.5 and 3.7.

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 (3.9).

Theorem 3.1. *The optimal ν-degree polynomial estimate of the state $x(k)$ of system* (1.1) *and of the unknown vector $\vartheta(k)$ are given by:*

$$\begin{aligned} \hat{x}_\nu(k) &= \mathcal{M}_n\widehat{X}^\nu(k) = \mathcal{M}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big], \\ \hat{\vartheta}_\nu(k) &= \mathcal{T}_n\widehat{X}^\nu(k) = \mathcal{T}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big], \end{aligned} \tag{3.10}$$

where:

$$\begin{aligned} \mathcal{M}_n &= \big[O_{n\times m} \quad \mathcal{M} \quad O_{n\times m(n^2+\cdots+n^\nu)}\big], \\ \mathcal{T}_n &= \big[I_m \quad O_{m\times m(n+\cdots+n^\nu)}\big], \end{aligned} \tag{3.11}$$

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

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

$$\begin{aligned} x(k) &= \mathcal{M}\big(\vartheta(k) \otimes x(k)\big) = \mathcal{M}X_1(k) = \mathcal{M}_nX^\nu(k) \\ \vartheta(k) &= \mathcal{T}_nX^\nu(k), \end{aligned} \tag{3.12}$$

so that the polynomial minimum variance state estimates in (3.10) are:

$$\begin{aligned} \hat{x}_\nu(k) &= \Pi\big[x(k)|L(Y_k^\nu)\big] = \Pi\big[\mathcal{M}_nX^\nu(k)|L(Y_k^\nu)\big] = \mathcal{M}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big] = \mathcal{M}_n\widehat{X}^\nu(k), \\ \hat{\vartheta}_\nu(k) &= \Pi\big[\vartheta(k)|L(Y_k^\nu)\big] = \Pi\big[\mathcal{T}_nX^\nu(k)|L(Y_k^\nu)\big] = \mathcal{T}_n\Pi\big[X^\nu(k)|L(Y_k^\nu)\big] = \mathcal{T}_n\widehat{X}^\nu(k). \end{aligned} \tag{3.13}$$

■

Remark 3.2. 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:

$$\mathrm{Cov}\big(x(k)-\hat{x}_\nu(k)\big) = \mathcal{M}_n\mathrm{Cov}\big(X^\nu(k)-\widehat{X}^\nu(k)\big)\mathcal{M}_n^T. \tag{3.14}$$

Remark 3.3. 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:

$$\hat{\mu}(k) : \ \|e_{\hat{\mu}(k)} - \widehat{\theta}_\nu(k)\|_\infty \le \|e_j - \widehat{\theta}_\nu(k)\|_\infty; \tag{3.15}$$

for $j = 1, \dots, m$. The motivation for this strategy is that the choice (3.15), when applied to the conditional expectation of $\vartheta(k)$, provides the Maximum Likelihood Estimate of $\mu(k)$. This happens because the components of $\widehat{\vartheta}(k) = I\!E\{\vartheta(k)|\sigma(Y_k)\}$ coincide with the conditional distribution of $\vartheta(k)$.

The following lemmas give some insights into the structure and properties of the model (3.9). All the results presented exploit the fact that, according to definition (3.8) and to identity (2.8), $\forall i, j, h \in \mathbb{Z}^+$:

$$X_j^{[h]} = \Theta_n^{h,j} X_{jh}, \quad X_i \otimes X_j = \Xi_{i,j} X_{i+j}, \tag{3.16}$$

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

$$\begin{aligned} \Theta_n^{h+1,j} &= \big(\Theta_n^{h,j} \otimes I_{mn^j}\big)\big(I_m \otimes C^T_{mn^j,n^{jh}}\big) \cdot \big(E_2 \otimes I_{n^{j(h+1)}}\big), \\ \Theta_n^{0,j} &= [1 \cdots 1] \in \mathbb{R}^{1\times m}, \\ \Xi_{i,j} &= \big(I_m \otimes C^T_{mn^j,n^i}\big)(E_2 \otimes I_{n^{i+j}}), \end{aligned} \tag{3.17}$$

with $C_{a,b}$ suitably dimensioned commutation matrices for the Kronecker product [4], and E_2 as in (2.8) (see [5] for more details).

Lemma 3.4. *The iterative equation of the component $X_j(k)$ defined in (3.8) is:*

$$\begin{aligned} X_j(k+1) &= \sum_{t_1=0}^{j} \mathbf{A}_{j,t_1}(k) X_{t_1}(k) + \mathcal{F}_j(k), \\ \mathcal{F}_j(k) &= \sum_{t_1=0}^{j} S_{t_1}^j(k) X_{t_1}(k), \end{aligned} \tag{3.18}$$

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

$$\mathbf{A}_{j,t_1}(k) = \big(I_m \otimes J_{t_1}^j(k)\big)\Xi_{0,t_1}, \tag{3.19}$$

$$S_{t_1}^j(k) = \big(I_m \otimes \mathcal{L}_{t_1}^j(k)\big)\Xi_{0,t_1}, \tag{3.20}$$

with:

$$J_{t_1}^j(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} L_t^j(k)\Big(I_{mn^{t_1}} \otimes \xi_{t_3}(k)\Big), \tag{3.21}$$

$$\mathcal{L}_{t_1}^j(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} L_t^j(k)\Big(I_{mn^{t_1}} \otimes \big(N^{[t_3]}(k) - \xi_{t_3}(k)\big)\Big), \tag{3.22}$$

$$L_t^j(k) = M_t^j\left(\widetilde{A}^{[t_1]} \otimes \overline{B}^{[t_2]}(k) \otimes \widetilde{F}^{[t_3]}\right) K_t^j, \tag{3.23}$$

$$K_t^j = \big(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\big)\big(I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}\big). \tag{3.24}$$

