



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

A. Germani, C. Manes, P. Palumbo

**POLYNOMIAL FILTERING FOR STOCHASTIC
SYSTEMS WITH MARKOVIAN SWITCHING
COEFFICIENTS**

R. 570 Giugno 2002

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

In this paper the suboptimal polynomial approach is followed to solve the state estimation problem for discrete-time systems subjected to Markovian switching and affected by additive noise (not necessarily Gaussian). The key point for the derivation of the optimal polynomial filter for the considered class of systems is the possibility to represent them as bilinear systems (linear drift with multiplicative noise). This goal is achieved by means of a suitable bilinear state space representation for the Markov jump process and through 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. It follows that for $\nu > 1$ better performances are obtained with respect to linear filters. Simulation results are reported as a validation of the theory.

Key words: Polynomial filtering, Kalman Filtering, Markov Processes, Stochastic Systems.

1. Introduction

Switching systems, also denoted hybrid systems or variable structure systems, are receiving a growing attention in recent years because of their importance from an applicative point of view, in that switching phenomena are normally present in many engineering problems (for a survey on hybrid systems control and applications see [3,8,14]). Many authors investigated the problem of state estimation for switching systems. Most papers in literature deal with variable structure systems with the switching coefficients modeled by a finite-state Markov Chain (see e.g. [1,4,7,9,10,12,15] for the discrete-time case and [13,16,17] for the continuous-time case).

In the framework of discrete-time systems, the problem of filtering the state of stochastic switching systems was first formulated in [1]. The authors pointed out the complexity of the exact solution of the problem and proposed an approximate solution. The same problem is considered in [10], where a partial observation of the switching process is assumed, and an almost-recursive implementation of the exact solution is derived. The drawback of this approach is that the complexity of the algorithm grows geometrically with time. In [7] a linear filter is implemented based on a clever use of the characteristic function associated to the Markovian jump. In [9] different approximate state estimators have been analyzed, without assuming observations on the switching process. All estimators proposed in [9] are iterative algorithms over a finite observation time, and do not allow a recursive implementation.

This paper proposes a minimum variance polynomial algorithm for the state estimation of a discrete-time variable structure system driven by Markovian switching coefficients. The polynomial methodology exploits the geometric interpretation of the minimum error variance estimator in a chosen class of measurements transformations. If the considered class of estimators forms a closed linear space, the optimal estimator can be computed as a projection. The polynomial approach considers the closed linear space of all polynomial output transformations of a chosen degree ν . Recently, this approach has led to important results in the field of suboptimal filtering of non Gaussian linear [5] and bilinear [6] systems. In [11] the authors presented the equations of the optimal linear filter for stochastic switching systems. The filter presented in [11] is equivalent, from a statistical point of view, to the filter proposed in [7] (differences lie on the methodologies adopted). 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.

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 described by the following set of equations:

$$\begin{aligned} x(k+1) &= A_{\mu(k)}x(k) + B_{\mu(k)}u(k) + F_{\mu(k)}N(k), & x(0) &= x_0, & k &\geq 0, \\ y(k) &= C_{\mu(k)}x(k) + D_{\mu(k)}u(k) + G_{\mu(k)}N(k), \end{aligned} \quad (2.1)$$

where the state $x(k)$ 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 . The matrices in (2.1) are forced to assume values on a finite set, depending on the jump parameter $\mu(k)$, which is a scalar Markov process, with finite range $\mathcal{R}(\mu) = \{1, \dots, m\} \subset \mathbb{N}$ and probability transition matrix $\Pi(k)$ defined by:

$$(\Pi(k))_{ij} = P(\mu(k+1) = i | \mu(k) = j), \quad i, j = 1, \dots, m, \quad (2.2)$$

with initial probability density

$$P(\mu(0) = i) = p_i, \quad i = 1, \dots, m, \quad \sum_{i=1}^m p_i = 1. \quad (2.3)$$

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

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

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

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

with \otimes the standard Kronecker product (for a quick survey on Kronecker products and their principal properties, see [6] and references therein). Note that, according to the noise statistics: $\xi_0(k) = 1$ and $\xi_1(k) = 0$. Moreover $N(k)$ is also independent of the Markov chain $\mu(k)$.

The initial state x_0 is a random variable, independent of both the noise sequence and the Markov chain, with finite and available moments up to the $2\nu^{\text{th}}$ -degree, named:

$$\mathbb{E}[x_0^{[j]}] = \zeta_j, \quad 1 \leq j \leq 2\nu. \quad (2.6)$$

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 the following state space realization for the multi-values Markov process as established by the following lemma:

Lemma 2.1. *Let $\{\theta(k) \in \mathbb{R}^m, k \in \mathbb{N}\}$ be a stochastic sequence assuming values in $\mathcal{E}_m = \{e_j, j = 1, \dots, m\}$, the natural basis in \mathbb{R}^m , according to the following stochastic recursive equation:*

$$\theta(k+1) = V(k)\theta(k), \quad \theta(0) = \theta_0, \quad k \in \mathbb{N}, \quad (2.7)$$

where θ_0 is a random variable assuming values in \mathcal{E}_m , with probabilities given by :

$$P(\theta_0 = e_i) = P(\mu(0) = i) = p_i, \quad i = 1, \dots, m, \quad (2.8)$$

(see (2.3)) and $\{V(k), k \in \mathbb{N}\}$ is a sequence of random matrices, whose j^{th} column sequence $\{V_j(k), k \in \mathbb{N}\}$ is a sequence of independent random vectors assuming values in \mathcal{E}_m , with probabilities given by :

$$P(V_j(k) = e_i) = (\Pi(k))_{ij}, \quad i, j = 1, \dots, m, \quad (2.9)$$

where the matrix $\Pi(k)$ is the same as defined in (2.2). Moreover, the column set of $V(k)$, together with θ_0 , is a set of independent random vectors. Then the system matrices $A_{\mu(k)}$, $B_{\mu(k)}$, $C_{\mu(k)}$, $D_{\mu(k)}$, $F_{\mu(k)}$, $G_{\mu(k)}$ can be represented as:

$$A_{\mu(k)} = \tilde{A} \cdot (\theta(k) \otimes I_n), \quad \tilde{A} = [A_1 \ \cdots \ A_m], \quad (2.10a)$$

$$B_{\mu(k)} = \tilde{B} \cdot (\theta(k) \otimes I_p), \quad \tilde{B} = [B_1 \ \cdots \ B_m], \quad (2.10b)$$

$$C_{\mu(k)} = \tilde{C} \cdot (\theta(k) \otimes I_n), \quad \tilde{C} = [C_1 \ \cdots \ C_m], \quad (2.10c)$$

$$D_{\mu(k)} = \tilde{D} \cdot (\theta(k) \otimes I_p), \quad \tilde{D} = [D_1 \ \cdots \ D_m], \quad (2.10d)$$

$$F_{\mu(k)} = \tilde{F} \cdot (\theta(k) \otimes I_b), \quad \tilde{F} = [F_1 \ \cdots \ F_m], \quad (2.10e)$$

$$G_{\mu(k)} = \tilde{G} \cdot (\theta(k) \otimes I_b), \quad \tilde{G} = [G_1 \ \cdots \ G_m]. \quad (2.10f)$$

Proof. The proof is developed only with reference to matrix $A_{\mu(k)}$ assuming, without loss of generality, that $A_i \neq A_j, \forall i \neq j$. Note, first, the identity induced by the Kronecker product:

$$\tilde{A} \cdot (\theta(k) \otimes I_n) \Big|_{\theta(k)=e_i} = [A_1 \ \cdots \ A_i \ \cdots \ A_m] \begin{bmatrix} O_{(i-1)n \times n} \\ I_n \\ O_{(m-i)n \times n} \end{bmatrix} = A_i, \quad (2.11)$$

so that the following identity is true:

$$P(\tilde{A} \cdot (\theta(k) \otimes I_n) = A_i) = P(\theta(k) = e_i). \quad (2.12)$$

The proof consists in verifying by induction that matrices $A_{\mu(k)}$ and $\tilde{A} \cdot (\theta(k) \otimes I_n)$ assume the same values with the same probabilities, for each $k \in \mathbb{N}$. Let $k = 0$. Then:

$$P(A_{\mu(0)} = A_i) = P(\mu(0) = i) = P(\theta_0 = e_i) = P(\tilde{A} \cdot (\theta_0 \otimes I_n) = A_i), \quad i = 1, \dots, m. \quad (2.13)$$

Assume, now, that $A_{\mu(k)}$ and $\tilde{A} \cdot (\theta(k) \otimes I_n)$ have the same probability distribution for a given $k \in \mathbb{N}$, that is:

$$P(A_{\mu(k)} = A_i) = P(\tilde{A} \cdot (\theta(k) \otimes I_n) = A_i), \quad i = 1, \dots, m. \quad (2.14)$$

Then:

$$\begin{aligned} P(A_{\mu(k+1)} = A_i) &= P(\mu(k+1) = i) = \sum_{j=1}^m P(\mu(k+1) = i | \mu(k) = j) P(\mu(k) = j) \\ &= \sum_{j=1}^m (\Pi(k))_{ij} P(A_{\mu(k)} = A_j). \end{aligned} \quad (2.15)$$

6.

