



ISTITUTO DI ANALISI DEI SISTEMI ED INFORMATICA
CONSIGLIO NAZIONALE DELLE RICERCHE

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

**QUADRATIC FILTERING FOR SIMULTANEOUS
STATE AND PARAMETER ESTIMATION OF
UNCERTAIN SYSTEMS**

R. 589 Giugno 2003

Domenico Di Martino – Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, Monteluco di Roio, 67040 L'Aquila, Italy. E-mail: dimartin@iasi.rm.cnr.it, domenico@ing.univaq.it.

Alfredo Germani – Dipartimento di Ingegneria Elettrica, Università degli studi dell'Aquila, Monteluco di Roio, 67040 L'Aquila, Italy. E-mail: germani@{iasi.rm.cnr.it, ing.univaq.it}.

Costanzo Manes – Dipartimento di Ingegneria Elettrica, Università degli studi dell'Aquila, Monteluco di Roio, 67040 L'Aquila, Italy. E-mail: manes@{iasi.rm.cnr.it, ing.univaq.it}.

Pasquale Palumbo – Istituto di Analisi dei Sistemi ed Informatica del CNR, Viale Manzoni 30, 00185 Roma, Italy. E-mail: palumbo@iasi.rm.cnr.it, ppalumbo@ing.univaq.it.

This work is supported by IASI-CNR and by EU Project HYBRIDGE

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

The aim of this paper is to investigate the problem of the joint estimation of both the state and parameters for a class of discrete-time linear systems driven by additive noise, not necessarily Gaussian. A recursive quadratic filter with respect to the observations is here proposed and implemented, by opportunely extending the state space also with the inclusion of the parameters vector. The algorithm is achieved with the systematic use of the Kronecker algebra, which constitutes a powerful tool for polynomial manipulations of vectors and matrices. Numerical simulations are also reported, showing the high performances of the proposed methods with respect to the usually adopted Extended Kalman Filter.

Key words: Nonlinear Filtering, Parameters Identification, Extended Kalman Filter

1. Introduction

Consider the following class of stochastic uncertain linear systems:

$$x(k+1) = A(\vartheta)x(k) + v(k), \quad x(0) = x_0, \quad k \geq 0, \quad (1.1)$$

$$y(k) = C(\vartheta)x(k) + w(k) \quad (1.2)$$

where $x(k) \in \mathbb{R}^n$ is the state of the system, $y(k) \in \mathbb{R}^q$ is the measured output, $\vartheta \in \mathbb{R}^m$ is the unknown vector of parameters, $v(k)$ and $w(k)$ are sequences of zero-mean, auto and mutually independent random vectors (white noise sequences). The most popular real time algorithm for simultaneous state and parameters estimation for this kind of systems is perhaps the Extended Kalman Filter (EKF) applied to an extended nonlinear system whose extended state is made of the original state and the parameter vector [1, 2, 3]. The diffusion of the EKF is due to its simplicity and on the fact that in many applications it provides good estimates. The EKF is based on the linear approximation of the extended system around the current estimate, and therefore it performs well in those cases in which the initial state estimate is good and the noises have low variance and approximately gaussian distribution. In such cases the state estimate remain close to the true state and the first-order Taylor expansion around such estimate remains a good approximation of the system dynamics. However, in the presence of high level non gaussian noises the state estimate deteriorates and the first-order approximation is no more a good model for the nonlinear system.

This paper deals with the problem of simultaneous filtering and parameters identification for systems of the type (1.1)–(1.2), in the case in which each component ϑ_i of the unknown parameter vector $\vartheta \in \mathbb{R}^m$ satisfies an interval constraint

$$\vartheta_i \in [\vartheta_{i,\min}, \vartheta_{i,\max}], \quad i = 1, \dots, m, \quad (1.3)$$

and the noise sequences $v(k)$ and $w(k)$ are far to be Gaussian. System matrices A and C are linearly dependent on the components of ϑ , thus modeling the case of both the state and the output matrices that are uncertain with respect to some of their entries. Such an assumption means that system (1.1)–(1.2) also belongs to the class of interval systems ([4, 5, 6]) so that it can be rewritten in the following form:

$$x(k+1) = A_0x(k) + \sum_{i=1}^m \vartheta_i A_i x(k) + v(k), \quad (1.4)$$

$$y(k) = C_0x(k) + \sum_{i=1}^m \vartheta_i C_i x(k) + w(k), \quad (1.5)$$

where $A_0 = A(0)$, $A_i = A(e_i) - A_0$, $C_0 = C(0)$ and $C_i = C(e_i) - C_0$, with e_i the i -th vector of the natural basis in \mathbb{R}^m .

By taking into account the unknown vector $\vartheta \in \mathbb{R}^m$ as a component of the augmented state $\tilde{X}(k) = [x(k)^T \quad \vartheta(k)^T]^T$, the uncertain system described by (1.4)–(1.5), endowed with the equation $\vartheta(k+1) = \vartheta(k)$, is bilinear with respect to the pair (x, ϑ) . Then, a way to solve both the filtering and the identification problem consists of implementing the Extended Kalman Filter for the bilinear system, that means to apply the optimal linear filter to the linear approximation of (1.4)–(1.5).

The novelty of the proposed approach is to use also the quadratic powers of the measurements in order to improve the performances of the standard EKF, by using a suitable extended state,

4.

named $X(k)$, containing all the second order powers of $\tilde{X}(k)$. Moreover, such an extended state $X(k)$ is chosen so that the extended output $Y(k) = [y^T(k) \ y^{[2]T}(k)]^T$ is related to $X(k)$ according to a linear equation, driven by a multiplicative noise. Here and in the following with square brackets are indicated the Kronecker powers ([7] and the references therein).

As far as the statistics of the noises $\{v(k) \in \mathbb{R}^n\}$ and $\{w(k) \in \mathbb{R}^q\}$ are concerned, it will be assumed the knowledge of their first four moments, namely

$$\mathbb{E}[v^{[i]}(k)] = \xi_i^v(k), \quad \mathbb{E}[w^{[i]}(k)] = \xi_i^w(k), \quad i = 1, \dots, 4. \quad (1.6)$$

The initial state x_0 is a random variable, not necessarily Gaussian, independent of the noise sequences, with finite and available moments up to the 4th order:

$$\mathbb{E}[x_0^{[i]}] = \zeta_i^0, \quad i = 1, \dots, 4. \quad (1.7)$$

Finally, for ϑ , it will be assumed to have a uniform distribution defined by the above mentioned interval constraints. As a consequence, its moments are given by:

$$\mathbb{E}[\vartheta_i^k] = \frac{1}{k+1} \sum_{l=0}^k \vartheta_{i,\min}^{k-l} \vartheta_{i,\max}^l, \quad k = 1, 2, \dots \quad (1.8)$$

According to standard Kronecker formalism, system (1.4)-(1.5) can be rewritten as follows:

$$x(k+1) = [A_0 \ \tilde{A}] \cdot \begin{bmatrix} x(k) \\ \vartheta \otimes x(k) \end{bmatrix} + v(k) \quad (1.9)$$

$$y(k) = [C_0 \ \tilde{C}] \cdot \begin{bmatrix} x(k) \\ \vartheta \otimes x(k) \end{bmatrix} + w(k), \quad (1.10)$$

where $\tilde{A} = [A_1 \dots A_m]$, $\tilde{C} = [C_1 \dots C_m]$ and \otimes is the Kronecker product. For the reader convenience there will be recalled some properties of the Kronecker Algebra (see Appendix A for details):

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

$$B \otimes A = C_{r_a, r_b}^T (A \otimes B) C_{c_a, c_b}; \quad (1.11)$$

- given the vectors $x, y \in \mathbb{R}^r$, it comes the quadratic Kronecker power expansion:

$$(x + y)^{[2]} = x^{[2]} + M_1^2(r)(x \otimes y) + y^{[2]} \quad (1.12)$$

where $M_1^2(r) \in \mathbb{R}^{r^2 \times r^2}$ is a suitable defined matrix coefficient.

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 $\mathbb{R}^{p \times q}$ and a zero vector in \mathbb{R}^p are denoted by $O_{p \times q}$ and O_p respectively, otherwise, no subscripts will be adopted.

2. Filtering of the extended system

In order to simultaneously estimate both the state of the system x and the unknown parameters vector ϑ , as already mentioned the latter is treated as a further state component with no dynamical evolution so that, by considering (1.9)-(1.10) with the additional state equation

$$\vartheta(k+1) = \vartheta(k), \quad (2.1)$$

(with $\vartheta(0)$ having moments given by (1.8)) the system is bilinear with respect to the pair (x, ϑ) . According to the minimum variance polynomial approach for linear and bilinear systems (see [7, 8]), in this work the optimal quadratic filter is investigated, so that the following extended measurements vector is defined:

$$Y(k) = \begin{bmatrix} y(k) \\ y^{[2]}(k) \end{bmatrix}, \quad (2.2)$$

Unfortunately, owing to the nonlinear feature of both the state and output equation, such a problem does not admit a finite dimensional solution. For instance, the extended output equation does not depend only on the augmented state $[x^T(k) \ \vartheta^T(k)]^T$, but also on the nonlinear terms $x^{[2]}(k)$, $\vartheta(k) \otimes x(k)$, $\vartheta(k) \otimes x^{[2]}(k)$, $\vartheta^{[2]}(k) \otimes x^{[2]}(k)$. Such a drawback is overcome by defining an extended state (whose components contain all of the previously mentioned terms), named $X(k)$, with respect to which the extended output becomes linear.