M_t^j in (3.23) are the matrix coefficients for the Kronecker power expansion [4], $t = (t_1, t_2, t_3)^T$ is a multi-index in $(\mathbb{Z}^+)^3$ and $\mathcal{R}_j = \{t \in (\mathbb{Z}^+)^3 : t_1 + t_2 + t_3 = j\}$. Moreover $\mathcal{F}(k) = [\mathcal{F}_0(k)^T \cdots \mathcal{F}_\nu(k)^T]^T$ is a sequence of zero-mean uncorrelated random vectors, whose covariance matrices $\Psi_{j,i}^{\mathcal{F}}(k) = \mathbb{E}[\mathcal{F}_j(k)\mathcal{F}_i(k)^T]$ are given by:

$$\Psi_{j,i}^{\mathcal{F}}(k) = \sum_{j_1=0}^{j}\sum_{i_1=0}^{i} \mathrm{st}^{-1}_{mn^{j_1},mn^{i_1}}\left(\Phi_{r_1,t_1}^{S,i,j}(k) \cdot \Xi_{r_1,t_1}\mathbb{E}\big[X_{r_1+t_1}(k)\big]\right), \tag{3.25}$$

with st^{-1} *the inverse of the stack operator [4] and:*

$$\begin{aligned}\Phi^{S,i,j}_{r_1,t_1}(k) =& I\!E\big[S^i_{r_1}(k)\otimes S^j_{t_1}(k)\big]\\ =&\big(I_m\otimes C^T_{mn^j,n^i}\big)\big(I_{m^2}\otimes \Phi^{\mathcal{L},j,i}_{t_1,r_1}(k)\big)\cdot\big(I_m\otimes C_{m^2n^{t_1},mn^{r_1}}\big)\big(\Xi_{0,r_1}\otimes\Xi_{0,t_1}\big),\end{aligned} \tag{3.26}$$

where:

$$\begin{aligned}\Phi^{\mathcal{L},j,i}_{t_1,r_1}(k) =& I\!E\big[\mathcal{L}^j_{t_1}(k)\otimes\mathcal{L}^i_{r_1}(k)\big]\\ =&\sum_{t_2,t_3}^{t\in\mathcal{R}_j}\sum_{r_2,r_3}^{r\in\mathcal{R}_i}\big(L^j_t(k)\otimes L^i_r(k)\big)\big(I_{mn^{t_1}}\otimes C^T_{mn^{r_1}b^{r_3},b^{t_3}}\big)\\ &\cdot\Big(I_{m^2n^{t_1+r_1}}\otimes\big(\xi_{r_3+t_3}(k)-\xi_{r_3}(k)\otimes\xi_{t_3}(k)\big)\Big)\big(I_{mn^{t_1}}\otimes C_{mn^{r_1},1}\big).\end{aligned} \tag{3.27}$$

Proof. The proof is a straightforward consequence of Lemma 3.2 in [5]. In that framework system (1.1) is a switching system, and $\mu(k)$ is a Markov chain with known transition probability matrix. The iterative equation (3.18) 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 $\{\mathcal{F}(k)\}$ 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 3.5. *The equations for the Kronecker powers of the measurements defined in* (3.4) *are:*

$$\begin{aligned}y^{[j]}(k) &= \sum_{t_1=0}^{j}\mathbf{C}_{j,t_1}(k)X_{t_1}(k)+\mathcal{G}_j(k),\\ \mathcal{G}_j(k) &= \sum_{t_1=0}^{j}\mathcal{T}^j_{t_1}(k)X_{t_1}(k),\end{aligned} \tag{3.28}$$

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

$$\mathbf{C}_{j,t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j}T^j_t(k)\Big(I_{mn^{t_1}}\otimes\xi_{t_3}(k)\Big), \tag{3.29}$$

$$\mathcal{T}^j_{t_1}(k) = \sum_{t_2,t_3}^{t\in\mathcal{R}_j}T^j_t(k)\Big(I_{mn^{t_1}}\otimes\big(N^{[t_3]}(k)-\xi_{t_3}(k)\big)\Big), \tag{3.30}$$

$$T^j_t(k) = M^j_t\Big(\widetilde{C}^{[t_1]}\otimes\overline{D}^{[t_2]}(k)\otimes\widetilde{G}^{[t_3]}\Big)K^j_t, \tag{3.31}$$

and K^j_t *as in* (3.24). $\mathcal{G}(k)=\big[\mathcal{G}_1(k)^T\dots\ \mathcal{G}_\nu(k)^T\big]^T$ *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:*

$$\Psi^{\mathcal{G}}_{j,i}(k) = \sum_{t_1=0}^{j}\sum_{r_1=0}^{i}\mathrm{st}^{-1}_{q^j,q^i}\Big(\Phi^{\mathcal{T},i,j}_{r_1,t_1}(k)\Xi_{r_1,t_1}I\!E\big[X_{r_1+t_1}(k)\big]\Big), \tag{3.32}$$

with:

$$\Phi^{\mathcal{T},i,j}_{r_1,t_1}(k) = \mathbb{E}\big[\mathcal{T}^i_{r_1}(k)\otimes\mathcal{T}^j_{t_1}(k)\big] = \sum_{r_2,r_3}^{r\in\mathcal{R}_i}\sum_{t_2,t_3}^{t\in\mathcal{R}_j}\big(T^i_r(k)\otimes T^j_t(k)\big)\big(I_{mn^{r_1}}\otimes C^T_{mn^{t_1}b^{t_3},b^{r_3}}\big) \cdot\Big(I_{m^2n^{r_1+t_1}}\otimes\big(\xi_{t_3+r_3}(k)-\xi_{t_3}(k)\otimes\xi_{r_3}(k)\big)\Big)\big(I_{mn^{r_1}}\otimes C_{mn^{t_1},1}\big). \tag{3.33}$$

Proof. The proof is a straightforward consequence of Lemma 3.3 in [5], according to the same remarks considered in the proof of Lemma 3.4. ■

Remark 3.6. It has to be stressed that, according to Lemmas 3.4 and 3.5, system (3.9) provides an *exact* generation model for the sequences $X^\nu(k)$ and $Y^\nu(k)$ (i.e. no approximation has been introduced).