Moreover:

$$P\left(\tilde{A}(\theta(k+1) \otimes I_n) = A_i\right) = P\left(\theta(k+1) = e_i\right) = \sum_{j=1}^m P\left(\theta(k+1) = e_i | \theta(k) = e_j\right) P\left(\theta(k) = e_j\right). \quad (2.16)$$

$\theta(k)$ assumes values in \mathcal{E}_m so that, from equation (2.7) it comes that the j -th column $V_j(k)$ of matrix $V(k)$ is the random variable $\theta(k+1)$ conditioned by $\theta(k)$ assuming the value e_j . This means that the probability in (2.16) becomes:

$$\begin{aligned} P\left(\tilde{A}(\theta(k+1) \otimes I_n) = A_i\right) &= \sum_{j=1}^m P\left(V_j(k) = e_i\right) P\left(\tilde{A}(\theta(k) \otimes I_n) = A_j\right) \\ &= \sum_{j=1}^m (\Pi(k))_{ij} P\left(\tilde{A}(\theta(k) \otimes I_n) = A_j\right), \end{aligned} \quad (2.17)$$

so that, from assumption (2.14), the Lemma is proved by induction. ■

Since $\theta(k) \in \mathcal{E}_m$, the following identity can be easily verified:

$$\theta^{[2]}(k) = E_2 \theta(k), \quad E_2 = [e_1^{[2]} \dots e_m^{[2]}]. \quad (2.18)$$

Moreover, note that, according to its definition, the random matrix $V(k)$ is independent of $\theta(k)$ and its mean value is given by the probability transition matrix: $\mathbb{E}[V(k)] = \Pi(k)$. The following lemma shows some properties concerning the statistics of $V(k)$.

Lemma 2.2. *Consider the zero-mean random matrix $\mathcal{V}(k) = V(k) - \Pi(k)$. Then:*

$$\mathbb{E}[\mathcal{V}_i(k) \otimes \mathcal{V}_j(k)] = \begin{cases} E_2 \cdot \Pi_i(k) - \Pi_i^{[2]}(k), & i = j, \\ O, & i \neq j, \end{cases} \quad (2.19)$$

with $\mathcal{V}_i(k)$ and $\Pi_i(k)$ the i^{th} columns of the matrices $\mathcal{V}(k)$ and $\Pi(k)$ respectively. Moreover,

$$\mathbb{E}[\mathcal{V}^{[2]}(k)] = \bar{\mathcal{V}}_2(k) E_2^T, \quad \bar{\mathcal{V}}_2(k) = [\mathbb{E}[\mathcal{V}_1^{[2]}(k)] \dots \mathbb{E}[\mathcal{V}_m^{[2]}(k)]]. \quad (2.20)$$

Proof. Recalling that the columns of $V(k)$ form a set of independent random vectors, as defined in lemma 2.1, equation (2.19) clearly comes for $i \neq j$. Let $i = j$. The range of each column $V_i(k)$, the same of $\theta(k)$, is the natural basis in \mathbb{R}^m , so that, according to (2.18), $V_i^{[2]}(k) = E_2 V_i(k)$. Then

$$\mathbb{E}[\mathcal{V}_i^{[2]}(k)] = \mathbb{E}[V_i^{[2]}(k)] - \Pi_i^{[2]}(k) = \mathbb{E}[E_2 \cdot V_i(k)] - \Pi_i^{[2]}(k) = E_2 \cdot \Pi_i(k) - \Pi_i^{[2]}(k). \quad (2.21)$$

In order to show (2.20), note that $\mathcal{V}(k) = \sum_{i=1}^m \mathcal{V}_i(k) e_i^T$. Then

$$\mathcal{V}^{[2]}(k) = \sum_{i=1}^m \sum_{j=1}^m (\mathcal{V}_i(k) e_i^T) \otimes (\mathcal{V}_j(k) e_j^T) = \sum_{i=1}^m \sum_{j=1}^m (\mathcal{V}_i(k) \otimes \mathcal{V}_j(k)) (e_i^T \otimes e_j^T). \quad (2.22)$$

According to (2.19) the mean value of $\mathcal{V}^{[2]}(k)$ is:

$$\mathbb{E}[\mathcal{V}^{[2]}(k)] = \sum_{i=1}^m \sum_{j=1}^m \mathbb{E}[\mathcal{V}_i(k) \otimes \mathcal{V}_j(k)] (e_i^T \otimes e_j^T) = \sum_{i=1}^m \mathbb{E}[\mathcal{V}_i^{[2]}(k)] e_i^{[2]T} = \bar{\mathcal{V}}_2(k) E_2^T. \quad (2.23)$$

■

By using the results of the previous lemmas and definitions, the stochastic system (2.1) admits the state space representation, summarized in the proposition reported below:

Proposition 2.3. *System (2.1) admits the following representation:*

$$\begin{aligned} x(k+1) &= \tilde{A} \cdot (\theta(k) \otimes x(k)) + \bar{B}(k)\theta(k) + \tilde{F} \cdot (\theta(k) \otimes N(k)), \\ \theta(k+1) &= \Pi(k)\theta(k) + \mathcal{V}(k)\theta(k), \quad k \geq 0, \\ y(k) &= \tilde{C} \cdot (\theta(k) \otimes x(k)) + \bar{D}(k)\theta(k) + \tilde{G} \cdot (\theta(k) \otimes N(k)), \end{aligned} \quad \begin{aligned} x(0) &= x_0, \\ \theta(0) &= \theta_0, \end{aligned} \quad (2.24)$$

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

$$\bar{B}(k) = \tilde{B}(I_m \otimes u(k)), \quad \bar{D}(k) = \tilde{D}(I_m \otimes u(k)). \quad (2.25)$$

Proof. According to lemma 2.1, system (2.1) can be put in the form:

$$\begin{aligned} x(k+1) &= \tilde{A} \cdot (\theta(k) \otimes I_n)x(k) + \tilde{B} \cdot (\theta(k) \otimes I_p)u(k) + \tilde{F} \cdot (\theta(k) \otimes I_b)N(k), \\ y(k) &= \tilde{C} \cdot (\theta(k) \otimes I_n)x(k) + \tilde{D} \cdot (\theta(k) \otimes I_p)u(k) + \tilde{G} \cdot (\theta(k) \otimes I_b)N(k), \end{aligned} \quad (2.26)$$

Now, by using the Kronecker product properties [6]:

$$\begin{aligned} (\theta(k) \otimes I_n)x(k) &= (\theta(k) \otimes I_n)(1 \otimes x(k)) = \theta(k) \otimes x(k), \\ (\theta(k) \otimes I_p)u(k) &= (\theta(k) \otimes I_p)(1 \otimes u(k)) = \theta(k) \otimes u(k) = (I_m \cdot \theta(k)) \otimes (u(k) \cdot 1) \\ &= (I_m \otimes u(k))\theta(k), \end{aligned} \quad (2.27)$$

so that (2.24) and (2.25) easily come. ■

3. The polynomial filter

It is well known that the optimal solution to the minimum variance filtering problem is given by the expectation value of the state conditioned by 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:

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

In the Gaussian case the optimal filter is a linear transformation of the measurements (the Kalman filter in the case of linear systems with Gaussian noise). Unfortunately, in the non Gaussian case, there is not a simple characterization of the conditional expectation, so that it is worthwhile to consider suboptimal estimates which have a simpler mathematical structure. The simplest suboptimal estimate is the optimal affine one. It consists in projecting the state onto the subspace $L(Y_k)$ of all the linear transformations of the output. For linear systems the optimal affine estimate is achieved by the Kalman filter. Suboptimal estimates comprised between the optimal linear and the conditional expectation can be considered by projecting onto subspaces greater than $L(Y_k)$, like subspaces of polynomial transformations of the measurements

[5, 6]. More in details, the subspace here considered is the following Hilbert space of a fixed degree ν polynomial transformations of the measurements:

$$\hat{x}_\nu(k) = \Pi[x(k)|L(Y_k^\nu)], \quad L(Y_k^\nu) = \text{span}\{Y^\nu(0), \dots, Y^\nu(k)\}, \quad (3.2)$$

with

$$Y_k^\nu = \begin{bmatrix} Y^\nu(0) \\ \vdots \\ Y^\nu(k) \end{bmatrix}, \quad Y^\nu(h) = \begin{bmatrix} Y_1(h) \\ \vdots \\ Y_\nu(h) \end{bmatrix}, \quad Y_i(h) = y^{[i]}(h), \quad h = 0, \dots, k, \quad (3.3)$$

and the extra-assumption that $\mathbb{E}\{|y^{[i]}(h)|^2\} < \infty$, for $i = 1, \dots, \nu$.

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

$$\hat{x}_\nu(k) = \mathcal{M}_n \hat{X}^\nu(k) = \mathcal{M}_n \Pi[X^\nu(k)|L(Y_k^\nu)], \quad \mathcal{M}_n = [O_{n \times m} \quad \mathcal{M} \quad O_{n \times m(n^2 + \dots + n^\nu)}], \quad (3.4)$$

with $\mathcal{M} = [I_n \dots I_n] \in \mathbb{R}^{n \times mn}$ and

$$X^\nu(k) = \begin{bmatrix} X_0(k) \\ \vdots \\ X_\nu(k) \end{bmatrix}, \quad X_i(k) = \theta(k) \otimes x^{[i]}(k). \quad (3.5)$$

Proof. According to the measurement equation in (2.24), all the output Kronecker powers depend on the vectors X_i defined in (3.5). Moreover, taking into account the range of the Markov parameter $\theta(k)$, then $x(k) = \mathcal{M}_n X^\nu(k)$ so that the polynomial minimum variance estimate in (3.2) is:

$$\begin{aligned} \hat{x}_\nu(k) &= \Pi[x(k)|L(Y_k^\nu)] = \Pi[\mathcal{M}_n X^\nu(k)|L(Y_k^\nu)] \\ &= \mathcal{M}_n \Pi[X^\nu(k)|L(Y_k^\nu)] = \mathcal{M}_n \hat{X}^\nu(k), \end{aligned} \quad (3.6)$$

■

The remaining of the paper is devoted to the computation of the projection in equation (3.6). The first step is to show that the sequences $\{X^\nu(k)\}$ and $\{Y^\nu(k)\}$ obey difference equations of the type

$$\begin{aligned} X^\nu(k+1) &= \mathbf{A}^\nu(k)X^\nu(k) + \mathcal{F}(k), \\ Y^\nu(k) &= \mathbf{C}^\nu(k)X^\nu(k) + \mathcal{G}(k), \end{aligned} \quad k \geq 0 \quad (3.7)$$

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

$$\begin{aligned} \mathcal{F}(k) &= \tilde{\mathcal{F}}(k, u(k), X^\nu(k), N(k)), \\ \mathcal{G}(k) &= \tilde{\mathcal{G}}(k, u(k), X^\nu(k), N(k)), \end{aligned} \quad (3.8)$$

with $\tilde{\mathcal{F}}$ and $\tilde{\mathcal{G}}$ 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 (3.7) is that, once $\mathbf{A}^\nu(k)$ and $\mathbf{C}^\nu(k)$ are known, together with the covariance matrices of the noise sequences $\mathcal{F}(k)$ and $\mathcal{G}(k)$, the minimum variance filter

for such a kind of bilinear system [6] can be used in order 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 (3.7). Before the statement of the Lemmas some notations must be introduced.

For each vector X_j , defined as in (3.5), there exist suitably defined matrices $\Theta_n^{h,j}$ and $\Xi_{i,j}$ such that:

$$X_j^{[h]} = \Theta_n^{h,j} X_{jh}, \quad X_i \otimes X_j = \Xi_{i,j} X_{i+j}, \quad \forall i, j, h \in \mathbb{N}, \quad (3.9)$$

(see Appendix for details on the matrices). Note that:

$$\Theta_n^{0,j} = [1 \cdots 1] \in \mathbb{R}^{1 \times m}, \quad \Theta_n^{1,j} = I_{mnj}, \quad \Theta_n^{2,0} = \Xi_{0,0} = E_2. \quad (3.10)$$

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 $\text{st}(A)$. The inverse operation is denoted $\text{st}_{r,c}^{-1}(\cdot)$, 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 $\text{st}^{-1}(\text{st}(A)) = A$.

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 $B \in \mathbb{R}^{r_b \times c_b}$

$$B \otimes A = C_{r_a, r_b}^T (A \otimes B) C_{c_a, c_b}, \quad (3.11)$$

(see [6]).

Lemma 3.2. *The iterative equation of the component $X_j(k)$ as defined in (3.5) is:*

$$X_j(k+1) = \sum_{t_1=0}^j \mathbf{A}_{j,t_1}(k) X_{t_1}(k) + \mathcal{F}_j(k), \quad \mathcal{F}_j(k) = \sum_{t_1=0}^j S_{t_1}^j(k) X_{t_1}(k), \quad (3.12)$$

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

$$\mathbf{A}_{j,t_1}(k) = (\Pi(k) \otimes J_{t_1}^j(k)) \Xi_{0,t_1}, \quad (3.13a)$$

$$S_{t_1}^j(k) = \left(\Pi(k) \otimes \mathcal{L}_{t_1}^j(k) + \mathcal{V}(k) \otimes J_{t_1}^j(k) + \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k) \right) \Xi_{0,t_1}, \quad (3.13b)$$

with

$$J_{t_1}^j(k) = \sum_{t_2, t_3}^{t \in \mathcal{R}_j} L_t^j(k) \left(I_{mn t_1} \otimes \xi_{t_3}(k) \right), \quad (3.14a)$$

$$\mathcal{L}_{t_1}^j(k) = \sum_{t_2, t_3}^{t \in \mathcal{R}_j} L_t^j(k) \left(I_{mn t_1} \otimes (N^{[t_3]}(k) - \xi_{t_3}(k)) \right), \quad (3.14b)$$

$$L_t^j(k) = M_t^j \left(\tilde{A}^{[t_1]} \otimes \tilde{B}^{[t_2]}(k) \otimes \tilde{F}^{[t_3]} \right) K_t^j, \quad (3.14c)$$

where K_t^j are given by (A.9b) in Appendix and M_t^j are the matrix coefficients for the polynomial Kronecker power expansion (see [6]). $t = (t_1, t_2, t_3)^T$ is a multi-index in \mathbb{N}^3 and $\mathcal{R}_j = \{t \in$

10.