Theorem 2.2 shows the explicit form of the extended state to be considered and the detailed derivation of the extended output equation. Before stating the theorem, it is useful to introduce the following lemma.

Lemma 2.1. *Consider the following random sequences: $\{z(k)\}$, $\{f(k)\}$, $\{g(k)\}$, such that $z(k)$ is independent of $f(h)$ and $g(h) \ \forall k \leq h$, and $\{f(k)\}$ and $\{g(k)\}$ are zero-mean auto and mutually independent. Let $\chi_1(z)$, $\chi_2(z)$, $\eta_1(f)$, $\eta_2(f)$, $\gamma(g)$ suitably dimensioned Borel functions, such that $\{\eta_i(f(k))\}$ and $\{\gamma(g(k))\}$ are sequences of zero-mean random vectors, $i = 1, 2$. Then:*

$$\mathbb{E} \left[\left(\chi_1(z(k)) \otimes \eta_1(f(k)) \right) \left(\chi_2(z(h)) \otimes \eta_2(f(h)) \right)^T \right] = 0 \quad \forall k \leq h, \quad (2.3)$$

$$\mathbb{E} \left[\left(\chi_1(z(k)) \otimes \eta_1(f(k)) \right) \left(\chi_2(z(h)) \otimes \gamma(g(h)) \right)^T \right] = 0 \quad \forall k, h, \quad (2.4)$$

and

$$\mathbb{E} \left[\chi_1(z(k)) \left(\chi_2(z(h)) \otimes \eta_1(f(h)) \right)^T \right] = 0 \quad \forall k \leq h. \quad (2.5)$$

Proof. Note that, by construction, $\eta_i(f(k))$ and $\gamma(g(k))$ are sequences of uncorrelated random vectors. Let $k \leq h$. Then:

$$\begin{aligned} & \mathbb{E} \left[\left(\chi_1(z(k)) \otimes \eta_1(f(k)) \right) \left(\chi_2(z(h)) \otimes \eta_2(f(h)) \right)^T \right] \\ &= st^{-1} \left(\mathbb{E} \left[\chi_2(z(h)) \otimes \eta_2(f(h)) \otimes \chi_1(z(k)) \otimes \eta_1(f(k)) \right] \right) \\ &= st^{-1} \left((I \otimes C^T) \mathbb{E} \left[\chi_2(z(h)) \otimes \chi_1(z(k)) \otimes \eta_1(f(k)) \right] \otimes \mathbb{E} \left[\eta_2(f(h)) \right] \right) = 0. \end{aligned} \quad (2.6)$$

In the same way, equation (2.4) and (2.5) are readily proved. ■

Theorem 2.2. Let $X(k) = [X_1^T(k) \cdots X_7^T(k)]^T$ be the extended state vector defined by

$$\begin{cases} X_1(k) = x(k) \in \mathbb{R}^n \\ X_2(k) = \vartheta(k) \in \mathbb{R}^m \\ X_3(k) = \vartheta(k) \otimes x(k) \in \mathbb{R}^{mn} \\ X_4(k) = x^{[2]}(k) \in \mathbb{R}^{n^2} \\ X_5(k) = \vartheta(k) \otimes x^{[2]}(k) \in \mathbb{R}^{mn^2} \\ X_6(k) = \vartheta^{[2]}(k) \otimes x^{[2]}(k) \in \mathbb{R}^{m^2n^2} \\ X_7(k) = \vartheta^{[2]}(k) \in \mathbb{R}^{m^2} \end{cases} \quad (2.7)$$

Then the dynamical evolution of the quadratic measurements vector $Y(k)$ defined in (2.2) is given by:

$$Y(k) = \mathcal{C}X(k) + \Gamma(k) + \mathcal{W}(k), \quad (2.8)$$

with the matrix \mathcal{C} decomposed according to definition (2.7) as

$$\mathcal{C} = \begin{bmatrix} \mathcal{C}_{11} & O & \mathcal{C}_{13} & O & O & O & O \\ O & O & O & \mathcal{C}_{24} & \mathcal{C}_{25} & \mathcal{C}_{26} & O \end{bmatrix}, \quad (2.9)$$

where

$$\begin{aligned} \mathcal{C}_{11} &= C_0, & \mathcal{C}_{13} &= \tilde{C}, & \mathcal{C}_{24} &= C_0^{[2]}, \\ \mathcal{C}_{25} &= M_1^2(q)(\tilde{C} \otimes C_0), & \mathcal{C}_{26} &= \tilde{C}^{[2]}(I_m \otimes C_{nm,n}^T) \end{aligned} \quad (2.10)$$

$\Gamma(k) = [O_q^T \quad \xi_2^{wT}(k)]^T$, $\mathcal{W}(k) = [\mathcal{W}_1^T(k) \quad \mathcal{W}_2^T(k)]^T$, where $\mathcal{W}_1(k) = w(k)$ and $\mathcal{W}_2(k)$ is the zero-mean, multiplicative noise sequence defined by:

$$\mathcal{W}_2(k) = w^{[2]}(k) - \xi_2^w(k) + D_1(X_1(k) \otimes w(k)) + D_2(X_3(k) \otimes w(k)), \quad (2.11)$$

with

$$D_1 = M_1^2(q)(C_0 \otimes I_q), \quad D_2 = M_1^2(q)(\tilde{C} \otimes I_q). \quad (2.12)$$

Moreover, $\mathcal{W}(k)$ is a sequence of uncorrelated random vectors.

Proof. The output equation immediately comes from (1.10). For the dynamics of its Kronecker square, by using the Kronecker product properties:

$$\begin{aligned} y^{[2]}(k) &= \left[C_0 x + \tilde{C}(\vartheta(k) \otimes x(k)) + w(k) \right]^{[2]} \\ &= \mathcal{C}_{24} X_4(k) + \mathcal{C}_{25} X_5(k) + \mathcal{C}_{26} X_6(k) + D_1(X_1(k) \otimes w(k)) + D_2(X_3(k) \otimes w(k)) + w^{[2]}(k), \end{aligned} \quad (2.13)$$

so that, according to (2.9), (2.10), (2.11) and (2.12), (2.8) is achieved. Moreover, note that from (2.11), the extended noise component $\mathcal{W}_2(k)$ can be written as $\mathcal{W}_2(k) = \chi_1(X(k)) \otimes \eta_1(w(k)) + \chi_2(X(k)) \otimes \eta_2(w(k)) + \chi_3(X(k)) \otimes \eta_3(w(k))$, with the sequences $\{X(k)\}$, $\{w(k)\}$ and $\{\eta_i(\cdot)\}$ satisfying the hypotheses of Lemma 2.1, so that

$$\mathbb{E} \left[\chi_i(X(k)) \otimes \eta_i(w(k)) \otimes \chi_j(X(h)) \otimes \eta_j(w(h)) \right] = 0, \quad (2.14)$$

$\forall i, j = 1, 2, 3, \forall k \neq h$, from which follows that $\mathcal{W}(k)$ is a sequence of zero-mean random vectors. ■

Theorem 2.3. *The equation of the extended state $X(k)$ defined by (2.7) in Theorem 2.2 is given by:*

$$X(k+1) = \mathcal{A}X(k) + \mathcal{U}(k) + \mathcal{V}(k) + \varphi(X(k)) \quad (2.15)$$

where the matrix \mathcal{A} is decomposed according to definition (2.7) as

$$\mathcal{A} = \begin{pmatrix} \mathcal{A}_{11} & O & \mathcal{A}_{13} & O & O & O & O \\ O & I_m & O & O & O & O & O \\ O & O & \mathcal{A}_{33} & O & O & O & O \\ O & O & O & \mathcal{A}_{44} & \mathcal{A}_{45} & \mathcal{A}_{46} & O \\ O & \mathcal{A}_{52} & O & O & \mathcal{A}_{55} & \mathcal{A}_{56} & O \\ O & O & O & O & O & \mathcal{A}_{66} & \mathcal{A}_{67} \\ O & O & O & O & O & O & I_{m^2} \end{pmatrix} \quad (2.16)$$

with

$$\begin{aligned} \mathcal{A}_{11} &= A_0, & \mathcal{A}_{13} &= \tilde{A}, & \mathcal{A}_{33} &= I_m \otimes A_0, \\ \mathcal{A}_{44} &= A_0^{[2]}, & \mathcal{A}_{45} &= M_1^2(n)(\tilde{A} \otimes A_0), \\ \mathcal{A}_{46} &= \tilde{A}^{[2]}(I_m \otimes C_{nm,n}^T), & \mathcal{A}_{52} &= I_m \otimes \xi_2^v, \\ \mathcal{A}_{55} &= I_m \otimes A_0^{[2]}, & \mathcal{A}_{66} &= I_{m^2} \otimes A_0^{[2]}, \\ \mathcal{A}_{56} &= I_m \otimes (M_1^2(n)(\tilde{A} \otimes A_0)), & \mathcal{A}_{67} &= I_{m^2} \otimes \xi_2^v, \end{aligned} \quad (2.17)$$

where

$$\mathcal{U}(k) = [O_n^T \quad O_m^T \quad O_{mn}^T \quad \xi_2^{vT}(k) \quad O_{mn^2}^T \quad O_{m^2n^2}^T \quad O_{m^2}^T]^T. \quad (2.18)$$