Lemma 3.7. *The noise sequences* $\{\mathcal{F}(k)\}$ *and* $\{\mathcal{G}(k)\}$ *are such that, for* $1\le i,j\le\nu$ *and* $\forall k,h\in\mathbb{Z}^+$*:*

$$\begin{aligned}\mathbb{E}\big[\mathcal{F}_j(k)\mathcal{G}^T_i(h)\big]&=0, \qquad \forall k\neq h\\ \mathbb{E}\big[\mathcal{F}_j(k)\mathcal{G}^T_i(k)\big]&=Q_{j,i}(k).\end{aligned} \tag{3.34}$$

with:

$$Q_{j,i}(k)=\sum_{t_1=0}^{j}\sum_{r_1=0}^{i}\mathrm{st}^{-1}_{mn^j,q^i}\Big(\mathcal{Q}^{i,j}_{r_1,t_1}(k)\Xi_{r_1,t_1}\mathbb{E}\big[X_{r_1+t_1}(k)\big]\Big), \tag{3.35}$$

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

Proof. The proof is a straightforward consequence of Lemma 3.4 in [5], according to the same remarks considered in the proof of Lemma 3.4. ■

According to Lemma 3.7, 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 (3.9) [2]; the straightforward algorithm is the following:

$$\begin{aligned}&\widehat{X}^\nu(0|-1)=\mathbb{E}\big[X^\nu(0)\big],\\ &\widehat{X}^\nu(k)=\widehat{X}^\nu(k|k-1)+\mathcal{K}(k)\big(Y^\nu(k)-\mathbf{C}^\nu(k)\widehat{X}^\nu(k|k-1)\big),\\ &\widehat{X}^\nu(k+1|k)=\mathbf{A}^\nu(k)\widehat{X}^\nu(k)+\mathcal{Z}(k)\big(Y^\nu(k)-\mathbf{C}^\nu(k)\widehat{X}^\nu(k|k-1)\big).\end{aligned} \tag{3.36}$$

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

$$P_P(0)=\mathrm{Cov}\big(X^\nu(0)\big), \tag{3.37}$$

$$\mathcal{Z}(k)=Q(k)\big(\mathbf{C}^\nu(k)P_P(k)\mathbf{C}^{\nu T}(k)+\Psi^{\mathcal{G}}(k)\big)^\dagger, \tag{3.38}$$

$$\mathcal{K}(k)=P_P(k)\mathbf{C}^{\nu T}(k)\big(\mathbf{C}^\nu(k)P_P(k)\mathbf{C}^{\nu T}(k)+\Psi^{\mathcal{G}}(k)\big)^\dagger, \tag{3.39}$$

$$P(k)=P_P(k)-\mathcal{K}(k)\mathbf{C}^\nu(k)P_P(k), \tag{3.40}$$

$$P_P(k+1)=\mathbf{A}^\nu(k)P(k)\mathbf{A}^{\nu T}(k)+\Psi^{\mathcal{F}}(k)-\mathcal{Z}(k)Q^T(k)-\mathbf{A}^\nu(k)\mathcal{K}(k)Q^T(k) \tag{3.41}$$

$$-Q(k)\mathcal{K}^T(k)\mathbf{A}^{\nu T}(k), \tag{3.42}$$

where in (3.42) the Moore-Penrose pseudoinverse has been used.

Remark 3.8. 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 (3.7).

Remark 3.9. Note that the recursive computation of $\Psi^{\mathcal{F}}(k)$, $\Psi^{\mathcal{G}}(k)$ and $Q(k)$ requires the computation of the expectations $I\!E\big[X_i(k)\big]$, $i = 1,\ldots,2\nu$ (see (3.25), (3.32) and (3.35)). These are the components of $I\!E\big[X^{\nu}(k)\big]$, and are recursively computed as

$$I\!E\big[X^{\nu}(k+1)\big] = \mathbf{A}^{2\nu}(k) I\!E\big[X^{\nu}(k)\big]. \tag{3.43}$$

4. Simulation results

This section reports simulation results referred to a system of the type (1.1), characterized by the following data:

- $x(k)\in I\!R^3$, $u(k)\in I\!R$, $y(k)\in I\!R^2$, $\mathcal{W}=\{1,2\}$;
- $A_1=\begin{bmatrix} 0.5 & 0 & 0.2 \\ -1.75 & 0.5 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}$, $B_1 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}$, $C_1 = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \end{bmatrix}$, $D_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$,
- $A_2=\begin{bmatrix} 0.3 & 0.25 & 0.1 \\ -1.75 & 0.5 & 0 \\ 0 & 1.2 & 1 \end{bmatrix}$; $B_2 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$, $C_2 = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, $D_2 = \begin{bmatrix} 0.5 \\ -0.2 \end{bmatrix}$;
- $F_1 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0.1 & 0 & 0 \\ 0 & 0.3 & 0 \end{bmatrix}$, $G_1 = \begin{bmatrix} 0 & 0 & 0.1 \\ 0 & 0 & -0.3 \end{bmatrix}$,
- $F_2 = \begin{bmatrix} 0 & 0.1 & 0 \\ 0.1 & 0.1 & 0 \\ -0.1 & 0 & 0 \end{bmatrix}$; $G_2 = \begin{bmatrix} 0 & 0 & 0.2 \\ 0 & 0 & 0.1 \end{bmatrix}$;
- the noise $N(k) \in I\!R^3$ has independent components, with distributions:

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

The distribution of N_3 is identical to that of N_1.
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 $k = 500$.