$\mathcal{N}^3 : t_1 + t_2 + t_3 = j$. $\mathcal{F}(k) = (\mathcal{F}_0(k)^T \cdots \mathcal{F}_\nu(k)^T)^T$ is a zero-mean sequence of uncorrelated random vectors, whose covariance matrices $\Psi_{j,i}^{\mathcal{F}}(k) = \mathbb{E}[\mathcal{F}_j(k)\mathcal{F}_i(k)^T]$ are:

$$\Psi_{j,i}^{\mathcal{F}}(k) = \sum_{j_1=0}^j \sum_{i_1=0}^i \text{st}_{mn^{j_1}, mn^{i_1}}^{-1} \left(\Phi_{r_1, t_1}^{S, i, j}(k) \cdot \Xi_{r_1, t_1} \mathbb{E}[X_{r_1+t_1}(k)] \right), \quad (3.15)$$

where $\Phi_{r_1, t_1}^{S, i, j}(k) = \mathbb{E}[S_{r_1}^i(k) \otimes S_{t_1}^j(k)]$ given by:

$$\begin{aligned} \Phi_{r_1, t_1}^{S, i, j}(k) &= (I_m \otimes C_{mn^j, n^i}^T) \left(\Pi^{[2]}(k) \otimes \Phi_{t_1, r_1}^{\mathcal{L}, j, i}(k) + (\bar{\mathcal{V}}_2(k) E_2^T) \otimes J_{t_1}^j(k) \otimes J_{r_1}^i(k) \right. \\ &\quad \left. + (\bar{\mathcal{V}}_2(k) E_2^T) \otimes \Phi_{t_1, r_1}^{\mathcal{L}, j, i}(k) \right) (I_m \otimes C_{m^2 n^{t_1}, mn^{r_1}}) (\Xi_{0, r_1} \otimes \Xi_{0, t_1}), \end{aligned} \quad (3.16)$$

with

$$\begin{aligned} \Phi_{t_1, r_1}^{\mathcal{L}, j, i}(k) &= \mathbb{E}[\mathcal{L}_{t_1}^j(k) \otimes \mathcal{L}_{r_1}^i(k)] = \sum_{t_2, t_3}^{t \in \mathcal{R}_j} \sum_{r_2, r_3}^{r \in \mathcal{R}_i} (L_t^j(k) \otimes L_r^i(k)) (I_{mn^{t_1}} \otimes C_{mn^{r_1} b^{r_3}, b^{t_3}}^T) \\ &\quad \cdot \left(I_{m^2 n^{t_1+r_1}} \otimes (\xi_{r_3+t_3}(k) - \xi_{r_3}(k) \otimes \xi_{t_3}(k)) \right) (I_{mn^{t_1}} \otimes C_{mn^{r_1}, 1}). \end{aligned} \quad (3.17)$$

Proof. Taking into account lemma A.2:

$$x^{[j]}(k+1) = \sum_{t_1=0}^j J_{t_1}^j(k) X_{t_1}(k) + \mu_j(k), \quad \mu_j(k) = \sum_{t_1=0}^j \mathcal{L}_{t_1}^j(k) X_{t_1}(k), \quad (3.18)$$

by naming the matrices $J_{t_1}^j(k)$, $\mathcal{L}_{t_1}^j(k)$, deterministic and random respectively, as in (3.14). Then:

$$\begin{aligned} X_j(k+1) &= \theta(k+1) \otimes x^{[j]}(k+1) \\ &= (\Pi(k)\theta(k) + \mathcal{V}(k)\theta(k)) \otimes \left(\sum_{t_1=0}^j J_{t_1}^j(k) X_{t_1}(k) + \mu_j(k) \right) \\ &= (\Pi(k)\theta(k)) \otimes \left(\sum_{t_1=0}^j J_{t_1}^j(k) X_{t_1}(k) \right) + (\mathcal{V}(k)\theta(k)) \otimes \left(\sum_{t_1=0}^j J_{t_1}^j(k) X_{t_1}(k) \right) \\ &\quad + \left((\Pi(k) + \mathcal{V}(k))\theta(k) \right) \otimes \left(\sum_{t_1=0}^j \mathcal{L}_{t_1}^j(k) X_{t_1}(k) \right). \end{aligned} \quad (3.19)$$

The first term of the sum gives:

$$\begin{aligned} (\Pi(k)\theta(k)) \otimes \left(\sum_{t_1=0}^j J_{t_1}^j(k) X_{t_1}(k) \right) &= \sum_{t_1=0}^j (\Pi(k) \otimes J_{t_1}^j(k)) (X_0(k) \otimes X_{t_1}(k)) \\ &= \sum_{t_1=0}^j (\Pi(k) \otimes J_{t_1}^j(k)) \Xi_{0, t_1} X_{t_1}(k) = \sum_{t_1=0}^j \mathbf{A}_{j, t_1}(k) X_{t_1}(k). \end{aligned} \quad (3.20)$$

According to (3.20), the sum of the other two terms of (3.19) gives:

$$\sum_{t_1=0}^j (\mathcal{V}(k) \otimes J_{t_1}^j(k)) \Xi_{0,t_1} X_{t_1}(k) + \sum_{t_1=0}^j \left((\Pi(k) + \mathcal{V}(k)) \otimes \mathcal{L}_{t_1}^j(k) \right) \Xi_{0,t_1} X_{t_1}(k), \quad (3.21)$$

so that equations (3.12), (3.13) are obtained.

The state noise $\mathcal{F}_j(k)$ is a zero-mean sequence, as it easily comes from the fact that $X_{t_1}(k)$ is uncorrelated with the zero-mean random matrix $S_{t_1}^j(k)$. It is also a sequence of uncorrelated random vectors in that, let $h \neq k$, for instance $h < k$:

$$\begin{aligned} \mathbb{E}[\mathcal{F}_j(k) \mathcal{F}_i(h)^T] &= \sum_{t_1=0}^j \sum_{r_1=0}^i \mathbb{E}[S_{t_1}^j(k) X_{t_1}(k) X_{r_1}(h)^T S_{r_1}^i(h)^T] \\ &= \sum_{t_1=0}^j \sum_{r_1=0}^i \mathbb{E}[S_{t_1}^j(k)] \mathbb{E}[X_{t_1}(k) X_{r_1}(h)^T S_{r_1}^i(h)^T] = O. \end{aligned} \quad (3.22)$$

Name $\Psi_{j,i}^{\mathcal{F}}(k) = \mathbb{E}[\mathcal{F}_j(k) \mathcal{F}_i(k)^T]$. According to the stack properties [6]:

$$\begin{aligned} \Psi_{j,i}^{\mathcal{F}}(k) &= \mathbb{E} \left[\sum_{t_1=0}^j \sum_{r_1=0}^i \left(S_{t_1}^j(k) X_{t_1}(k) \right) \left(S_{r_1}^i(k) X_{r_1}(k) \right)^T \right] \\ &= \mathbb{E} \left[\sum_{t_1=0}^j \sum_{r_1=0}^i \text{st}_{mn^{j_1}, mn^{i_1}}^{-1} \left(\left(S_{r_1}^i(k) X_{r_1}(k) \right) \otimes \left(S_{t_1}^j(k) X_{t_1}(k) \right) \right) \right] \\ &= \mathbb{E} \left[\sum_{t_1=0}^j \sum_{r_1=0}^i \text{st}_{mn^{j_1}, mn^{i_1}}^{-1} \left(\left(S_{r_1}^i(k) \otimes S_{t_1}^j(k) \right) \cdot \left(X_{r_1}(k) \otimes X_{t_1}(k) \right) \right) \right] \\ &= \sum_{t_1=0}^j \sum_{r_1=0}^i \text{st}_{mn^{j_1}, mn^{i_1}}^{-1} \left(\mathbb{E} \left[S_{r_1}^i(k) \otimes S_{t_1}^j(k) \right] \cdot \Xi_{r_1, t_1} \mathbb{E} \left[X_{r_1+t_1}(k) \right] \right). \end{aligned} \quad (3.23)$$

Developing computations for $\mathbb{E} \left[S_{r_1}^i(k) \otimes S_{t_1}^j(k) \right]$ in the last of (3.23) one has

$$\begin{aligned} \Phi_{r_1, t_1}^{S, i, j}(k) &= \mathbb{E} \left[S_{r_1}^i(k) \otimes S_{t_1}^j(k) \right] \\ &= \mathbb{E} \left[\left(\Pi(k) \otimes \mathcal{L}_{r_1}^i(k) + \mathcal{V}(k) \otimes J_{r_1}^i(k) + \mathcal{V}(k) \otimes \mathcal{L}_{r_1}^i(k) \right) \right. \\ &\quad \left. \otimes \left(\Pi(k) \otimes \mathcal{L}_{t_1}^j(k) + \mathcal{V}(k) \otimes J_{t_1}^j(k) + \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k) \right) \right] \left(\Xi_{0, r_1} \otimes \Xi_{0, t_1} \right). \end{aligned} \quad (3.24)$$