$\mathcal{U}(k)$ is a deterministic input and $\mathcal{V}(k) = [\mathcal{V}_1^T(k) \quad O^T \quad \mathcal{V}_3^T(k) \quad \mathcal{V}_4^T(k) \quad \mathcal{V}_5^T(k) \quad \mathcal{V}_6^T(k) \quad O^T]^T$ is a zero-mean, uncorrelated multiplicative noise sequence given by

$$\left\{ \begin{array}{l} \mathcal{V}_1(k) = v(k) \in \mathbb{R}^n \\ \mathcal{V}_3(k) = X_2(k) \otimes v(k) \in \mathbb{R}^{mn} \\ \mathcal{V}_4(k) = v^{[2]}(k) - \xi_2^v(k) + Q_1(X_1(k) \otimes v(k)) + Q_2(X_3(k) \otimes v(k)) \in \mathbb{R}^{m^2} \\ \mathcal{V}_5(k) = X_2(k) \otimes (v^{[2]}(k) - \xi_2^v(k)) + Q_3(X_3(k) \otimes v(k)) \\ \quad + Q_4(X_2(k) \otimes X_3(k) \otimes v(k)) \in \mathbb{R}^{mn^2} \\ \mathcal{V}_6(k) = X_7(k) \otimes (v^{[2]}(k) - \xi_2^v(k)) + Q_5(X_2(k) \otimes X_3(k) \otimes v(k)) \\ \quad + Q_6(X_7(k) \otimes X_3(k) \otimes v(k)) \in \mathbb{R}^{m^2n^2} \end{array} \right. \quad (2.19)$$

where

$$\begin{aligned} Q_1 &= M_1^2(n)(A_0 \otimes I_n) \\ Q_2 &= M_1^2(n)(\tilde{A} \otimes I_n) \\ Q_3 &= I_m \otimes (M_1^2(n)(A_0 \otimes I_n)) \\ Q_4 &= I_m \otimes (M_1^2(n)(\tilde{A} \otimes I_n)) \\ Q_5 &= I_{m^2} \otimes (M_1^2(n)(A_0 \otimes I_n)) \\ Q_6 &= I_{m^2} \otimes (M_1^2(n)(\tilde{A} \otimes I_n)). \end{aligned} \quad (2.20)$$

Moreover $\mathcal{V}(k)$ is a sequence of uncorrelated random variables. All the nonlinearities are embodied into the term $\varphi(X(k)) = [O_n^T \ O_m^T \ \varphi_3(X) \ O_{n^2}^T \ \varphi_5(X) \ \varphi_6(X) \ O_{m^2}^T]^T$ which is a nonlinear map of the components of the extended state X given by:

$$\begin{cases} \varphi_3(X(k)) = \Phi_3(X_2(k) \otimes X_3(k)) \in \mathbb{R}^{mn} \\ \varphi_5(X(k)) = \Phi_5(X_2(k) \otimes X_6(k)) \in \mathbb{R}^{mn^2} \\ \varphi_6(X(k)) = \Phi_{6,a}(X_2(k) \otimes X_6(k)) + \Phi_{6,b}(X_7(k) \otimes X_6(k)) \in \mathbb{R}^{m^2n^2} \end{cases} \quad (2.21)$$

with

$$\begin{aligned} \Phi_3 &= I_m \otimes \tilde{A}, \\ \Phi_5 &= I_m \otimes \left(\tilde{A}^{[2]} (I_m \otimes C_{nm,n}^T) \right), \\ \Phi_{6,a} &= I_{m^2} \otimes \left(M_1^2(n) \left(\tilde{A} \otimes A_0 \right) \right), \\ \Phi_{6,b} &= I_{m^2} \otimes \left(\tilde{A}^{[2]} (I_m \otimes C_{nm,n}^T) \right). \end{aligned} \quad (2.22)$$

Proof. The dynamics of X_1 , X_2 and X_7 easily comes from (1.9) and (2.1). For the others:

$$\begin{aligned} X_3(k+1) &= \vartheta(k) \otimes \left[A_0 x(k) + \tilde{A} (\vartheta(k) \otimes x(k)) + v(k) \right] \\ &= (I_m \otimes A_0) (\vartheta(k) \otimes x(k)) + (I_m \otimes \tilde{A}) (\vartheta^{[2]}(k) \otimes x(k)) + \vartheta(k) \otimes v(k) \\ &= \mathcal{A}_{33} X_3(k) + \mathcal{V}_2(k) + \varphi_3(X(k)), \end{aligned} \quad (2.23)$$

$$\begin{aligned} X_4(k+1) &= \left[A_0 x(k) + \tilde{A} (\vartheta(k) \otimes x(k)) + v(k) \right]^{[2]} \\ &= \mathcal{A}_{44} X_4(k) + \mathcal{A}_{45} X_5(k) + \mathcal{A}_{46} X_6(k) + v^{[2]}(k) + Q_1 (X_1(k) \otimes v(k)) + Q_2 (X_3(k) \otimes v(k)), \end{aligned} \quad (2.24)$$

$$\begin{aligned} X_5(k+1) &= \vartheta(k) \otimes X_4(k+1) \\ &= \mathcal{A}_{55} X_5(k) + \mathcal{A}_{56} X_6(k) + \Phi_5 (X_2(k) \otimes X_6(k)) + Q_3 (X_3(k) \otimes v(k)) \\ &\quad + Q_4 (X_2(k) \otimes X_3(k) \otimes v(k)) + X_2(k) \otimes v^{[2]}(k) \end{aligned} \quad (2.25)$$

$$\begin{aligned} X_6(k+1) &= \vartheta(k) \otimes X_5(k+1) \\ &= \mathcal{A}_{66} X_6(k) + \Phi_{6,a} (X_2(k) \otimes X_6(k)) + \Phi_{6,b} (X_7(k) \otimes X_6(k)) + Q_5 (X_2(k) \otimes X_3(k) \otimes v(k)) \\ &\quad + Q_6 (X_7(k) \otimes X_3(k) \otimes v(k)) + X_7(k) \otimes v^{[2]}(k), \end{aligned} \quad (2.26)$$

according to the Kronecker product properties. The proof is completed by following the same sketch as in Theorem 2.2. The fact that $\{\mathcal{V}(k)\}$ is a sequence of uncorrelated random vectors comes readily from Lemma 2.1. ■

3. The filtering algorithm

As previously mentioned, the proposed filtering algorithm is based on the EKF applied to the extended model equations:

$$\begin{aligned} X(k+1) &= \mathcal{A}X(k) + \mathcal{U}(k) + \mathcal{V}(k) + \varphi(X(k)) \\ Y(k) &= \mathcal{C}X(k) + \Gamma(k) + \mathcal{W}(k). \end{aligned} \quad (3.1)$$

where it is worthwhile to remark that the unique nonlinearities appear in the term $\varphi(X)$. According to the extended state formulation, the unknown parameter ϑ is treated as a random vector with known initial distribution. This means that all the moments of $\vartheta(k)$, $\forall k$ are available, and can be used in the computation of both the covariance matrices of the extended noises and the mean value of the extended state, as it will be clearer later.

