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

A. Germani, C. Manes, P. Palumbo

POLYNOMIAL FILTERING AND IDENTIFICATION FOR DISCRETE-TIME NONLINEAR UNCERTAIN STOCHASTIC SYSTEMS

R. 655 Dicembre 2006

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

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

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

ISSN: 1128–3378

Collana dei Rapporti
dell'Istituto di Analisi dei Sistemi ed Informatica, CNR

viale Manzoni 30, 00185 ROMA, Italy

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

Abstract

This paper deals with the problem of system identification and state estimation for nonlinear uncertain stochastic system, in the discrete-time framework. A polynomial filtering algorithm is here proposed and implemented by opportunely extending the state space with the inclusion of the unknown vector of parameters, so that the filtering and identification problems are simultaneously solved. Such an algorithm is achieved by applying the optimal polynomial filter of a chosen degree μ to the Carleman approximation of the same degree of the extended nonlinear system.

Key words: Polynomial filtering, Extended Kalman Filter, Carleman approximation, System identification, Uncertain systems.

1. Introduction

In this work it is investigated the problem of the simultaneous filtering and parameters identification for the following nonlinear uncertain stochastic system:

$$\begin{aligned} x(k+1) &= f\big(\theta, x(k)\big) + v(k), \quad x(0) = x_0, \quad k \geq 0 \\ y(k) &= h\big(\theta, x(k)\big) + w(k) \end{aligned} \tag{1.1}$$

where $x(k) \in I\!R^n$ is the state of the system, $y(k) \in I\!R^q$ is the measured output, $\theta \in I\!R^m$ is the unknown vector of parameters, $v(k)$ and $w(k)$ are sequences of zero-mean, auto and mutually independent random vectors, not necessarily Gaussian, x_0 is a random variable, independent of both the sequences $v(k)$ and $w(k)$. The most popular real time algorithm for simultaneous state and parameter estimation for this kind of systems is the Extended Kalman Filter (EKF) applied to the extended system, whose state is made of the original state and the parameter vector [8, 3]. The diffusion of the EKF is due to its simplicity and to the fact that in many applications it provides good estimates. The EKF is based on the linear approximation of a nonlinear 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 remains 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 parameter identification for system (1.1), in the case in which each component θ_i of the unknown parameter vector $\theta \in I\!R^m$ satisfies an internal constraint

$$\theta_i \in [\theta_{i,\min}, \theta_{i,\mathrm{Max}}], \quad i = 1, \ldots, m. \tag{1.2}$$

The state and output transition maps $f : I\!R^m \times I\!R^n \mapsto I\!R^n$ and $h : I\!R^m \times I\!R^n \mapsto I\!R^q$ are nonlinear maps, smooth with respect to both the parameters and the state vectors. Like the EKF-based algorithm previously mentioned, also the one here proposed considers the unknown parameters as further components of an extended state, including the original state components: $X^T(k) = \begin{bmatrix} x^T(k) & \theta^T(k) \end{bmatrix}^T$, with $\theta(k+1) = \theta(k)$. However, here the Polynomial Extended Kalman Filter (PEKF) is adopted, instead of the standard EKF. Such an algorithm [5] belongs to the group of nonlinear filtering methods where the nonlinear system is approximated by a the stochastic system for which known filtering procedures are available (the EKF, for instance, and all the existing modified versions [4, 10, 6]; an effective modification of the EKF is the Unscented Kalman Filter (UKF) [7], that uses the so-called *unscented transform* for the state and output prediction steps in the EKF equations.) The PEKF is obtained by the application of the optimal polynomial filter of [1, 2] to the Carleman approximation of a nonlinear system (see [9, 11]), whereas the standard EKF applies the classical Kalman filter to the linear approximation of nonlinear systems. The Carleman approximation of order μ of a nonlinear system is achieved by suitably defining an extended state made of the Kronecker powers of the original state up to a given order μ. The analogous definition of an extended output is also required for the construction of a polynomial filter. In the stochastic discrete-time framework the Carleman approximated system consists of a bilinear system (linear drift and multiplicative noise) with respect to the extended state. The extended output turns out to be a linear