Figures 1–3 report the 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

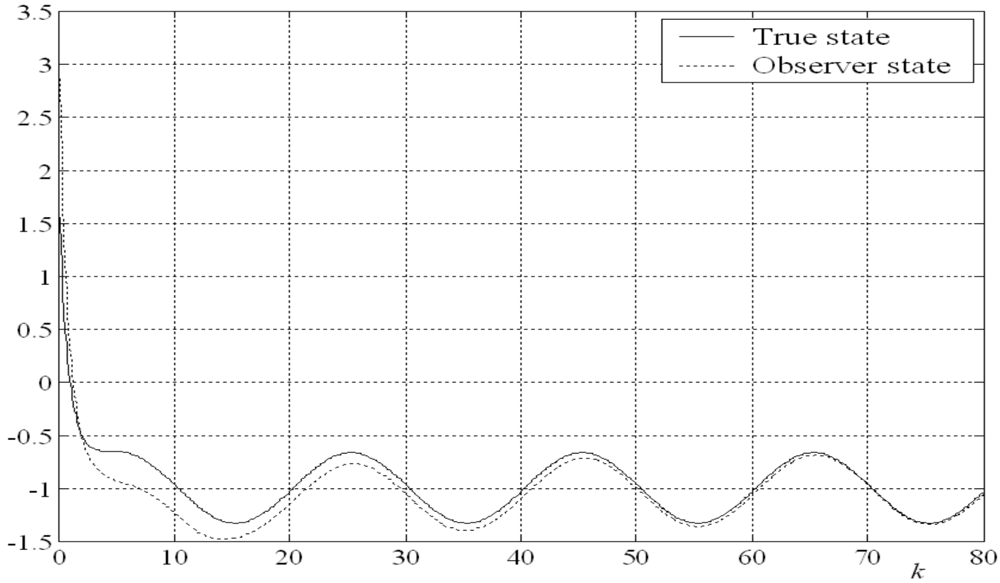

Figure 1: **True and estimated** $x_1(k)$**.**

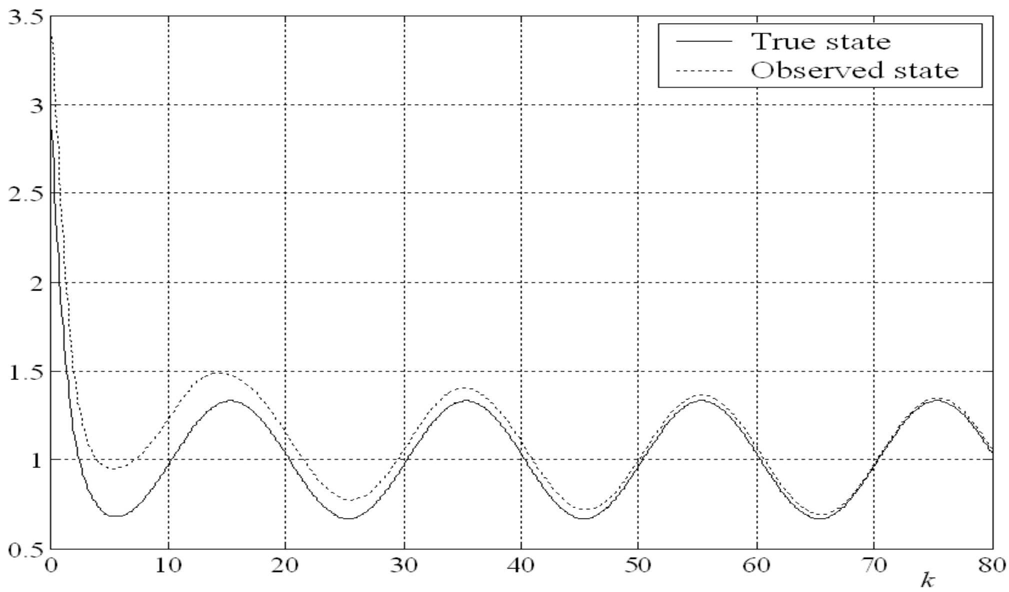

Figure 2: **True and estimated state** $x_2(k)$**.**

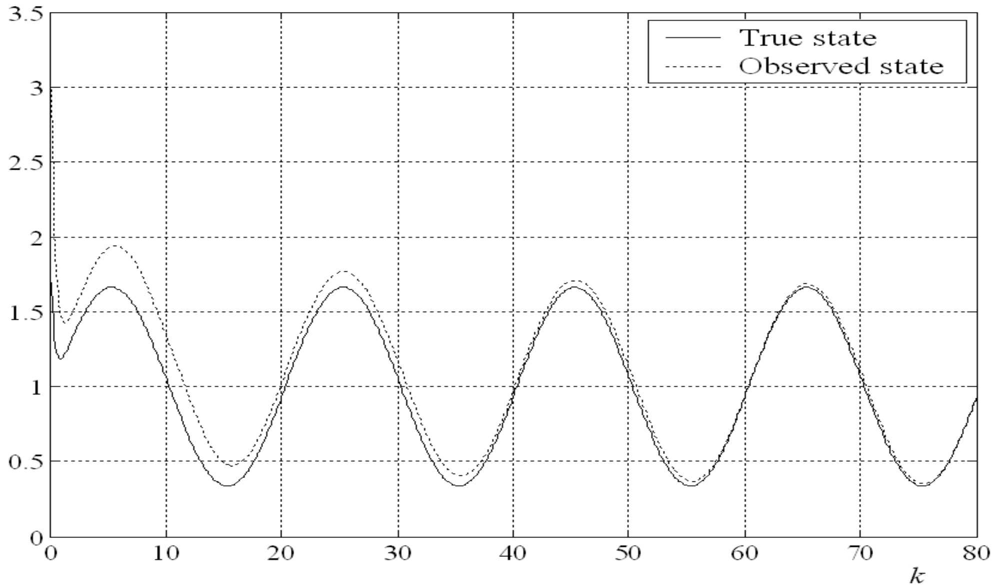

Figure 3: **True and estimated state $x_3(k)$.**

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 sensible improvement of the second order filter with respect to the first order one, proving the goodness of the developed theory.

A. Kronecker Algebra

For the ease of the reader, in this Appendix are reported some useful results on the Kronecker algebra. The proofs and other further details can be found in [1] and [4]. Let M and N be matrices of dimensions $r \times s$ and $p \times q$ respectively, then the Kronecker product $M \otimes N$ is defined as the $(r \cdot p) \times (s \cdot q)$ matrix

$$M \otimes N = \begin{bmatrix} m_{11}N & \dots & m_{1s}N \\ \vdots & \ddots & \vdots \\ m_{r1}N & \dots & m_{rs}N \end{bmatrix}, \tag{A.1}$$

where the m_{ij} are the entries of M.

Definition A.1. Let M be an $r \times s$ matrix:

$$M = \begin{bmatrix} m_1 & m_2 & \dots & m_s \end{bmatrix}, \tag{A.2}$$

where m_i denotes the i-th column of M. The stack of M is defined as the $r \cdot s$ vector:

$$M = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{bmatrix}. \tag{A.3}$$

Observe that a vector as in (A.3) can be reduced to a matrix M as in (A.2), once it is known the number of the rows r of the original matrix, by considering the inverse operation of the stack denoted by st^{-1}. More generally, let m be a vector in $I\!R^{\mu}$, and r be a divisor of μ. Then the $r \times (\mu/r)$ matrix given by $M = \mathrm{st}^{-1}(m, r)$ is defined so that:

$$\mathrm{st}(M) = m. \tag{A.4}$$

In presence of vectors $m \in I\!R^{(\mu^2)}$, that is their length is given by a square, the notation $\mathrm{st}^{-1}(m)$ has to be considered as a short version of $\mathrm{st}^{-1}(m, \mu)$.