As it is expected from the optimal polynomial approach, the quadratic filter requires the knowledge of the first four noises moments, given by (1.6), appearing in the covariance matrices of the extended noises (see Appendix B). Such an information is not used by the classical EKF, which just requires the first and second order moments: this justifies the adopted procedure with respect to the classical EKF. In order to compute the covariances of the sequences $\mathcal{V}(k)$ and $\mathcal{W}(k)$ it is necessary the knowledge of the expected values $Z_i(k) = \mathbb{E}[X_i(k)]$, whose dynamical equations, according to (2.15), are given by

$$Z(k+1) = \mathcal{A}Z(k) + \mathcal{U}(k) + \mathbb{E}[\varphi(X(k))]. \quad (3.2)$$

Note that, thanks to (2.21), the nonlinear term in (3.2) contains the following mean values:

$$\mathbb{E}[\vartheta^{[i]}(k) \otimes x^{[j]}(k)] = \mathbb{E}[\vartheta^{[i]}(k)] \otimes \mathbb{E}[x^{[j]}(k)], \quad (3.3)$$

with $i = 2, 3, 4$ and $j = 1, 2$. The first mean value is computed from the statistics of $\vartheta(0)$, according to the fact that the parameter $\vartheta(k)$ obeys to the stationary equation (2.1); the other can be referred to as $Z_1(k)$ or $Z_4(k)$ depending on whether $j = 1$ or $j = 2$. That means, equation (3.2) can be rewritten as:

$$Z(k+1) = \mathcal{A}Z(k) + \mathcal{U}(k) + \tilde{\varphi}(Z(k)), \quad (3.4)$$

with $\tilde{\varphi}(Z)$ a suitable defined nonlinear map. Moreover the extended state and measurement noises are uncorrelated, that means, for $i = 1, \dots, 7$ and $j = 1, 2$:

$$\mathbb{E}[\mathcal{V}_i(k)\mathcal{W}_j^T(h)] = 0, \quad \forall k, h, \quad (3.5)$$

as it comes from Lemma 2.1. Finally, the first order approximation of the nonlinear map $\varphi(X(k))$ around the current estimate $\hat{X}(k)$ is needed. According to (2.21) it comes that each nonlinear term has the form:

$$\Phi(X_i \otimes X_j), \quad i, j \in \{1, \dots, 7\} \quad (3.6)$$

so that, its linear approximation is:

$$\begin{aligned} X_i(k) \otimes X_j(k) &\cong \hat{X}_i(k) \otimes \hat{X}_j(k) \\ &+ (I \otimes \hat{X}_j(k)) \cdot (X_i(k) - \hat{X}_i(k)) + (\hat{X}_i(k) \otimes I)(X_j(k) - \hat{X}_j(k)). \end{aligned} \quad (3.7)$$

Similarly to the classical EKF [10], the proposed Filtering Algorithm is a recursive estimation scheme whose performances depend on the specific application. A better behavior with respect to the classical EKF is expected because a higher degree approximation of the nonlinearity appearing in the extended system is adopted. Here follows a description of the algorithm, where some formulas reported in the Appendix are used.

The proposed Filtering Algorithm

I) Computation of the initial conditions of the filter:

$$\widehat{X}(0|-1) = \mathbb{E}[X(0)] = \zeta_1^0, \quad \text{a priori estimate of the initial extended state;} \quad (3.8)$$

$$P_P(0) = \text{Cov}(X(0)), \quad \text{covariance of the a priori estimate;} \quad (3.9)$$

$$Z(0) = \zeta_1^0, \quad \text{initialization of (3.2);} \quad (3.10)$$

$$k = -1, \quad \text{inicialization of the counter;} \quad (3.11)$$

II) computation of the extended output prediction:

$$\widehat{Y}(k+1|k) = \mathcal{C}\widehat{X}(k+1|k) + \Gamma(k+1) \quad (3.12)$$

(note that the matrix \mathcal{C} does not depend on time);

III) computation of the output noise covariance:

$$\Psi^{\mathcal{W}}(k) = \Psi^{\mathcal{W}}(k, \widehat{X}(k)) \quad \text{using (B.22);} \quad (3.13)$$

(note that $\Psi^{\mathcal{W}}(k+1)$ requires the knowledge of $Z_i(k+1)$, $i = 1, \dots, 7$)

IV) computation of the Kalman gain:

$$K(k+1) = P_P(k+1)\mathcal{C}^T \left(\mathcal{C}P_P(k+1)\mathcal{C}^T + \Psi^{\mathcal{W}}(k+1) \right)^{-1} \quad (3.14)$$

V) computation of the error covariance matrix:

$$P(k+1) = \left(I_N - K(k+1)\mathcal{C} \right) P_P(k+1), \quad (3.15)$$

with $N = n + m + mn + n^2 + mn^2 + m^2n^2 + m^2$;

VI) computation of the state estimate $\widehat{X}(k+1)$:

$$\widehat{X}(k+1) = \widehat{X}(k+1|k) + K(k+1)(Y(k+1) - \widehat{Y}(k+1|k)); \quad (3.16)$$

VII) increment of the counter: $k = k + 1$;

VIII) computation of the extended state prediction:

$$\widehat{X}(k+1|k) = \mathcal{A}\widehat{X}(k) + \mathcal{U}(k) + \varphi(\widehat{X}(k)); \quad (3.17)$$

IX) computation of the first order approximation of the extended state equation around $\widehat{X}(k)$:

$$\widetilde{\mathcal{A}}(k) = \mathcal{A} + \frac{\partial \varphi(X)}{\partial X} \Big|_{\widehat{X}(k)}, \quad \text{using (D.3)} \quad (3.18)$$

$$\Psi^{\mathcal{V}}(k) = \Psi^{\mathcal{V}}(k, \widehat{X}(k)) \quad \text{using (B.21)} \quad (3.19)$$

(also in this case, $\Psi^{\mathcal{V}}(k)$ requires $Z_i(k)$, $i = 1, \dots, 7$);

X) computation of the one-step prediction error covariance matrix:

$$P_P(k+1) = \tilde{\mathcal{A}}(k)P(k)\tilde{\mathcal{A}}(k)^T + \Psi^{\mathcal{V}}(k); \quad (3.20)$$

XI) computation of $Z_i(k+1)$, $i = 1, \dots, 7$, using (3.2). GOTO STEP **II**.

4. Simulation results

Some significative results are here reported in order to show the effectiveness of the proposed algorithm.

Consider the following system:

$$\begin{aligned} x_1(k+1) &= x_2(k) + v_1(k), \\ x_2(k+1) &= \vartheta x_1(k) + 0.3x_2(k) + v_2(k), \\ y(k) &= x_1(k) - x_2(k) + w(k), \end{aligned} \quad (4.1)$$

with v_1, v_2, w mutually independent noise sequences with asymmetrical distributions

$$P\{v_i(k) = -0.4\} = 0.9, \quad P\{v_i(k) = 3.6\} = 0.1, \quad (4.2)$$

for $i = 1, 2$ and

$$P\{w(k) = 1.2\} = 0.2, \quad P\{w(k) = -0.3\} = 0.8. \quad (4.3)$$

These distributions are assumed known, so that the moments up to the 4th order can be computed, as required by the estimation algorithm. For the intialization of the algorithm also the distribution of the initial state x_0 is required. For this purpose, the initial state x_0 is assumed gaussian with zero mean and covariance I_n .

It is assumed that the system (4.1) is stable. This means that necessarily the unknown parameter ϑ belongs to the interval $[-1, 0.7]$. In the reported simulation the true values for the parameter ϑ and the initial state x_0 are

$$\theta = 0.4, \quad x_0 = \begin{bmatrix} 10 \\ 8 \end{bmatrix}, \quad (4.4)$$

The initial estimate of the parameter is chosen in the center of the admissible interval, that is $\hat{\vartheta}_0 = -1.5$, whereas the initial state estimate is taken equal to its mean value 0. In order to test the goodness of the quadratic approach, the proposed filtering algorithm has been compared with the Extended Kalman Filter. The reported results refer to simulation over 1000 steps. Figures 1 and 2 report the true states, the quadratic estimates and the EKF estimates, while figure 3 report the true parameter and its quadratic and EKF estimates. Due to the high frequency behavior of the state only the first 50 simulation step are plotted in fig.'s 1 and 2.

The computation of the mean square estimation errors over the 1000 simulation steps, for both the quadratic algorithm and the EKF algorithm, gives the following result

$$\begin{aligned} \sigma_{1,\text{Quad}}^2 &= 0.8223, & \sigma_{1,\text{EKF}}^2 &= 1.4404, \\ \sigma_{2,\text{Quad}}^2 &= 0.4210, & \sigma_{2,\text{EKF}}^2 &= 0.8826. \end{aligned} \quad (4.5)$$

These data show that in this example the proposed algorithm improves in a very effective manner the performances of the EKF, by decreasing the mean square errors for more than 40%.

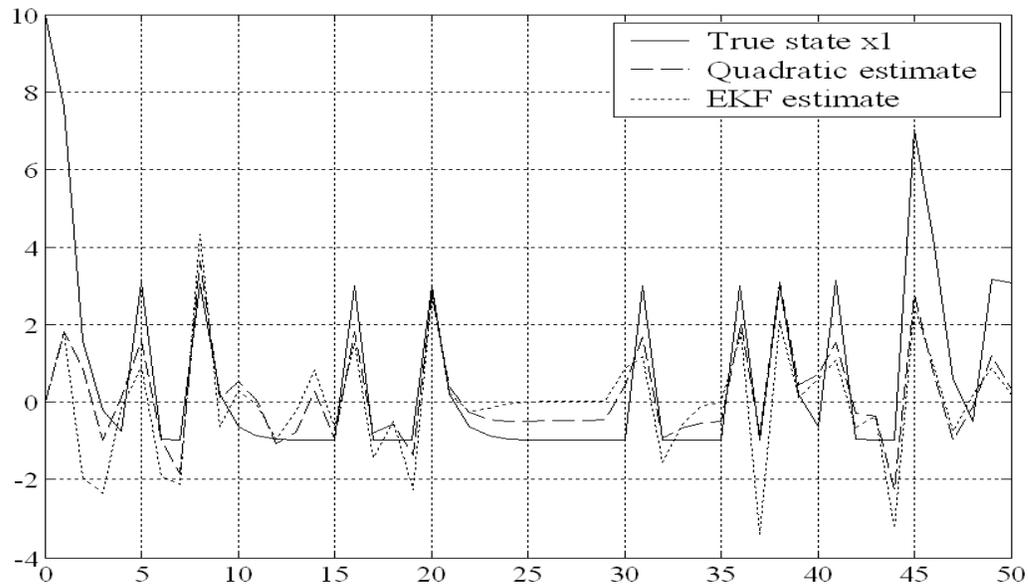


Figure 1: True and estimated state: the first component.