function of the extended state, corrupted by multiplicative noise. Once the approximation is obtained, the recursive equations of the optimal polynomial filter of order μ are available and can be applied with no further approximations (see [1, 2]). It is interesting to note that the implementation of the PEKF of a given degree μ does not require the complete knowledge of the noises distributions: only the moments up to order 2μ are needed. When $\mu = 1$ the PEKF reduces to the classical EKF. As in the case of the classical EKF, the Polynomial Extended Kalman Filter (PEKF) is a time-varying recursive algorithm whose performances depend on the specific application. A better behavior with respect to the classical EKF is expected for two reasons: i) a higher degree of approximation of the nonlinear system is adopted; ii) the optimal polynomial estimate is implemented for the approximate system, instead of the linear Kalman estimate of the EKF.

2. Carleman approximation of stochastic systems

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, considering the additional state equation $\theta(k+1) = \theta(k)$, system (1.1) becomes:

$$\begin{aligned} X(k+1) &= f_e\big(X(k)\big) + v_e(k), \qquad X(0) = X_0 = \begin{pmatrix} x_0^T & \theta^T \end{pmatrix}^T, \quad k \geq 0 \\ y(k) &= h_e\big(X(k)\big) + w(k), \end{aligned} \tag{2.1}$$

where $f_e : I\!R^{n_e} \mapsto I\!R^{n_e}$, $h_e : I\!R^{n_e} \mapsto I\!R^q$, with $n_e = n + m$, are:

$$f_e\big(X\big) = \begin{bmatrix} f\big(X_2, X_1\big) \\ X_2 \end{bmatrix}, \qquad h_e\big(X\big) = h\big(X_2, X_1\big), \qquad v_e(k) = \begin{bmatrix} v(k) \\ 0 \end{bmatrix} \tag{2.2}$$

according to the natural decomposition of the extended vector

$$X(k) = \begin{bmatrix} X_1(k) \\ X_2(k) \end{bmatrix}, \qquad X_1(k) \in I\!R^n, \qquad X_2(k) \in I\!R^m. \tag{2.3}$$

The Carleman approximation, here applied to system (2.1) in order to implement the filtering algorithm, is based on the following steps. Choose an integer μ and consider the sequences $X^{[i]}(k)$ and $y^{[i]}(k)$ of the Kronecker powers of the extended state and output of system (2.1) for $i = 1, \ldots, \mu$ (here superscripts in square brackets denote the Kronecker powers of vectors and matrices; for a quick survey on the Kronecker algebra see [2]). The update equations for these sequences are

$$X^{[i]}(k+1) = \Big(f_e\big(X(k)\big) + v_e(k)\Big)^{[i]}, \qquad y^{[i]}(k) = \Big(h_e\big(X(k)\big) + w(k)\Big)^{[i]}. \tag{2.4}$$

Under standard analyticity hypotheses the nonlinear functions $(f_e + v_e)^{[i]}$ and $(h_e + w)^{[i]}$ can be approximated in a suitable neighborhood of a given point $\widetilde{X}$ using Taylor polynomials of degree μ:

$$\Big(f_e\big(X(k)\big) + v_e(k)\Big)^{[i]} \approx \sum_{j=0}^{\mu} F_{i,j}(\widetilde{X})\big(X(k) - \widetilde{X}\big)^{[j]} + \sum_{j=0}^{\mu} \varphi_{i,j}\big(\widetilde{X}, v_e(k)\big)\big(X(k) - \widetilde{X}\big)^{[j]}, \tag{2.5}$$