In case of vectors Kronecker products, it is easy to verify that, if $u \in I\!R^r$ and $v \in I\!R^s$, the i-th entry of $u \otimes v$ is given by

$$(u \otimes v)_i = u_l \cdot v_m; \quad l = \left[\frac{i-1}{s}\right] + 1, \quad m = |i-1|_s + 1, \tag{A.5}$$

where $[\cdot]$ and $|\cdot|_s$ denote integer part and s-modulo respectively. Moreover, the Kronecker power of M is defined as

$$M^{[0]} = 1 \in I\!R, \tag{A.6a}$$

$$M^{[l]} = M \otimes M^{[l-1]} \quad l \geq 1. \tag{A.6b}$$

Some useful properties of the Kronecker product and stack operation are the followings:

$$(A+B)\otimes(C+D) = A\otimes C + A\otimes D + B\otimes C + B\otimes D \tag{A.7a}$$

$$A\otimes(B\otimes C) = (A\otimes B)\otimes C \tag{A.7b}$$

$$(A\cdot C)\otimes(B\cdot D) = (A\otimes B)\cdot(C\otimes D) \tag{A.7c}$$

$$(A\otimes B)^T = A^T\otimes B^T \tag{A.7d}$$

$$\mathrm{st}(A\cdot B\cdot C) = (C^T\otimes A)\cdot st(B) \tag{A.7e}$$

$$u\otimes v = \mathrm{st}(v\cdot u^T) \tag{A.7f}$$

$$\mathrm{tr}(A\otimes B) = \mathrm{tr}(A)\cdot \mathrm{tr}(B) \tag{A.7g}$$

Other useful properties can be found in [1].

A generalized version of (A.7c), often used throughout the paper is the following:

$$(A_1\cdot B_1)\otimes(A_2\cdot B_2)\otimes(A_3\cdot B_3) = (A_1\otimes A_2\otimes A_3)\cdot(B_1\otimes B_2\otimes B_3). \tag{A.8}$$

According to its definition (A.1), the Kronecker product is not commutative. However, the following result holds:

Lemma A.2. *For any given pair of matrices $A\in I\!R^{r\times s}$, $B\in I\!R^{n\times m}$, it is:*

$$B\otimes A = C_{r,n}^T(A\otimes B)C_{s,m}, \tag{A.9}$$

where $C_{r,n}$, $C_{s,m}$ are defined so that, denoted $\{C_{u,v}\}_{h,l}$ their (h,l) entries:

$$\{C_{u,v}\}_{h,l} = \begin{cases} 1, & \text{if } l = (|h-1|_v)u + \left(\left[\frac{h-1}{v}\right]+1\right); \\ 0, & \text{otherwise.}\end{cases} \tag{A.10}$$