According to the Kronecker algebra [6], the first term of the mean value in (3.24) can be written as:

$$\mathbb{E} \left[\Pi(k) \otimes \mathcal{L}_{r_1}^i(k) \otimes \Pi(k) \otimes \mathcal{L}_{t_1}^j(k) \right] = (I_m \otimes C_{mn^j, n^i}^T) \left(\Pi^{[2]}(k) \otimes \Phi_{t_1, r_1}^{\mathcal{L}, j, i}(k) \right) (I_m \otimes C_{m^2 n^{t_1}, mn^{r_1}}). \quad (3.25)$$

Following the same procedure, taking into account that $\mathcal{L}_{t_1}^j(k)$ is uncorrelated with $\mathcal{V}(k)$, the following results are readily obtained:

$$\mathbb{E} \left[\Pi(k) \otimes \mathcal{L}_{r_1}^i(k) \otimes \mathcal{V}(k) \otimes J_{t_1}^j(k) \right] = \Pi(k) \otimes \mathbb{E} \left[\mathcal{L}_{r_1}^i(k) \right] \otimes \mathbb{E} \left[\mathcal{V}(k) \right] \otimes J_{t_1}^j(k) = O, \quad (3.26)$$

$$\mathbb{E}\left[\Pi(k) \otimes \mathcal{L}_{r_1}^i(k) \otimes \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k)\right] = \Pi(k) \otimes \mathbb{E}\left[\mathcal{L}_{r_1}^i(k) \otimes \mathbb{E}[\mathcal{V}(k)] \otimes \mathcal{L}_{t_1}^j(k)\right] = O, \quad (3.27)$$

$$\mathbb{E}\left[\mathcal{V}(k) \otimes J_{r_1}^i(k) \otimes \Pi(k) \otimes \mathcal{L}_{t_1}^j(k)\right] = \mathbb{E}\left[\mathcal{V}(k)\right] \otimes J_{r_1}^i(k) \otimes \Pi(k) \otimes \mathbb{E}\left[\mathcal{L}_{t_1}^j(k)\right] = O, \quad (3.28)$$

$$\begin{aligned} \mathbb{E}\left[\mathcal{V}(k) \otimes J_{r_1}^i(k) \otimes \mathcal{V}(k) \otimes J_{t_1}^j(k)\right] \\ = (I_m \otimes C_{mn^j, n^i}^T) \left((\bar{V}_2(k) E_2^T) \otimes J_{t_1}^j(k) \otimes J_{r_1}^i(k) \right) (I_m \otimes C_{m^2 n^{t_1}, mn^{r_1}}), \end{aligned} \quad (3.29)$$

$$\mathbb{E}\left[\mathcal{V}(k) \otimes J_{r_1}^i(k) \otimes \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k)\right] = \mathbb{E}\left[\mathcal{V}(k) \otimes J_{r_1}^i(k) \otimes \mathcal{V}(k)\right] \otimes \mathbb{E}\left[\mathcal{L}_{t_1}^j(k)\right] = O, \quad (3.30)$$

$$\mathbb{E}\left[\mathcal{V}(k) \otimes \mathcal{L}_{r_1}^i(k) \otimes \Pi(k) \otimes \mathcal{L}_{t_1}^j(k)\right] = \mathbb{E}\left[\mathcal{V}(k)\right] \otimes \mathbb{E}\left[\mathcal{L}_{r_1}^i(k) \otimes \Pi(k) \otimes \mathcal{L}_{t_1}^j(k)\right] = O, \quad (3.31)$$

$$\mathbb{E}\left[\mathcal{V}(k) \otimes \mathcal{L}_{r_1}^i(k) \otimes \mathcal{V}(k) \otimes J_{t_1}^j(k)\right] = \mathbb{E}\left[\mathcal{V}(k) \otimes \mathbb{E}[\mathcal{L}_{r_1}^i(k)] \otimes \mathcal{V}(k)\right] \otimes J_{t_1}^j(k) = O, \quad (3.32)$$

$$\begin{aligned} \mathbb{E}\left[\mathcal{V}(k) \otimes \mathcal{L}_{r_1}^i(k) \otimes \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k)\right] \\ = (I_m \otimes C_{mn^j, n^i}^T) \left((\bar{V}_2(k) E_2^T) \otimes \Phi_{t_1, r_1}^{\mathcal{L}, j, i}(k) \right) (I_m \otimes C_{m^2 n^{t_1}, mn^{r_1}}), \end{aligned} \quad (3.33)$$

so that equation (3.16) is achieved. The mean value $\Phi_{t_1, r_1}^{\mathcal{L}, j, i}(k)$ in (3.17) comes from equation (A.10) of lemma A.2 in Appendix. ■

Lemma 3.3. *The measurements equations for the output Y^ν defined in (3.3) are:*

$$Y_j(k) = \sum_{t_1=0}^j \mathbf{C}_{j, t_1}(k) X_{t_1}(k) + \mathcal{G}_j(k), \quad \mathcal{G}_j(k) = \sum_{t_1=0}^j \mathcal{T}_{t_1}^j(k) X_{t_1}(k) \quad (3.34)$$

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

$$\mathbf{C}_{j, t_1}(k) = \sum_{\substack{t \in \mathcal{R}_j \\ t_2, t_3}} T_t^j(k) \left(I_{mn^{t_1}} \otimes \xi_{t_3}(k) \right), \quad (3.35a)$$

$$\mathcal{T}_{t_1}^j(k) = \sum_{\substack{t \in \mathcal{R}_j \\ t_2, t_3}} T_t^j(k) \left(I_{mn^{t_1}} \otimes (N^{[t_3]}(k) - \xi_{t_3}(k)) \right), \quad (3.35b)$$

$$T_t^j(k) = M_t^j \left(\tilde{C}^{[t_1]} \otimes \bar{D}^{[t_2]}(k) \otimes \tilde{G}^{[t_3]} \right) K_t^j, \quad (3.35c)$$

and K_t^j as in (A.9b) in Appendix. Moreover $\mathcal{G}(k) = (\mathcal{G}_1(k)^T, \dots, \mathcal{G}_\nu(k)^T)^T$ is a zero-mean sequence of uncorrelated random vectors, whose covariance matrices $\Psi_{j, i}^{\mathcal{G}}(k) = \mathbb{E}[\mathcal{G}_j(k) \mathcal{G}_i(k)^T]$ are:

$$\Psi_{j, i}^{\mathcal{G}}(k) = \sum_{t_1=0}^j \sum_{r_1=0}^i \text{st}_{q^j, q^i}^{-1} \left(\Phi_{r_1, t_1}^{\mathcal{T}, i, j}(k) \Xi_{r_1, t_1} \mathbb{E}[X_{r_1+t_1}(k)] \right), \quad (3.36)$$

with

$$\begin{aligned} \Phi_{r_1, t_1}^{\mathcal{T}, i, j}(k) &= \mathbb{E} \left[\mathcal{T}_{r_1}^i(k) \otimes \mathcal{T}_{t_1}^j(k) \right] = \sum_{r_2, r_3}^{r \in \mathcal{R}_i} \sum_{t_2, t_3}^{t \in \mathcal{R}_j} \left(T_r^i(k) \otimes T_t^j(k) \right) (I_{mn^{r_1}} \otimes C_{mn^{t_1} b^{t_3}, b^{r_3}}^T) \\ &\quad \cdot \left(I_{m^2 n^{r_1+t_1}} \otimes \left(\xi_{t_3+r_3}(k) - \xi_{t_3}(k) \otimes \xi_{r_3}(k) \right) \right) (I_{mn^{r_1}} \otimes C_{mn^{t_1}, 1}). \end{aligned} \quad (3.37)$$

Proof. The proof is a straightforward consequence of Lemma A.2, except equation (3.36), that can be derived through the same steps used to prove (3.15) in Lemma 3.2. ■

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

$$\begin{aligned} \mathbb{E}[\mathcal{F}_j(k) \mathcal{G}_i^T(h)] &= 0, \quad k \neq h, \quad 0 \leq j \leq \nu, \\ \mathbb{E}[\mathcal{F}_j(k) \mathcal{G}_i^T(k)] &= Q_{j,i}(k), \quad 0 < i \leq \nu, \quad \forall k, h \in \mathbb{N}. \end{aligned} \quad (3.38)$$

with

$$Q_{j,i}(k) = \sum_{t_1=0}^j \sum_{r_1=0}^i \text{st}_{mn^j, q^i}^{-1} \left(\mathcal{Q}_{r_1, t_1}^{i,j}(k) \Xi_{r_1, t_1} \mathbb{E}[X_{r_1+t_1}(k)] \right) \quad (3.39)$$

where $\mathcal{Q}_{r_1, t_1}^{i,j}(k) = \mathbb{E} \left[\mathcal{T}_{r_1}^i(k) \otimes S_{t_1}^j(k) \right]$.