$$\Big(h_e\big(X(k)\big) + w(k)\Big)^{[i]} \approx \sum_{j=0}^{\mu} H_{i,j}(\widetilde{X})\big(X(k) - \widetilde{X}\big)^{[j]} + \sum_{j=0}^{\mu} \psi_{i,j}\big(\widetilde{X}, w(k)\big)\big(X(k) - \widetilde{X}\big)^{[j]}, \tag{2.6}$$

where $\varphi_{i,j}\big(\widetilde{X}, v_e(k)\big)$ and $\psi_{i,j}\big(\widetilde{X}, w(k)\big)$ are suitably defined polynomials of $v_e(k)$ and $w(k)$ (see [5]), and

$$F_{i,j}(X) = \frac{1}{j!}\left(\nabla_X^{[j]} \otimes f_e^{[i]}\right), \quad H_{i,j}(X) = \frac{1}{j!}\left(\nabla_X^{[j]} \otimes h_e^{[i]}\right); \tag{2.7}$$

the operator $\nabla_x^{[j]}\otimes$ applied to a function $\alpha = \alpha(x) : \mathbb{R}^n \mapsto \mathbb{R}^i$ is defined as

$$\nabla_x^{[0]} \otimes \alpha = \alpha, \qquad \nabla_x^{[j+1]} \otimes \alpha = \nabla_x \otimes \nabla_x^{[j]} \otimes \alpha, \qquad j \geq 1, \tag{2.8}$$

with $\nabla_x = [\partial/\partial x_1 \ \cdots \ \partial/\partial x_n]$. Note that $\nabla_x \otimes \alpha$ is the standard Jacobian of the vector function α.

Remark 2.1. From a computational point of view, the Carleman coefficients $F_{1,s}$ and $H_{1,s}$ can be achieved directly from the original maps:

$$F_{1,s}(X) = \frac{1}{s!}\big(\nabla_X^{[s]} \otimes f_e(X)\big) = \frac{1}{s!}\left(\nabla_X^{[s]} \otimes \begin{bmatrix} f\big(X_2, X_1\big) \\ X_2 \end{bmatrix}\right) = \frac{1}{s!}\begin{bmatrix} \nabla_X^{[s]} \otimes f(\theta, x) \\ \nabla_X^{[s]} \otimes I_d(\theta) \end{bmatrix}_{(\theta,x)=(X_2,X_1)}. \tag{2.9}$$

$$H_{1,s}(X) = \frac{1}{s!}\big(\nabla_X^{[s]} \otimes h_e(X)\big) = \frac{1}{s!}\left(\nabla_X^{[s]} \otimes h(X_2, X_1)\right) = \frac{1}{s!}\left(\nabla_X^{[s]} \otimes h(\theta, x)\right)_{(\theta,x)=(X_2,X_1)}. \tag{2.10}$$

with

$$\nabla_X^{[s]} \otimes I(\theta) = \begin{bmatrix} \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial \theta} \end{bmatrix}^{[s]} \otimes I_d(\theta) = \begin{cases} \theta, & \text{for } \ s = 0, \\ \begin{bmatrix} O_{m\times n} & I_m \end{bmatrix}, & \text{for } \ s = 1, \\ O_{m\times n_e^s}, & \text{for } \ s > 1. \end{cases} \tag{2.11}$$

As it will be clearer from the Appendix, all the Carleman coefficients related to the powers of f_e and h_e can be achieved from $F_{1,s}$ and $H_{1,s}$, and so from (2.9) and (2.10). ∎