Proposition A.3. *For any given matrices A,B,C,D, having dimensions $n_A\times m_A$, $n_B\times m_B$, $n_C\times m_C$, $n_D\times m_D$ respectively:*

$$A\otimes B\otimes C\otimes D = \left(I_{n_A}\otimes C_{n_Cn_D,n_B}^T\right)(A\otimes C\otimes D\otimes B)\left(I_{m_A}\otimes C_{m_Cm_D,m_B}\right). \tag{A.11}$$

Proof.
By applying property (A.7b), (A.7c) and lemma A.2:

$$\begin{aligned} A\otimes B\otimes C\otimes D &= \Big(A\otimes\big(B\otimes(C\otimes D)\big)\Big)\\ &= \Big(A\otimes\big(C_{n_Cn_D,n_B}^T(C\otimes D\otimes B)C_{m_Cm_D,m_B}\big)\Big)\\ &= \big(I_{n_A}\otimes C_{n_Cn_D,n_B}^T\big)\Big(A\otimes\big((C\otimes D\otimes B)C_{m_Cm_D,m_B}\big)\Big)\\ &= \big(I_{n_A}\otimes C_{n_Cn_D,n_B}^T\big)\big(A\otimes C\otimes D\otimes B\big)\big(I_{m_A}\otimes C_{m_Cm_D,m_B}\big). \end{aligned} \tag{A.12}$$

■

Remark A.4. Observe that $C_{1,1}=1$, hence in the vector case when $a\in I\!R^r$ and $b\in I\!R^n$, (A.9) becomes

$$b\otimes a = C_{r,n}^T(a\otimes b). \tag{A.13}$$

Moreover, in the vector case the commutation matrices satisfy also the following recursive formula.

Lemma A.5. *Let $a, b \in I\!R^n$ and $l \in I\!N$. Then*

$$b^{[l]} \otimes a = G_l(n)(a \otimes b^{[l]}), \tag{A.14}$$

with the sequence $\{G_l(n) = C^T_{n,n^l}\}$ given by the following recursive equations

$$G_1(n) = C^T_{n,n}, \tag{A.15a}$$
$$G_l(n) = (I_{n,1} \otimes G_{l-1}(n)) \cdot (G_1(n) \otimes I_{n,l-1}), \qquad l > 1, \tag{A.15b}$$

where $I_{n,r}$ is the identity matrix in $I\!R^{n^r}$.

A binomial formula can be found for the Kronecker power, which generalizes the classical Newton one.

Lemma A.6. *Let $a, b \in I\!R^n$. For any integer $h \geq 0$ the matrix coefficients of the following binomial power formula:*

$$(a+b)^{[h]} = \sum_{k=0}^{h} M^h_k(n)(a^{[k]} \otimes b^{[h-k]}) \tag{A.16}$$

constitute a set of matrices $\{M^h_0(n), ..., M^h_h(n); \quad M^h_k(n) \in I\!R^{n^h \times n^h}\}$ such that:

$$M^h_h(n) = M^h_0(n) = I_{n,h}, \tag{A.17a}$$
$$M^h_j(n) = (M^{h-1}_j(n) \otimes I_{n,1}) + (M^{h-1}_{j-1}(n) \otimes I_{n,1}) \cdot (I_{n,j-1} \otimes G_{h-j}(n)), \quad 1 \leq j \leq h-1, \tag{A.17b}$$

where $G_l(n)$ and $I_{n,l}$ are as in Lemma A.4.

Lemma A.6 can also be generalized to the polynomial case. Obviously, given any polynomial $a_1 + ... + a_p$, $a_i \in I\!R^n$, $1 \leq i \leq p$, $p \in I\!N$, its h-th Kronecker power admits a representation as:

$$(a_1 + a_2 + ... + a_p)^{[h]} = \sum_{\substack{h_1,...,h_p \geq 0 \\ h_1+...+h_p = h}} M^h_{h_1,...,h_p}\big(a_1^{[h_1]} \otimes a_2^{[h_2]} \otimes \cdots \otimes a_p^{[h_p]}\big) \tag{A.18}$$

where $M^h_{h_1,...,h_p}$ are suitable matrices. The definition of symbols $M^l_{l_1,...,l_s}$ is extended, with $l > 0$ when at least one of the l_i's is negative, as

$$M^l_{l_1,...,l_s} = O_{n^l \times n^l}. \tag{A.19}$$

Moreover the following statement can be proved:

Lemma A.7. *The matrices $M^h_{h_1,...,h_p} \in I\!R^{n^h \times n^h}$ in* (A.18) *satisfy the recursive formula:*

$$M^h_{h_1,...,h_p} = I_1, \qquad h = 1 \tag{A.20a}$$
$$M^h_{h_1,...,h_p} = \sum_{1 \leq i \leq p-1} \big(M^{h-1}_{h_1,...,h_i-1,...,h_p} \otimes I_1\big)\big(I_{h_1+...+h_i-1} \otimes G_{h_{i+1}+...+h_p}\big) + M^{h-1}_{h_1,...,h_p-1} \otimes I_1, \qquad h > 1. \tag{A.20b}$$

Proof. The proof can be found in [4]. ∎

B. Extended matrices derivation

The appendix contains two Lemmas showing the computation of the matrices used in equations (3.16) and in Lemmas 3.4,3.5 and 3.7.