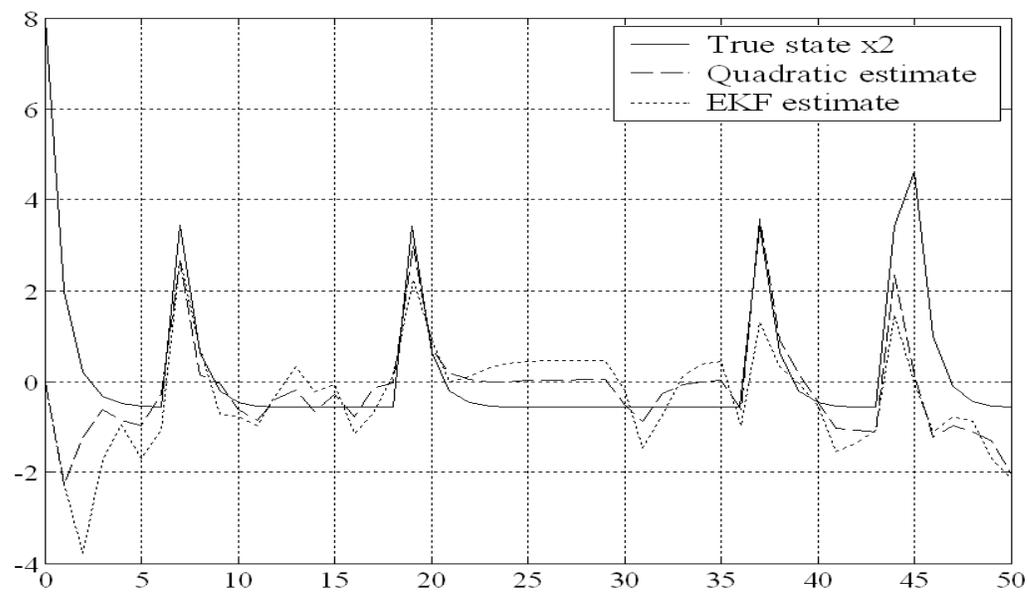


Figure 2: True and estimated state: the second component.

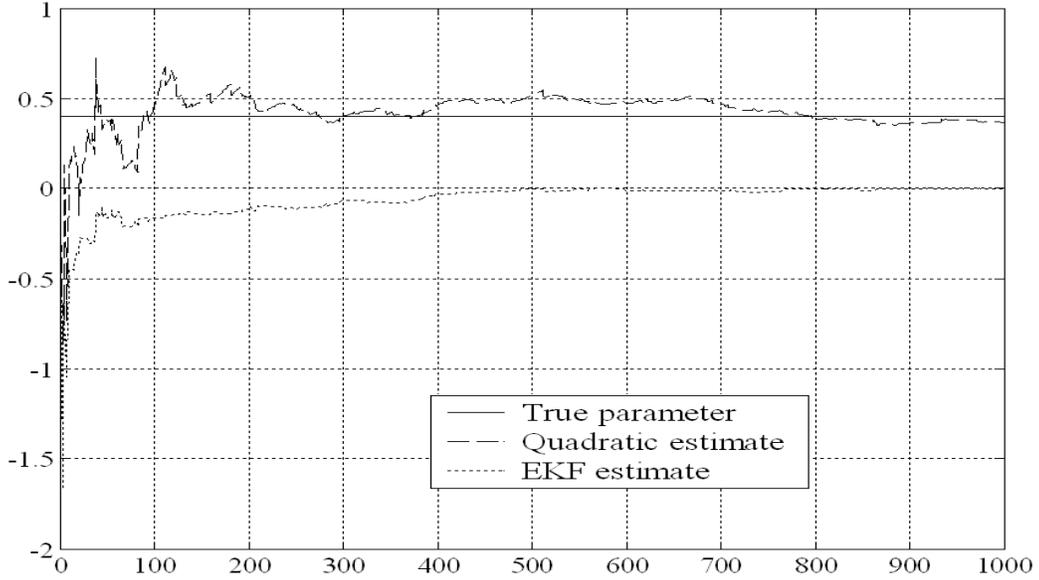


Figure 3: True and estimated parameter ϑ_0 .

5. Conclusions and future developments

The problem of the simultaneous state and parameters estimation for the class of systems described by (1.4)-(1.5) affected by additive noises, not necessarily Gaussian, has been investigated in this paper. The filtering algorithm here proposed is based on two steps: first the system is extended by considering the state variables shown by Lemma 2.2, obtaining a global system where all the nonlinearities are enclosed in the function φ whose components are given by (2.21). Next, the filter of the approximating system is computed, by exploiting not only the measurements but also their square. Numerical simulations show good performances of the proposed filter. In particular, in the presence of nongaussian noises, in most cases the proposed algorithm provides better state estimate of the uncertain system when compared to the Extended Kalman Filter.

The extension of these results to the Borel subspace of the polynomial transformations of the output will be object of future research.

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 [7]. 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

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

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

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

$$M = [m_1 \quad m_2 \quad \dots \quad m_s], \quad (\text{A.2})$$

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

$$M = \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{bmatrix}. \quad (\text{A.3})$$

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 \mathbb{R}^μ , and r be a divisor of μ . Then the $r \times (\mu/r)$ matrix given by $M = \text{st}^{-1}(m, r)$ is defined so that:

$$\text{st}(M) = m. \quad (\text{A.4})$$

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

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

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

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

$$M^{[0]} = 1 \in \mathbb{R}, \quad (\text{A.6a})$$

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

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

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

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

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

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

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

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

$$\text{tr}(A \otimes B) = \text{tr}(A) \cdot \text{tr}(B) \quad (\text{A.7g})$$

Other useful properties can be found in [7].

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

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

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 $A \in \mathbb{R}^{r \times s}$, $B \in \mathbb{R}^{n \times m}$, it is:

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

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

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

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:

$$A \otimes B \otimes C \otimes D = (I_{n_A} \otimes C_{n_C n_D, n_B}^T) (A \otimes C \otimes D \otimes B) (I_{m_A} \otimes C_{m_C m_D, m_B}). \quad (\text{A.11})$$

Proof.

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

$$\begin{aligned} A \otimes B \otimes C \otimes D &= (A \otimes (B \otimes (C \otimes D))) \\ &= (A \otimes (C_{n_C n_D, n_B}^T (C \otimes D \otimes B) C_{m_C m_D, m_B})) \\ &= (I_{n_A} \otimes C_{n_C n_D, n_B}^T) (A \otimes ((C \otimes D \otimes B) C_{m_C m_D, m_B})) \\ &= (I_{n_A} \otimes C_{n_C n_D, n_B}^T) (A \otimes C \otimes D \otimes B) (I_{m_A} \otimes C_{m_C m_D, m_B}). \end{aligned} \quad (\text{A.12})$$

■

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

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

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

Lemma A.5. Let $a, b \in \mathbb{R}^n$ and $l \in \mathbb{N}$. Then

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

with the sequence $\{G_l(n) = C_{n,n}^T\}$ given by the following recursive equations

$$G_1(n) = C_{n,n}^T, \quad (\text{A.15a})$$

$$G_l(n) = (I_{n,1} \otimes G_{l-1}(n)) \cdot (G_1(n) \otimes I_{n,l-1}), \quad l > 1, \quad (\text{A.15b})$$

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

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

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

$$(a + b)^{[h]} = \sum_{k=0}^h M_k^h(n) (a^{[k]} \otimes b^{[h-k]}) \quad (\text{A.16})$$

constitute a set of matrices $\{M_0^h(n), \dots, M_h^h(n); \quad M_k^h(n) \in \mathbb{R}^{n^h \times n^h}\}$ such that:

$$M_h^h(n) = M_0^h(n) = I_{n,h}, \quad (\text{A.17a})$$

$$M_j^h(n) = (M_j^{h-1}(n) \otimes I_{n,1}) + (M_{j-1}^{h-1}(n) \otimes I_{n,1}) \cdot (I_{n,j-1} \otimes G_{h-j}(n)), \quad 1 \leq j \leq h-1, \quad (\text{A.17b})$$

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 $a_1 + \dots + a_p$, $a_i \in \mathbb{R}^n$, $1 \leq i \leq p$, $p \in \mathbb{N}$, its h -th Kronecker power admits a representation as:

$$(a_1 + a_2 + \dots + a_p)^{[h]} = \sum_{\substack{h_1, \dots, h_p \geq 0 \\ h_1 + \dots + h_p = h}} M_{h_1, \dots, h_p}^h (a_1^{[h_1]} \otimes a_2^{[h_2]} \otimes \dots \otimes a_p^{[h_p]}) \quad (\text{A.18})$$

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

$$M_{l_1, \dots, l_s}^l = O_{n^l \times n^l}. \quad (\text{A.19})$$

Moreover the following statement can be proved:

Lemma A.7. The matrices $M_{h_1, \dots, h_p}^h \in \mathbb{R}^{n^h \times n^h}$ in (A.18) satisfy the recursive formula:

$$M_{h_1, \dots, h_p}^h = I_1, \quad h = 1 \quad (\text{A.20a})$$

$$M_{h_1, \dots, h_p}^h = \sum_{1 \leq i \leq p-1} (M_{h_1, \dots, h_{i-1}, \dots, h_p}^{h-1} \otimes I_1) (I_{h_1 + \dots + h_{i-1}} \otimes G_{h_{i+1} + \dots + h_p}) \\ + M_{h_1, \dots, h_{p-1}}^{h-1} \otimes I_1, \quad h > 1. \quad (\text{A.20b})$$

Proof. The proof can be found in [7]. ■

B. Computation of the covariances matrices

The explicit expression for the covariances of the extended state and output noises can be directly derived from (2.11) and (2.19), by exploiting the Kronecker product properties and the mutual independence between each state component and the noise at the same instant.