Proof. Analogously to the proof in Lemma 3.2, equation (3.38) is easily verified for $h \neq k$. Following the same passages in (3.23) one has

$$\mathbb{E}[\mathcal{F}_j(k) \mathcal{G}_i(k)^T] = \sum_{t_1=0}^j \sum_{r_1=0}^i \text{st}_{mn^j, q^i}^{-1} \left(\mathcal{Q}_{r_1, t_1}^{i,j}(k) \cdot \Xi_{r_1, t_1} \mathbb{E}[X_{r_1+t_1}(k)] \right), \quad (3.40)$$

with:

$$\begin{aligned} \mathcal{Q}_{r_1, t_1}^{i,j}(k) &= \mathbb{E} \left[\mathcal{T}_{r_1}^i(k) \otimes S_{t_1}^j(k) \right] \\ &= \sum_{r_2, r_3}^{r \in \mathcal{R}_i} \mathbb{E} \left[\left(T_r^i(k) \left(I_{mn^{r_1}} \otimes (N^{[r_3]}(k) - \xi_{r_3}(k)) \right) \right) \right. \\ &\quad \left. \otimes \left(\left(\Pi(k) \otimes \mathcal{L}_{t_1}^j(k) + \mathcal{V}(k) \otimes J_{t_1}^j(k) + \mathcal{V}(k) \otimes \mathcal{L}_{t_1}^j(k) \right) \Xi_{0, t_1} \right) \right]. \end{aligned} \quad (3.41)$$

Now both $N(k)$ and $\mathcal{L}_{t_1}^j(k)$ are uncorrelated with respect to $\mathcal{V}(k)$, so that the mean value in equation (3.41) reduces to:

$$\begin{aligned} &\mathbb{E} \left[\left(T_r^i(k) \left(I_{mn^{r_1}} \otimes (N^{[r_3]}(k) - \xi_{r_3}(k)) \right) \right) \otimes \left(\left(\Pi(k) \otimes \mathcal{L}_{t_1}^j(k) \right) \Xi_{0, t_1} \right) \right] \\ &= (T_r^i(k) \otimes I_{mn^j}) \mathbb{E} \left[I_{mn^{r_1}} \otimes (N^{[r_3]}(k) - \xi_{r_3}(k)) \otimes \Pi(k) \otimes \mathcal{L}_{t_1}^j(k) \right] (I_{mn^{r_1}} \otimes \Xi_{0, t_1}) \\ &= (T_r^i(k) \otimes I_{mn^j}) (I_{mn^{r_1}} \otimes C_{mn^j, b^{r_3}}^T) \left(I_{mn^{r_1}} \otimes \Pi(k) \right. \\ &\quad \left. \otimes \mathbb{E} \left[\mathcal{L}_{t_1}^j(k) \otimes (N^{[r_3]}(k) - \xi_{r_3}(k)) \right] \right) (I_{mn^{r_1}} \otimes C_{m^2 n^{t_1}, 1}) (I_{mn^{r_1}} \otimes \Xi_{0, t_1}). \end{aligned} \quad (3.42)$$

14.

By using (3.14b), the mean value in (3.42) becomes:

$$\begin{aligned}
& \sum_{t_2, t_3}^{t \in \mathcal{R}_j} \mathbb{E} \left[\left(L_t^j(k) \left(I_{mn^{t_1}} \otimes (N^{[t_3]}(k) - \xi_{t_3}(k)) \right) \right) \otimes (N^{[r_3]}(k) - \xi_{r_3}(k)) \right] \\
&= \sum_{t_2, t_3}^{t \in \mathcal{R}_j} (L_t^j(k) \otimes I_{br_3}) \left(I_{mn^{t_1}} \otimes \mathbb{E} \left[(N^{[t_3]}(k) - \xi_{t_3}(k)) \otimes (N^{[r_3]}(k) - \xi_{r_3}(k)) \right] \right) \\
&= \sum_{t_2, t_3}^{t \in \mathcal{R}_j} (L_t^j(k) \otimes I_{br_3}) \left(I_{mn^{t_1}} \otimes (\xi_{t_3+r_3}(k) - \xi_{t_3}(k) \otimes \xi_{r_3}(k)) \right). \tag{3.43}
\end{aligned}$$

■

Proposition 3.5. *According to Theorem 3.1, the ν -th degree polynomial filtering algorithm is the following:*

$$\hat{x}_\nu(k) = \mathcal{M}_n \hat{X}^\nu(k), \tag{3.44a}$$

$$\hat{X}^\nu(k) = \hat{X}^\nu(k|k-1) + \mathcal{K}(k) \left(Y^\nu(k) - \mathbf{C}^\nu(k) \hat{X}^\nu(k|k-1) \right), \tag{3.44b}$$

$$\begin{aligned}
\hat{X}^\nu(k+1|k) &= \left(\mathbf{A}^\nu(k) - (\mathbf{A}^\nu(k) \mathcal{K}(k) + \mathcal{Z}(k)) \mathbf{C}^\nu(k) \right) \hat{X}^\nu(k|k-1) \\
&\quad + (\mathbf{A}^\nu(k) \mathcal{K}(k) + \mathcal{Z}(k)) Y^\nu(k). \tag{3.44c}
\end{aligned}$$

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

$$\mathcal{Z}(k) = Q(k) \left(\mathbf{C}^\nu(k) \mathcal{P}_P(k) \mathbf{C}^{\nu T}(k) + \Psi^{\mathcal{G}}(k) \right)^\dagger \tag{3.45a}$$

$$\begin{aligned}
\mathcal{P}_P(k+1) &= \mathbf{A}^\nu(k) \mathcal{P}(k) \mathbf{A}^{\nu T}(k) + \Psi^{\mathcal{F}}(k) - \mathcal{Z}(k) Q^T(k) \\
&\quad - \mathbf{A}^\nu(k) \mathcal{K}(k) Q^T(k) - Q(k) \mathcal{K}^T(k) \mathbf{A}^{\nu T}(k) \tag{3.45b}
\end{aligned}$$

$$\mathcal{P}(k) = \mathcal{P}_P(k) - \mathcal{K}(k) \mathbf{C}^\nu(k) \mathcal{P}_P(k) \tag{3.45c}$$

$$\mathcal{K}(k) = \mathcal{P}_P(k) \mathbf{C}^{\nu T}(k) \left(\mathbf{C}^\nu(k) \mathcal{P}_P(k) \mathbf{C}^{\nu T}(k) + \Psi^{\mathcal{G}}(k) \right)^\dagger \tag{3.45d}$$

where in (3.45a, d) the Moore-Penrose pseudoinverse has been used. Matrices $\Psi^{\mathcal{F}}(k)$, $\Psi^{\mathcal{G}}(k)$ are, respectively the extended state and measurements noise covariance matrices, given by equations (3.15), defined in Lemma 3.2, and (3.36) defined in Lemma 3.3; $Q(k)$ is the covariance matrix between $\mathcal{F}(k)$ and $\mathcal{G}(k)$ sequences at the same instants, given by equation (3.39) of lemma 3.4. Note that, according to their definition (see (3.15), (3.36) and (3.39)), in order to compute the noise covariance matrices $\Psi^{\mathcal{F}}(k)$, $\Psi^{\mathcal{G}}(k)$ and $Q(k)$ of (3.45), the following 2ν order deterministic system has to be computed:

$$\mathbb{E}[X^\nu(k+1)] = \mathbf{A}^{2\nu}(k) \mathbb{E}[X^\nu(k)], \tag{3.46}$$

which gives the evolution of the mean value of the extended state.