Lemma B.1. *Let $x \in I\!R^n$ and $X_j = \vartheta \otimes x^{[j]}$, with ϑ taking values in the natural basis of $I\!R^m$. Then $\forall i, j, h \in I\!N$:*

$$X_j^{[h]} = \Theta_n^{h,j} X_{jh}, \qquad X_i \otimes X_j = \Xi_{i,j} X_{i+j}, \tag{B.1}$$

with

$$\begin{aligned} \Theta_n^{h+1,j} &= \big(\Theta_n^{h,j} \otimes I_{mn^j}\big)\big(I_m \otimes C^T_{mn^j,n^{jh}}\big)\big(E_2 \otimes I_{n^{j(h+1)}}\big), \qquad h > 0 \\ \Theta_n^{0,j} &= [1 \cdots 1]. \end{aligned} \tag{B.2}$$

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

Proof. The first equation in (B.1) is proved by induction: it is clearly true for $h = 0$. Let it be true for $h = k$. Then:

$$\begin{aligned} X_j^{[k+1]} &= X_j^{[k]} \otimes X_j = \big(\Theta_n^{k,j} X_{jk}\big) \otimes X_j = \big(\Theta_n^{k,j} \otimes I_{mn^j}\big)\big(\vartheta \otimes x^{[jk]} \otimes \vartheta \otimes x^{[j]}\big) \\ &= \big(\Theta_n^{k,j} \otimes I_{mn^j}\big)\big(I_m \otimes C^T_{mn^j,n^{jk}}\big)\big(\vartheta^{[2]} \otimes x^{[j(k+1)]}\big) \\ &= \big(\Theta_n^{k,j} \otimes I_{mn^j}\big)\big(I_m \otimes C^T_{mn^j,n^{jk}}\big)\big(E_2 \otimes I_{n^{j(k+1)}}\big) X_{j(k+1)} = \Theta_n^{k+1,j} X_{j(k+1)}. \end{aligned} \tag{B.3}$$

The second equation in (B.1), easily comes:

$$X_i \otimes X_j = \vartheta \otimes x^{[i]} \otimes \vartheta \otimes x^{[j]} = \big(I_m \otimes C^T_{mn^j,n^i}\big)\big(\vartheta^{[2]} \otimes x^{[i+j]}\big) = \big(I_m \otimes C^T_{mn^j,n^i}\big)\big(E_2 \otimes I_{n^{i+j}}\big) X_{i+j}. \tag{B.4}$$

■

Lemma B.2. *Let ϑ be a random vector taking values in the natural basis of $I\!R^m$, N be a random vector taking values in $I\!R^b$, with finite and available moments, named:*

$$I\!E\big[N^{[j]}\big] = \xi_j, \qquad j \in I\!N, \tag{B.5}$$

and $z \in I\!R^n$, $v \in I\!R^p$ random vectors such that:

$$v = \Gamma_1\big(\vartheta \otimes z\big) + \Gamma_0 \vartheta + \widetilde{\Gamma}\big(\vartheta \otimes N\big) \tag{B.6}$$

with Γ_1, Γ_0, $\widetilde{\Gamma}$ matrices of suitable dimensions. Moreover, suppose that $\big\{\vartheta, N, z\big\}$ is a triple of independent random vectors. Then, for each $j \in I\!N$:

$$v^{[j]} = \sum_{t_1=0}^{j} \mathcal{H}^j_{t_1} Z_{t_1} + w_j, \qquad w_j = \sum_{t_1=0}^{j} \mathcal{W}^j_{t_1} Z_{t_1}, \tag{B.7}$$

with $Z_j = \vartheta \otimes z^{[j]}$ and w_j zero-mean random vectors. Matrices $\mathcal{H}^j_{t_1}$, $\mathcal{W}^j_{t_1}$, deterministic and random respectively, are given by:

$$\mathcal{H}^j_{t_1} = \sum_{t_2,t_3}^{t \in \mathcal{R}_j} W^j_t \big(I_{mn^{t_1}} \otimes \xi_{t_3}\big), \qquad \mathcal{W}^j_{t_1} = \sum_{t_2,t_3}^{t \in \mathcal{R}_j} W^j_t \Big(I_{mn^{t_1}} \otimes \big(N^{[t_3]} - \xi_{t_3}\big)\Big), \tag{B.8}$$

with $t = (t_1, t_2, t_3)^T$ *a multi-index in* $I\!N^3$ *and* $\mathcal{R}_j = \{t \in I\!N^3 : \ t_1 + t_2 + t_3 = j\}$. *The matrices* W_t^j *in* (B.8) *are defined by*

$$W_t^j = M_t^j \big(\Gamma_1^{[t_1]} \otimes \Gamma_0^{[t_2]} \otimes \widetilde{\Gamma}^{[t_3]}\big) K_t^j, \tag{B.9a}$$

$$K_t^j = \big(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\big)\big(I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}\big)\big(\Xi_{t_1,0} \otimes I_{b^{t_3}}\big). \tag{B.9b}$$

M_t^j *are the matrix coefficients for the polynomial Kronecker power expansion [4]. Moreover, the second order moments of* $\mathcal{W}_{t_1}^j(k)$ *are the following:*

$$\begin{aligned} \Phi_{t_1,r_1}^{\mathcal{W},j,i}(k) = I\!E\big[\mathcal{W}_{t_1}^j \otimes \mathcal{W}_{r_1}^i(k)\big] = \sum_{t_2,t_3}^{t\in\mathcal{R}_j} \sum_{r_2,r_3}^{r\in\mathcal{R}_i} \big(W_t^j \otimes W_r^i\big)\big(I_{mn^{t_1}} \otimes C^T_{mn^{r_1}b^{r_3},b^{t_3}}\big) \\ \cdot \Big(I_{m^2n^{t_1+r_1}} \otimes \big(\xi_{r_3+t_3} - \xi_{r_3} \otimes \xi_{t_3}\big)\Big)\big(I_{mn^{t_1}} \otimes C_{mn^{r_1},1}\big) \end{aligned} \tag{B.10}$$