For convenience, the block expressions of the noise are reported:

$$\mathcal{V}(k) = [\mathcal{V}_1(k) \quad 0 \quad \mathcal{V}_3(k) \quad \mathcal{V}_4(k) \quad \mathcal{V}_5(k) \quad \mathcal{V}_6(k) \quad 0]^T \in \mathbb{R}^N, \quad (\text{B.1})$$

$$\mathcal{W}(k) = [\mathcal{W}_1(k) \quad \mathcal{W}_2(k)]^T \in \mathbb{R}^{q+q^2}, \quad (\text{B.2})$$

so that the covariances can also be blocks-built. By defining $\Psi_{i,j}^{\mathcal{V}} = \mathbb{E} [\mathcal{V}_i(k) \mathcal{V}_j^T(k)]$ and $\Psi_{i,j}^{\mathcal{W}} = \mathbb{E} [\mathcal{W}_i(k) \mathcal{W}_j^T(k)]$ respectively the i, j -th block of the state noise and of the output noise covariance and taking into account the explicit expressions of every single terms, the following

results hold:

$$\Psi_{1,1}^{\mathcal{Y}}(k) = st^{-1}\xi_2^v, \quad (\text{B.3})$$

$$\Psi_{1,3}^{\mathcal{Y}}(k) = Z_2^T \otimes st^{-1}\xi_2^v, \quad (\text{B.4})$$

$$\Psi_{4,1}^{\mathcal{Y}}(k) = st^{-1}\xi_3^v + Q_1 (Z_1 \otimes st^{-1}\xi_2^v) + Q_2 (Z_3 \otimes st^{-1}\xi_2^v), \quad (\text{B.5})$$

$$\Psi_{5,1}^{\mathcal{Y}}(k) = Z_2 \otimes st^{-1}\xi_3^v + Q_3 (Z_3 \otimes st^{-1}\xi_2^v) + Q_4 (\mathbb{E} [X_2 \otimes X_3] \otimes st^{-1}\xi_2^v), \quad (\text{B.6})$$

$$\Psi_{6,1}^{\mathcal{Y}}(k) = Z_7 \otimes st^{-1}\xi_3^v + Q_5 (\mathbb{E} [X_2 \otimes X_3] \otimes st^{-1}\xi_2^v) + Q_6 (\mathbb{E} [X_7 \otimes X_3] \otimes st^{-1}\xi_2^v), \quad (\text{B.7})$$

$$\Psi_{3,3}^{\mathcal{Y}}(k) = st^{-1}Z_7 \otimes st^{-1}\xi_2^v, \quad (\text{B.8})$$

$$\Psi_{4,3}^{\mathcal{Y}}(k) = Z_2^T \otimes st^{-1}\xi_3^v + Q_1 (\mathbb{E} [X_1 X_2^T] \otimes st^{-1}\xi_2^v) + Q_2 (\mathbb{E} [X_3 X_2^T] \otimes st^{-1}\xi_2^v), \quad (\text{B.9})$$

$$\begin{aligned} \Psi_{5,3}^{\mathcal{Y}}(k) &= \mathbb{E} [X_2 X_2^T] \otimes st^{-1}\xi_3^v + Q_3 (\mathbb{E} [X_3 X_2^T] \otimes st^{-1}\xi_2^v) + \\ &+ Q_4 (\mathbb{E} [(X_2 \otimes X_3) X_2^T] \otimes st^{-1}\xi_2^v), \end{aligned} \quad (\text{B.10})$$

$$\begin{aligned} \Psi_{6,3}^{\mathcal{Y}}(k) &= \mathbb{E} [X_7 X_2^T] \otimes st^{-1}\xi_3^v + Q_5 (\mathbb{E} [(X_2 \otimes X_3) X_2^T] \otimes st^{-1}\xi_2^v) \\ &+ Q_6 (\mathbb{E} [(X_7 \otimes X_3) X_2^T] \otimes st^{-1}\xi_2^v), \end{aligned} \quad (\text{B.11})$$

$$\begin{aligned} \Psi_{4,4}^{\mathcal{Y}}(k) &= st^{-1}\xi_4^v - \xi_2^v \xi_2^{vT} + (Z_1^T \otimes st^{-1}\xi_3^v) Q_1^T + (Z_3^T \otimes st^{-1}\xi_3^v) Q_2^T + Q_1 (Z_1 \otimes st^{-1}\xi_3^v) \\ &+ Q_2 (Z_3 \otimes st^{-1}\xi_3^v) + Q_1 (\mathbb{E} [X_1 X_1^T] \otimes st^{-1}\xi_2^v) Q_1^T + Q_1 (\mathbb{E} [X_1 X_3^T] \otimes st^{-1}\xi_2^v) Q_2^T \\ &+ Q_2 (\mathbb{E} [X_3 X_1^T] \otimes st^{-1}\xi_2^v) Q_1^T + Q_2 (\mathbb{E} [X_3 X_3^T] \otimes st^{-1}\xi_2^v) Q_2^T, \end{aligned} \quad (\text{B.12})$$

$$\begin{aligned} \Psi_{5,4}^{\mathcal{Y}}(k) &= Z_2 \otimes st^{-1}\xi_4^v - Z_2 \otimes (\xi_2^v \xi_2^{vT}) + (\mathbb{E} [X_2 X_1^T] \otimes st^{-1}\xi_3^v) Q_1^T + (\mathbb{E} [X_2 X_3^T] \otimes st^{-1}\xi_3^v) Q_2^T \\ &+ Q_3 (Z_3 \otimes st^{-1}\xi_3^v) + Q_4 (\mathbb{E} [X_2 \otimes X_3] \otimes st^{-1}\xi_3^v) + Q_3 (\mathbb{E} [X_3 X_3^T] \otimes st^{-1}\xi_2^v) Q_2^T \\ &+ Q_3 (\mathbb{E} [X_3 X_1^T] \otimes st^{-1}\xi_2^v) Q_1^T + Q_4 (\mathbb{E} [(X_2 \otimes X_3) X_1^T] \otimes st^{-1}\xi_2^v) Q_1^T \\ &+ Q_4 (\mathbb{E} [(X_2 \otimes X_3) X_3^T] \otimes st^{-1}\xi_2^v) Q_2^T, \end{aligned} \quad (\text{B.13})$$

$$\begin{aligned} \Psi_{6,4}^{\mathcal{Y}}(k) &= Z_7 \otimes st^{-1}\xi_4^v - Z_7 \otimes (\xi_2^v \xi_2^{vT}) + (\mathbb{E} [X_7 X_1^T] \otimes st^{-1}\xi_3^v) Q_1^T + (\mathbb{E} [X_7 X_3^T] \otimes st^{-1}\xi_3^v) Q_2^T \\ &+ Q_5 (\mathbb{E} [X_2 \otimes X_3] \otimes st^{-1}\xi_3^v) + Q_6 (\mathbb{E} [X_7 \otimes X_3] \otimes st^{-1}\xi_3^v) \\ &+ Q_5 (\mathbb{E} [(X_2 \otimes X_3) X_1^T] \otimes st^{-1}\xi_2^v) Q_1^T + Q_5 (\mathbb{E} [(X_2 \otimes X_3) X_3^T] \otimes st^{-1}\xi_2^v) Q_2^T \\ &+ Q_6 (\mathbb{E} [(X_7 \otimes X_3) X_1^T] \otimes st^{-1}\xi_2^v) Q_1^T + Q_6 (\mathbb{E} [(X_7 \otimes X_3) X_3^T] \otimes st^{-1}\xi_2^v) Q_2^T, \end{aligned} \quad (\text{B.14})$$