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 (3.7), that has a multiplicative noise structure (see equations (3.34), (3.12), 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 already demonstrated in [6]. ■

Remark 3.6. 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, denoted $\mathcal{P}(k)$ in the algorithm described in Proposition 3.5, as follows

$$\mathbb{E}\left[(x(k) - \hat{x}_\nu(k))(x(k) - \hat{x}_\nu(k))^T\right] = \mathcal{M}_n \mathcal{P}(k) \mathcal{M}_n^T \quad (3.47)$$

4. Numerical simulations

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

- $x(k) \in \mathbb{R}^2$, $u(k), y(k) \in \mathbb{R}$, $\mathcal{R}(\mu) = \{1, 2, 3\}$;
- $A_1 = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix}$, $A_2 = \begin{bmatrix} 0.50 & 0.25 \\ -1.75 & 0.50 \end{bmatrix}$, $A_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$;
- $B_1 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$, $B_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $B_3 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$;
- $C_1 = [1 \ 1]$, $C_2 = [1 \ 0]$, $C_3 = [2 \ 1]$, $D_1 = 1$, $D_2 = 0.5$, $D_3 = 0$;
- $F_1 = \begin{bmatrix} 0.1 & 0 & 0 \\ 0.1 & 0 & 0 \end{bmatrix}$, $F_2 = \begin{bmatrix} 0 & 0.1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $F_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0.1 & 0 \end{bmatrix}$;
- $G_1 = [0 \ 0 \ 0.1]$, $G_2 = [0 \ 0 \ 0.2]$, $G_3 = [0 \ 0 \ 0.04]$;
- the deterministic input throughout the simulation is: $u(k) \equiv 1$, $k \geq 0$;
- the noise $N(k) = (N_1(k), N_2(k), N_3(k))^T$ has independent components, following discrete asymmetric distributions:

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

(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

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

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

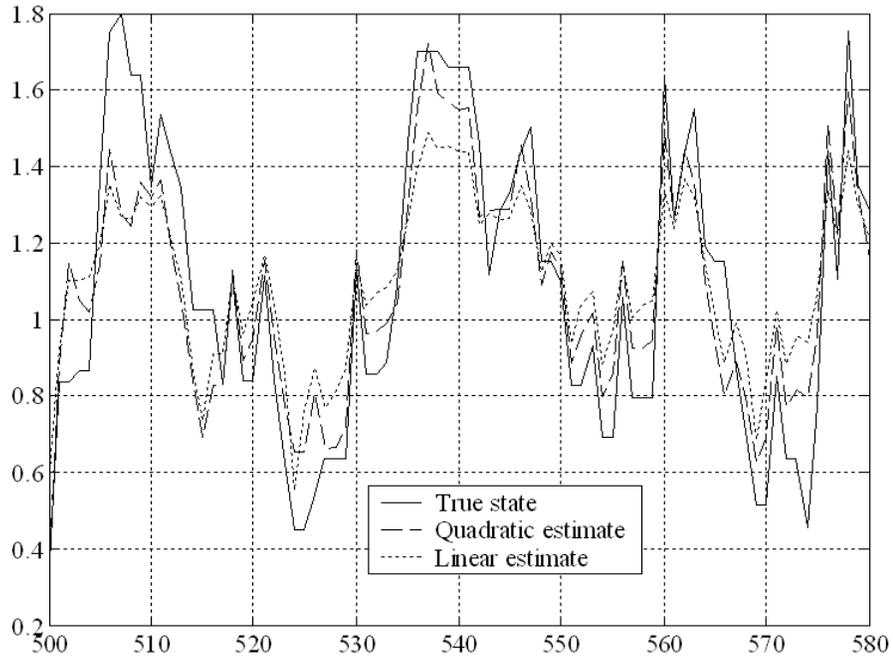


Fig. 4.1 – True and estimated states: the first component.

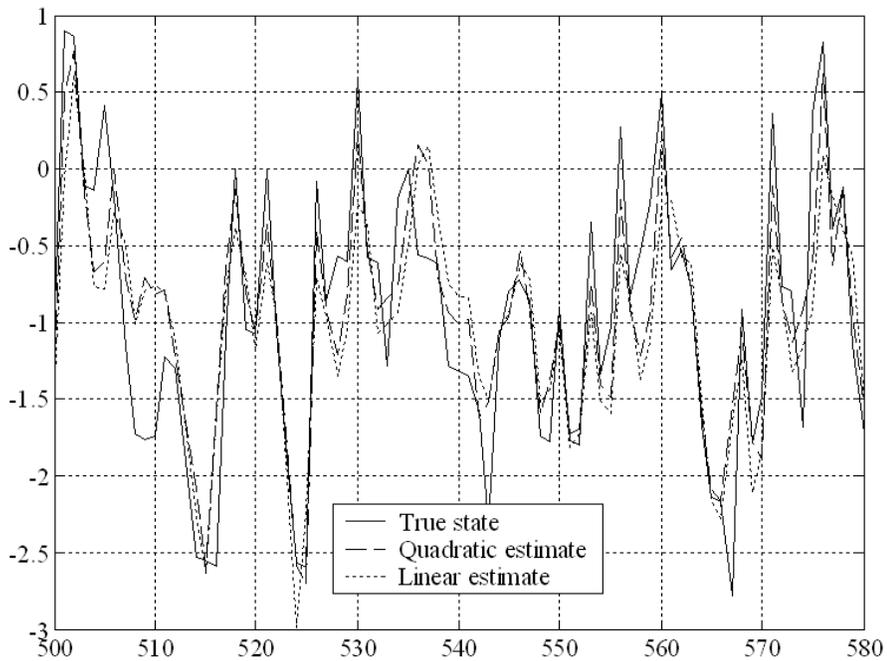


Fig. 4.2 – True and estimated states: the second component.

Appendix: the Kronecker algebra

The appendix contains two Lemmas showing the computation of the matrices used in equations (3.9) and in Lemmas 3.2, 3.3.

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

$$X_j^{[h]} = \Theta_n^{h,j} X_{jh}, \quad X_i \otimes X_j = \Xi_{i,j} X_{i+j}, \quad (\text{A.1})$$

with

$$\begin{aligned} \Theta_n^{h+1,j} &= (\Theta_n^{h,j} \otimes I_{mn^j}) (I_m \otimes C_{mn^j, n^{jh}}^T) (E_2 \otimes I_{n^{j(h+1)}}), \quad h > 0, \\ \Theta_n^{0,j} &= [1 \cdots 1]. \end{aligned} \quad (\text{A.2})$$

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

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

$$\begin{aligned} X_j^{[k+1]} &= X_j^{[k]} \otimes X_j = (\Theta_n^{k,j} X_{jk}) \otimes X_j = (\Theta_n^{k,j} \otimes I_{mn^j}) (\theta \otimes x^{[jk]} \otimes \theta \otimes x^{[j]}) \\ &= (\Theta_n^{k,j} \otimes I_{mn^j}) (I_m \otimes C_{mn^j, n^{jk}}^T) (\theta^{[2]} \otimes x^{[j(k+1)]}) \\ &= (\Theta_n^{k,j} \otimes I_{mn^j}) (I_m \otimes C_{mn^j, n^{jk}}^T) (E_2 \otimes I_{n^{j(k+1)}}) X_{j(k+1)} = \Theta_n^{k+1,j} X_{j(k+1)}. \end{aligned} \quad (\text{A.3})$$

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

$$X_i \otimes X_j = \theta \otimes x^{[i]} \otimes \theta \otimes x^{[j]} = (I_m \otimes C_{mn^j, n^i}^T) (\theta^{[2]} \otimes x^{[i+j]}) = (I_m \otimes C_{mn^j, n^i}^T) (E_2 \otimes I_{n^{i+j}}) X_{i+j}. \quad (\text{A.4})$$