Matrices $C_{a,b}$ *are the commutation matrices for a Kronecker product (see [4]).*

Proof. Applying the Newton formula to the Kronecker powers [4], it comes:

$$\begin{aligned} v^{[j]} &= \Big(\Gamma_1 Z_1 + \Gamma_0 Z_0 + \widetilde{\Gamma}(\vartheta \otimes N)\Big)^{[j]} \\ &= \sum_{t_1,t_2,t_3}^{t\in\mathcal{R}_j} M_t^j \bigg(\big(\Gamma_1^{[t_1]} Z_1^{[t_1]}\big) \otimes \big(\Gamma_0^{[t_2]} Z_0^{[t_2]}\big) \otimes \Big(\widetilde{\Gamma}^{[t_3]}\big(\vartheta \otimes N\big)^{[t_3]}\Big)\bigg) \\ &= \sum_{t_1,t_2,t_3}^{t\in\mathcal{R}_j} M_t^j \big(\Gamma_1^{[t_1]} \otimes \Gamma_0^{[t_2]} \otimes \widetilde{\Gamma}^{[t_3]}\big)\Big(Z_1^{[t_1]} \otimes Z_0^{[t_2]} \otimes \big(\vartheta \otimes N\big)^{[t_3]}\Big). \end{aligned} \tag{B.11}$$

By using equations (3.16) and the Kronecker properties [4], the last factor in the sum (B.11) becomes:

$$\begin{aligned} Z_1^{[t_1]} \otimes Z_0^{[t_2]} \otimes \big(\vartheta \otimes N\big)^{[t_3]} &= \big(\Theta_n^{t_1,1} Z_{t_1}\big) \otimes \big(\Theta_n^{t_2,0} Z_0\big) \otimes \Big(\Theta_b^{t_3,1}\big(\vartheta \otimes N^{[t_3]}\big)\Big) \\ &= \big(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\big)\big(Z_{t_1} \otimes \vartheta^{[2]} \otimes N^{[t_3]}\big) \\ &= \big(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\big)\big(I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}\big)\big(Z_{t_1} \otimes Z_0 \otimes N^{[t_3]}\big) \\ &= \big(\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}\big)\big(I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}\big)\big(\Xi_{t_1,0} \otimes I_{b^{t_3}}\big)\big(Z_{t_1} \otimes N^{[t_3]}\big) \\ &= K_t^j\Big(Z_{t_1} \otimes \big(\xi_{t_3} + N^{[t_3]} - \xi_{t_3}\big)\Big) = K_t^j\big(Z_{t_1} \otimes \xi_{t_3}\big) + K_t^j\Big(Z_{t_1} \otimes \big(N^{[t_3]} - \xi_{t_3}\big)\Big) \\ &= K_t^j\big(I_{mn^{t_1}} \otimes \xi_{t_3}\big)Z_{t_1} + K_t^j\Big(I_{mn^{t_1}} \otimes \big(N^{[t_3]} - \xi_{t_3}\big)\Big)Z_{t_1}, \end{aligned} \tag{B.12}$$

so that, substituting (B.12) in (B.11), by using (B.8) and (B.9b), equations (B.7) come. According to the independence of N and Z_{t_1}, note that w_j is a zero-mean random vector. Equation (B.10) is below obtained, by using the commutation formula for the Kronecker products [4]:

$$\begin{aligned} \Phi_{t_1,r_1}^{\mathcal{W},j,i} &= \sum_{t_2,t_3}^{t\in\mathcal{R}_j} \sum_{r_2,r_3}^{r\in\mathcal{R}_i} \big(W_t^j \otimes W_r^i\big) I\!E\Big[I_{mn^{t_1}} \otimes \big(N^{[t_3]} - \xi_{t_3}\big) \otimes I_{mn^{r_1}} \otimes \big(N^{[r_3]} - \xi_{r_3}\big)\Big] \\ &= \sum_{t_2,t_3}^{t\in\mathcal{R}_j} \sum_{r_2,r_3}^{r\in\mathcal{R}_i} \big(W_t^j \otimes W_r^i\big)\big(I_{mn^{t_1}} \otimes C^T_{mn^{r_1}b^{r_3},b^{t_3}}\big) \\ &\quad \cdot \bigg(I_{m^2n^{t_1+r_1}} \otimes I\!E\Big[\big(N^{[r_3]} - \xi_{r_3}\big) \otimes \big(N^{[t_3]} - \xi_{t_3}\big)\Big]\bigg)\big(I_{mn^{t_1}} \otimes C_{mn^{r_1},1}\big). \end{aligned} \tag{B.13}$$

By writing explicitly the mean value, the lemma is proved. ■

References

[1] R. Bellman, *Introduction to matrix analysis*, McGraw-Hill, 1970.

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

[3] 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.

[4] 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.

[5] A. Germani, C. Manes and P. Palumbo, "Polynomial Filtering for Stochastic Systems with Markovian Switching Coefficients," proc. of 42^{th} *Conf. on Decision and Control*, Maui, Hawaii, USA, 2003.

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

[7] 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.

[8] 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.

[9] 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.

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

[11] 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.

[12] A. Germani, C. Manes, P. Palumbo, "Polynomial filtering for stochastic non-Gaussian descriptor systems", *to appear on IEEE Trans. Circ. and Syst. II*, July, 2004.