$$\begin{aligned} \Psi_{5,5}^{\mathcal{Y}}(k) &= \mathbb{E} (X_2 X_2^T) \otimes st^{-1}\xi_4^v - \mathbb{E} (X_2 X_2^T) \otimes (\xi_2^v \xi_2^{vT}) + (\mathbb{E} (X_2 X_3^T) \otimes st^{-1}\xi_3^v) Q_3^T \\ &+ (\mathbb{E} [X_2 (X_2 \otimes X_3)^T] \otimes st^{-1}\xi_3^v) Q_4^T + Q_3 (\mathbb{E} [X_3 X_2^T] \otimes st^{-1}\xi_3^v) \\ &+ Q_3 (\mathbb{E} [X_3 (X_2 \otimes X_3)^T] \otimes st^{-1}\xi_2^v) Q_4^T + Q_4 (\mathbb{E} [(X_2 \otimes X_3) X_2^T] \otimes st^{-1}\xi_3^v) \\ &+ Q_3 (\mathbb{E} [X_3 X_3^T] \otimes st^{-1}\xi_2^v) Q_3^T + Q_4 (\mathbb{E} [(X_2 \otimes X_3) X_3^T] \otimes st^{-1}\xi_3^v) Q_3^T \\ &+ Q_4 (\mathbb{E} [(X_2 \otimes X_3) (X_2 \otimes X_3)^T] \otimes st^{-1}\xi_2^v) Q_4^T, \end{aligned} \quad (\text{B.15})$$

$$\begin{aligned} \Psi_{6,5}^{\mathcal{Y}}(k) &= \mathbb{E} (X_7 X_2^T) \otimes st^{-1}\xi_4^v - \mathbb{E} (X_7 X_2^T) \otimes (\xi_2^v \xi_2^{vT}) + (\mathbb{E} (X_7 X_3^T) \otimes st^{-1}\xi_3^v) Q_3^T \\ &+ (\mathbb{E} [x_7 (X_2 \otimes X_3)^T] \otimes st^{-1}\xi_3^v) Q_4^T + Q_5 (\mathbb{E} [(X_2 \otimes X_3) X_2^T] \otimes st^{-1}\xi_3^v) \\ &+ Q_5 (\mathbb{E} [(X_2 \otimes X_3) X_3^T] \otimes st^{-1}\xi_2^v) Q_3^T + Q_6 (\mathbb{E} [(X_7 \otimes X_3) X_2^T] \otimes st^{-1}\xi_3^v) \\ &+ Q_5 (\mathbb{E} [(X_2 \otimes X_3) (X_2 \otimes X_3)^T] \otimes st^{-1}\xi_2^v) Q_4^T + Q_6 (\mathbb{E} [(X_7 \otimes X_3) X_3^T] \otimes st^{-1}\xi_3^v) Q_4^T \\ &+ Q_6 (\mathbb{E} [(X_7 \otimes X_3) (X_2 \otimes X_3)^T] \otimes st^{-1}\xi_2^v) Q_3^T, \end{aligned} \quad (\text{B.16})$$

$$\begin{aligned}
\Psi_{6,6}^{\mathcal{V}}(k) &= \mathbb{E}(X_7 X_7^T) \otimes st^{-1} \xi_4^v - \mathbb{E}(X_7 X_7^T) \otimes (\xi_2^v \xi_2^{vT}) + (\mathbb{E}[X_7 (X_2 \otimes X_3^T)]) \otimes st^{-1} \xi_3^v Q_5^T \\
&\quad + \left(\mathbb{E}[X_7 (X_7 \otimes X_3)^T] \otimes st^{-1} \xi_3^v \right) Q_6^T + Q_5 \left(\mathbb{E}[(X_2 \otimes X_3) X_7^T] \otimes st^{-1} \xi_3^v \right) \\
&\quad + Q_5 \left(\mathbb{E}[(X_2 \otimes X_3) (X_2 \otimes X_3)^T] \otimes st^{-1} \xi_2^v \right) Q_5^T + Q_6 \left(\mathbb{E}[(X_7 \otimes X_3) X_7^T] \otimes st^{-1} \xi_3^v \right) \\
&\quad + Q_5 \left(\mathbb{E}[(X_2 \otimes X_3) (X_7 \otimes X_3)^T] \otimes st^{-1} \xi_2^v \right) Q_6^T \\
&\quad + Q_6 \left(\mathbb{E}[(X_7 \otimes X_3) (X_2 \otimes X_3)^T] \otimes st^{-1} \xi_2^v \right) Q_5^T \\
&\quad + Q_6 \left(\mathbb{E}[(X_7 \otimes X_3) (X_7 \otimes X_3)^T] \otimes st^{-1} \xi_2^v \right) Q_6^T, \tag{B.17}
\end{aligned}$$

$$\Psi_{1,1}^{\mathcal{W}}(k) = st^{-1} \xi_2^w, \tag{B.18}$$

$$\Psi_{2,1}^{\mathcal{W}}(k) = st^{-1} \xi_3^w + D_1 (Z_1 \otimes st^{-1} \xi_2^w) + D_2 (Z_3 \otimes st^{-1} \xi_2^w), \tag{B.19}$$

$$\begin{aligned}
\Psi_{2,2}^{\mathcal{W}}(k) &= st^{-1} \xi_4^w - \xi_2^w \xi_2^{wT} + (Z_1^T \otimes st^{-1} \xi_3^w) D_1^T + (Z_3^T \otimes st^{-1} \xi_3^w) D_2^T + D_1 (Z_1 \otimes st^{-1} \xi_3^w) \\
&\quad + D_2 (Z_3 \otimes st^{-1} \xi_3^w) + D_1 (\mathbb{E}[X_1 X_1^T] \otimes st^{-1} \xi_2^w) D_1^T + D_1 (\mathbb{E}[X_1 X_3^T] \otimes st^{-1} \xi_2^w) D_2^T \\
&\quad + D_2 (\mathbb{E}[X_3 X_1^T] \otimes st^{-1} \xi_2^w) D_1^T + D_2 (\mathbb{E}[X_3 X_3^T] \otimes st^{-1} \xi_2^w) D_2^T, \tag{B.20}
\end{aligned}$$

so that the covariance matrices assume the following blocks-form:

$$\Psi^{\mathcal{V}}(k) = \begin{pmatrix} \Psi_{1,1}^{\mathcal{V}}(k) & 0 & \Psi_{3,1}^{\mathcal{V}T}(k) & \Psi_{4,1}^{\mathcal{V}T}(k) & \Psi_{5,1}^{\mathcal{V}T}(k) & \Psi_{6,1}^{\mathcal{V}T}(k) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Psi_{3,1}^{\mathcal{V}}(k) & 0 & \Psi_{3,3}^{\mathcal{V}}(k) & \Psi_{4,3}^{\mathcal{V}T}(k) & \Psi_{5,3}^{\mathcal{V}T}(k) & \Psi_{6,3}^{\mathcal{V}T}(k) & 0 \\ \Psi_{4,1}^{\mathcal{V}}(k) & 0 & \Psi_{4,3}^{\mathcal{V}}(k) & \Psi_{4,4}^{\mathcal{V}}(k) & \Psi_{5,4}^{\mathcal{V}T}(k) & \Psi_{6,4}^{\mathcal{V}T}(k) & 0 \\ \Psi_{5,1}^{\mathcal{V}}(k) & 0 & \Psi_{5,3}^{\mathcal{V}}(k) & \Psi_{5,4}^{\mathcal{V}}(k) & \Psi_{5,5}^{\mathcal{V}}(k) & \Psi_{6,5}^{\mathcal{V}T}(k) & 0 \\ \Psi_{6,1}^{\mathcal{V}}(k) & 0 & \Psi_{6,3}^{\mathcal{V}}(k) & \Psi_{6,4}^{\mathcal{V}}(k) & \Psi_{6,5}^{\mathcal{V}}(k) & \Psi_{6,6}^{\mathcal{V}}(k) & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{B.21}$$

$$\Psi^{\mathcal{W}}(k) = \begin{pmatrix} \Psi_{1,1}^{\mathcal{W}}(k) & \Psi_{2,1}^{\mathcal{W}T}(k) \\ \Psi_{2,1}^{\mathcal{W}}(k) & \Psi_{2,2}^{\mathcal{W}}(k) \end{pmatrix}, \tag{B.22}$$

As already noted, the covariances depend on the mean values $Z_i(k) = \mathbb{E}[X_i(k)]$, $i = 1, \dots, 7$ and on the expected value of the Kronecker products of the state variables, that can be easily computed by taking into account their definition and the independence of $\vartheta(k)$ and $x(k)$, $\forall k$. In fact, by using the stack property (A.7f), it can be readily proved that every expected value appearing in the expressions (B.21)-(B.22) assume the form $\mathbb{E}[\vartheta^{[r]}(k)] \otimes \mathbb{E}[x^{[s]}(k)]$, with $r \leq 6$ and $s \leq 2$.