■

Lemma A.2. *Let θ be a random vector taking values in the natural basis of \mathbb{R}^m , N be a random vector taking values in \mathbb{R}^b , with finite and available moments, named:*

$$\mathbb{E}[N^{[j]}] = \xi_j, \quad j \in \mathbb{N}, \quad (\text{A.5})$$

and $z \in \mathbb{R}^n$, $v \in \mathbb{R}^p$ random vectors such that:

$$v = \Gamma_1(\theta \otimes z) + \Gamma_0\theta + \tilde{\Gamma}(\theta \otimes N) \quad (\text{A.6})$$

with $\Gamma_1, \Gamma_0, \tilde{\Gamma}$ matrices of suitable dimensions. Moreover, suppose that $\{\theta, N, z\}$ is a triple of independent random vectors. Then, for each $j \in \mathbb{N}$:

$$v^{[j]} = \sum_{t_1=0}^j \mathcal{H}_{t_1}^j Z_{t_1} + w_j, \quad w_j = \sum_{t_1=0}^j \mathcal{W}_{t_1}^j Z_{t_1}, \quad (\text{A.7})$$

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

$$\mathcal{H}_{t_1}^j = \sum_{t_2, t_3}^{t \in \mathcal{R}_j} W_t^j (I_{mn^{t_1}} \otimes \xi_{t_3}), \quad \mathcal{W}_{t_1}^j = \sum_{t_2, t_3}^{t \in \mathcal{R}_j} W_t^j \left(I_{mn^{t_1}} \otimes (N^{[t_3]} - \xi_{t_3}) \right), \quad (\text{A.8})$$

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

$$W_t^j = M_t^j (\Gamma_1^{[t_1]} \otimes \Gamma_0^{[t_2]} \otimes \tilde{\Gamma}^{[t_3]}) K_t^j, \quad (\text{A.9a})$$

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

M_t^j are the matrix coefficients for the polynomial Kronecker power expansion [6]. Moreover, the second order moments of $W_{t_1}^j(k)$ are the following:

$$\begin{aligned} \Phi_{t_1, r_1}^{W, j, i}(k) &= \mathbb{E}[W_{t_1}^j \otimes W_{r_1}^i(k)] = \sum_{t_2, t_3} \sum_{r_2, r_3}^{t \in \mathcal{R}_j, r \in \mathcal{R}_i} (W_t^j \otimes W_r^i) (I_{mn^{t_1}} \otimes C_{mn^{r_1} b^{r_3}, b^{t_3}}^T) \\ &\quad \cdot \left(I_{m^2 n^{t_1 + r_1}} \otimes (\xi_{r_3 + t_3} - \xi_{r_3} \otimes \xi_{t_3}) \right) (I_{mn^{t_1}} \otimes C_{mn^{r_1}, 1}) \end{aligned} \quad (\text{A.10})$$

Matrices $C_{a,b}$ are the commutation matrices for a Kronecker product (see (3.11) and [6]).

Proof. Applying the Newton formula to the Kronecker powers [6], it comes:

$$\begin{aligned} v^{[j]} &= \left(\Gamma_1 Z_1 + \Gamma_0 Z_0 + \tilde{\Gamma}(\theta \otimes N) \right)^{[j]} \\ &= \sum_{t_1, t_2, t_3}^{t \in \mathcal{R}_j} M_t^j \left((\Gamma_1^{[t_1]} Z_1^{[t_1]}) \otimes (\Gamma_0^{[t_2]} Z_0^{[t_2]}) \otimes \left(\tilde{\Gamma}^{[t_3]}(\theta \otimes N)^{[t_3]} \right) \right) \\ &= \sum_{t_1, t_2, t_3}^{t \in \mathcal{R}_j} M_t^j (\Gamma_1^{[t_1]} \otimes \Gamma_0^{[t_2]} \otimes \tilde{\Gamma}^{[t_3]}) \left(Z_1^{[t_1]} \otimes Z_0^{[t_2]} \otimes (\theta \otimes N)^{[t_3]} \right). \end{aligned} \quad (\text{A.11})$$

By using equations (3.9) and the Kronecker properties [6], the last factor in the sum (A.11) becomes:

$$\begin{aligned} Z_1^{[t_1]} \otimes Z_0^{[t_2]} \otimes (\theta \otimes N)^{[t_3]} &= (\Theta_n^{t_1,1} Z_{t_1}) \otimes (\Theta_n^{t_2,0} Z_0) \otimes \left(\Theta_b^{t_3,1} (\theta \otimes N^{[t_3]}) \right) \\ &= (\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}) (Z_{t_1} \otimes \theta^{[2]} \otimes N^{[t_3]}) \\ &= (\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}) (I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}) (Z_{t_1} \otimes Z_0 \otimes N^{[t_3]}) \\ &= (\Theta_n^{t_1,1} \otimes \Theta_n^{t_2,0} \otimes \Theta_b^{t_3,1}) (I_{mn^{t_1}} \otimes E_2 \otimes I_{b^{t_3}}) (\Xi_{t_1,0} \otimes I_{b^{t_3}}) (Z_{t_1} \otimes N^{[t_3]}) \\ &= K_t^j \left(Z_{t_1} \otimes (\xi_{t_3} + N^{[t_3]} - \xi_{t_3}) \right) = K_t^j (Z_{t_1} \otimes \xi_{t_3}) + K_t^j \left(Z_{t_1} \otimes (N^{[t_3]} - \xi_{t_3}) \right) \\ &= K_t^j (I_{mn^{t_1}} \otimes \xi_{t_3}) Z_{t_1} + K_t^j \left(I_{mn^{t_1}} \otimes (N^{[t_3]} - \xi_{t_3}) \right) Z_{t_1}, \end{aligned} \quad (\text{A.12})$$

so that, substituting (A.12) in (A.11), by using (A.8) and (A.9), equations (A.7) come. According to the independence of N and Z_{t_1} , note that w_j is a zero-mean random vector. Equation (A.10) is below obtained, by using the commutation formula for the Kronecker products [6]:

$$\begin{aligned} \Phi_{t_1, r_1}^{W, j, i} &= \sum_{t_2, t_3} \sum_{r_2, r_3}^{t \in \mathcal{R}_j, r \in \mathcal{R}_i} (W_t^j \otimes W_r^i) \mathbb{E} \left[I_{mn^{t_1}} \otimes (N^{[t_3]} - \xi_{t_3}) \otimes I_{mn^{r_1}} \otimes (N^{[r_3]} - \xi_{r_3}) \right] \\ &= \sum_{t_2, t_3} \sum_{r_2, r_3}^{t \in \mathcal{R}_j, r \in \mathcal{R}_i} (W_t^j \otimes W_r^i) (I_{mn^{t_1}} \otimes C_{mn^{r_1} b^{r_3}, b^{t_3}}^T) \\ &\quad \cdot \left(I_{m^2 n^{t_1 + r_1}} \otimes \mathbb{E} \left[(N^{[r_3]} - \xi_{r_3}) \otimes (N^{[t_3]} - \xi_{t_3}) \right] \right) (I_{mn^{t_1}} \otimes C_{mn^{r_1}, 1}). \end{aligned} \quad (\text{A.13})$$

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

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] A. Balluchi, L. Benvenuti, M. D. Di Benedetto, C. Pinello, A. L. Sangiovanni-Vincentelli, "Automotive Engine Control and Hybrid Systems: Challenge and Opportunities," *Proc. of IEEE*, Vol. 88, No. 7, pp. 888–912, 2000.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] M. D. Di Benedetto, A. Sangiovanni-Vincentelli (Eds): *Hybrid Systems: Computation and Control 2001*, LNCS 2034, Springer-Verlag 2001.
- [9] 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.
- [10] 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.
- [11] A. Germani, C. Manes and P. Palumbo, "State Estimation for a Class of Stochastic Variable Structure Systems," *proc. of 41th Conf. on Decision and Control*, pp. 2514–2519, Las Vegas, Nevada, 2002.
- [12] 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.
- [13] J. L. Hibey, C. D. Charalambous, "Conditional Densities for Continuous-Time Nonlinear Hybrid Systems with Applications to Fault Detection," *IEEE Trans. Automat. Contr.*, Vol. 44, No. 11, pp. 2164–2169, 1999.
- [14] K. H. Johansson, M. Egerstedt, J. Lygeros, S. Sastry, "On the regularization of Zeno hybrid automata. Hybrid control systems," *Systems Control Lett.*, Vol. 38, no. 3, 141–150, 1999.
- [15] 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.
- [16] 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.
- [17] Q. Zhang, "Hybrid Filtering for Linear Systems with Non-Gaussian Disturbances," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 1, pp. 50–61, 2000.