For instance, let us consider the maximum power term, appearing in (B.17):

$$\begin{aligned}
\mathbb{E}[(X_7 \otimes X_3) (X_7 \otimes X_3)^T] &= st^{-1} \left(\mathbb{E}[(X_7 \otimes X_3)^{[2]}] \right) \\
&= \left(I_{m^2} \otimes \mathcal{C}_{m^3n, mn}^T \right) st^{-1} \left(\mathbb{E}[(X_7^{[2]} \otimes X_3^{[2]})] \right) = \left(I_{m^2} \otimes \mathcal{C}_{m^3n, mn}^T \right) st^{-1} \left(\mathbb{E}[(\vartheta^{[6]} \otimes x^{[2]})] \right) \\
&= \left(I_{m^2} \otimes \mathcal{C}_{m^3n, mn}^T \right) st^{-1} \left(\mathbb{E}[\vartheta^{[6]}] \otimes \mathbb{E}[x^{[2]}] \right) = \left(I_{m^2} \otimes \mathcal{C}_{m^3n, mn}^T \right) st^{-1} (\xi_6^{\vartheta_0} \otimes Z_4). \tag{B.23}
\end{aligned}$$

being all the statistics of $\vartheta^{[i]}(k)$ stationary, $\forall i$. The proof is so completed.

C. Mutual covariance between extended state and output noises

This Appendix shows the uncorrelation between the extended state and output noises appearing in (2.8) and (2.15). As usual, the mutual covariance is computed block by block, taking into account that the state, the state noise $v(k)$ and the measurement noise $w(k)$ are mutually independent at the same instant.

The proof that $\mathbb{E} [\mathcal{V}_i(k)\mathcal{W}_1^T(k)] = 0 \forall k$, $i = 1, 3, 4, 5, 6$ is immediate: the same holds for $\mathbb{E} [\mathcal{V}_i(k)\mathcal{W}_2^T(k)] = 0 \forall k$, $i = 1, 3, 4$. For the remaining terms, here follows the computation of $\mathbb{E} [\mathcal{V}_4(k)\mathcal{W}_2^T(k)]$:

$$\begin{aligned}
\mathbb{E} [\mathcal{V}_4(k)\mathcal{W}_2^T(k)] &= \mathbb{E} [(v^{[2]} - \xi_2^v) ((X_1^T \otimes w^T) D_1^T + (X_3^T \otimes w^T) D_2^T)] \\
&+ \mathbb{E} [(v^{[2]} - \xi_2^v) (w^{[2]} - \xi_2^w)^T] + \mathbb{E} [(Q_1 (X_1 \otimes v) + Q_2 (X_3 \otimes w)) (w^{[2]} - \xi_2^w)^T] \\
&+ Q_1 \mathbb{E} [(X_1 \otimes v) (X_1^T \otimes w^T)] D_1^T + Q_1 \mathbb{E} [(X_1 \otimes v) (X_3^T \otimes w^T)] D_2^T \\
&+ Q_2 \mathbb{E} [(X_3 \otimes v) (X_1^T \otimes w^T)] D_1^T + Q_2 \mathbb{E} [(X_3 \otimes v) (X_3^T \otimes w^T)] D_2^T \\
&= (\mathbb{E} [X_1^T] \otimes \mathbb{E} [(v^{[2]} - \xi_2^v) w^T]) D_1^T + (\mathbb{E} [X_3^T] \otimes \mathbb{E} [(v^{[2]} - \xi_2^v) w^T]) D_2^T \\
&+ \mathbb{E} [(v^{[2]} - \xi_2^v)] \mathbb{E} [(w^{[2]} - \xi_2^w)^T] + (Q_1 \mathbb{E} [X_1] + Q_2 \mathbb{E} [X_3]) \otimes \mathbb{E} [v (w^{[2]} - \xi_2^w)^T] \\
&+ Q_1 (\mathbb{E} [X_1 X_1^T] \otimes \mathbb{E} [vw^T]) D_1^T + Q_1 (\mathbb{E} [X_1 X_3^T] \otimes \mathbb{E} [vw^T]) D_2^T \\
&+ Q_2 (\mathbb{E} [X_3 X_1^T] \otimes \mathbb{E} [vw^T]) D_1^T + Q_2 (\mathbb{E} [X_3 X_3^T] \otimes \mathbb{E} [vw^T]) D_2^T = 0, \quad (\text{C.1})
\end{aligned}$$

since each addend of the sum is zero thanks to the independence of $v(k)$ and $w(k)$. The result is that the state and output noises are uncorrelated at every instant. For the sake of brevity, the computations relative to $\mathbb{E} [\mathcal{V}_i(k)\mathcal{W}_2^T(k)]$, $i = 5, 6$ are not reported since they are perfectly analogous to the previous.

D. About the linearization of the nonlinear map $\varphi(X(k))$

The nonlinear map

$$\varphi(X(k)) = \begin{bmatrix} 0 \\ 0 \\ \Phi_3(X_2(k) \otimes X_3(k)) \\ 0 \\ \Phi_5(X_2(k) \otimes X_6(k)) \\ \Phi_{6,a}(X_2(k) \otimes X_6(k)) + \Phi_{6,b}(X_7(k) \otimes X_6(k)) \end{bmatrix}, \quad (\text{D.1})$$

reported for convenience, has to be developed up to the first order at each step k of the filtering algorithm around the current estimate of the state, $\hat{X}(k)$, in order to derive the equivalent dynamical matrix \tilde{A} for the computation of the error prediction covariance matrix $P_p(k)$.

The nonlinear terms appearing in (D.1) are of the form $X_i(k) \otimes X_j(k)$, $X_i \in \mathbb{R}^{n_i}$, $X_j \in \mathbb{R}^{n_j}$, so that the first order Taylor power expansion around an arbitrary point $(\tilde{X}_i(k), \tilde{X}_j(k))$ is given by

$$\begin{aligned}
X_i(k) \otimes X_j(k) &\cong \tilde{X}_i(k) \otimes \tilde{X}_j(k) + (I_{n_i} \otimes \tilde{X}_j(k)) (X_i(k) - \tilde{X}_i(k)) \\
&+ (\tilde{X}_i(k) \otimes I_{n_j}) (X_j(k) - \tilde{X}_j(k)). \quad (\text{D.2})
\end{aligned}$$

By applying Eq. (D.2) to (D.1) the matrix \tilde{A} can be easily derived. It is given by $\tilde{A}(k) = \mathcal{A} + \bar{A}(k)$, where $\bar{A}(k)$ is:

$$\bar{A}(k) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Phi_3(I_m \otimes \tilde{X}_3) & \Phi_3(\tilde{X}_2 \otimes I_{nm}) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Phi_5(I_m \otimes \tilde{X}_6) & 0 & 0 & 0 & \Phi_5(\tilde{X}_2 \otimes I_{n^2m^2}) & 0 \\ 0 & \Phi_{6,a}(I_m \otimes \tilde{X}_6) & 0 & 0 & 0 & \Phi_{6,a}(\tilde{X}_2 \otimes I_{n^2m^2}) & \Phi_{6,b}(I_{m^2} \otimes \tilde{X}_6) \\ 0 & 0 & 0 & 0 & 0 & +\Phi_{6,b}(\tilde{X}_7 \otimes I_{n^2m^2}) & 0 \end{pmatrix}. \quad (\text{D.3})$$

Remark D.1. Note that also the deterministic input $\mathcal{U}(k)$ changes consequently to the expansion of φ , but the new input $\tilde{\mathcal{U}}(k)$ is **not used** in the filtering algorithm, since the computation of the state prediction does not require any approximations.

References

- [1] R.E. Kopp and R.J. Oxford, "Linear regression applied to system identification for adaptive control systems," *AIAA J.*, pp. 2300–2306, Oct. 1963.
- [2] H. Cox, "On the estimation of state variables and parameters for noisy dynamic systems," *IEEE Trans. Automat. Contr.*, Vol. AC-9, pp. 5–12, Jan. 1964.
- [3] J.B. Farison, R.E. Graham, and R.C. Shelton, "Identification and control of linear discrete systems," *IEEE Trans. Automat. Contr.*, Vol. AC-12, pp. 438–442, Au. 1967.
- [4] A.C. Bartlett, A. Tesi, A. Vicino, "Frequency response of uncertain systems with interval plants," *IEEE Trans. Automat. Contr.*, Vol. 38, No. 6, pp. 929–933, June 1993.
- [5] E. De Santis, A. Vicino, "Diagonal dominance for robust stability of MIMO interval systems," *IEEE Trans. Automat. Contr.*, Vol. 41, No. 6, pp. 871–875, June 1996.
- [6] Hu Sanqing, Jun Wang, "On stabilization of a new class of linear time-invariant interval systems via constant state feedback control," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 11, pp. 2106–2111, Nov. 2000.
- [7] F. Carravetta, A. Germani, 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.
- [8] F. Carravetta, A. Germani, M. Raimondi, "Polynomial filtering for linear discrete-time non-Gaussian systems," *SIAM Journal on Control and Optimization*, Vol. 34, No. 5, pp. 1666–1690, 1996.
- [9] L. Ljung, "Linear time-invariant models of non-linear time-varying Systems," *European Journal of Control*, Vol. 7, issue 2–3, 2001.
- [10] A. V. Balakrishnan, *Kalman Filtering Theory*, New York: Optimization Software, 1984.