The expansion of the powers of the binomials in the summations in eq.s (2.5) and (2.6) allows to write these as polynomials of $X(k)$ of degree μ (see [5]). The substitution of the j-th power of $X(k)$ in the summations with a vector $X_j^\mu(k)$ of the same dimension (recall that $X^{[j]}(k) \in \mathbb{R}^{n_e^j}$), and of the i-th power of $y(k)$ with a vector $Y_i^\mu(k) \in \mathbb{R}^{q^i}$ in the output equations, yeld the recursive equations of the Carleman approximation of order μ around $\widetilde{X}$:

$$\begin{aligned} X_i^\mu(k+1) &= \sum_{j=1}^{\mu} A_{i,j}^\mu\big(k, \widetilde{X}\big) X_i^\mu(k) + u_m^\mu\big(k, \widetilde{X}\big) + v_i^\mu\big(k, \widetilde{X}\big), \\ Y_i^\mu(k) &= \sum_{j=1}^{\mu} C_{i,j}^\mu\big(k, \widetilde{X}\big) X_j^\mu(k) + \gamma_i^\mu\big(k, \widetilde{X}\big) + w_i^\mu\big(k, \widetilde{X}\big), \end{aligned} \qquad \begin{aligned} &i = 1, \ldots, \mu, \\ &X_i^\mu(0) = X_0^{[i]}. \end{aligned} \tag{2.12}$$

The 2μ equations (2.12) of the Carleman approximation of system (1.1) can be put in the following compact form

$$\begin{aligned} X^\mu(k+1) &= \mathcal{A}^\mu\big(k, \widetilde{X}\big) X^\mu(k) + \mathcal{U}^\mu\big(k, \widetilde{X}\big) + V^\mu\big(k, \widetilde{X}\big), \\ Y^\mu(k) &= \mathcal{C}^\mu\big(k, \widetilde{X}\big) X^\mu(k) + \Gamma^\mu\big(k, \widetilde{X}\big) + W^\mu\big(k, \widetilde{X}\big), \end{aligned} \tag{2.13}$$

where

$$X^{\mu}(k) = \begin{bmatrix} X_1^{\mu}(k) \\ \vdots \\ X_{\mu}^{\mu}(k) \end{bmatrix} \in I\!R^{n_\mu}, \quad Y^{\mu}(k) = \begin{bmatrix} Y_1^{\mu}(k) \\ \vdots \\ Y_{\mu}^{\mu}(k) \end{bmatrix} \in I\!R^{q_\mu}, \qquad \begin{aligned} n_\mu &= \sum_{j=1}^{\mu} n_e^j, \\ q_\mu &= \sum_{j=1}^{\mu} q^j, \end{aligned} \tag{2.14}$$

$$\mathcal{A}^{\mu} = \begin{bmatrix} A_{1,1}^{\mu} & \cdots & A_{1,\mu}^{\mu} \\ \vdots & \ddots & \vdots \\ A_{\mu,1}^{\mu} & \cdots & A_{\mu,\mu}^{\mu} \end{bmatrix}, \qquad \mathcal{U}^{\mu} = \begin{bmatrix} u_1^{\mu} \\ \vdots \\ u_{\mu}^{\mu} \end{bmatrix}, \qquad V^{\mu} = \begin{bmatrix} v_1^{\mu} \\ \vdots \\ v_{\mu}^{\mu} \end{bmatrix}, \tag{2.15}$$

$$\mathcal{C}^{\mu} = \begin{bmatrix} C_{1,1}^{\mu} & \cdots & C_{1,\mu}^{\mu} \\ \vdots & \ddots & \vdots \\ C_{\mu,1}^{\mu} & \cdots & C_{\mu,\mu}^{\mu} \end{bmatrix}, \qquad \Gamma^{\mu} = \begin{bmatrix} \gamma_1^{\mu} \\ \vdots \\ \gamma_{\mu}^{\mu} \end{bmatrix}, \qquad W^{\mu} = \begin{bmatrix} w_1^{\mu} \\ \vdots \\ w_{\mu}^{\mu} \end{bmatrix}, \tag{2.16}$$

From eq. (A.9), in Appendix, the noises $V^{\mu}\big(k, \widetilde{X}\big)$ and $W^{\mu}\big(k, \widetilde{X}\big)$ are bilinear functions of the extended state $X^{\mu}(k)$ and of zero-mean random vectors uncorrelated with $X^{\mu}(k)$ of the type $(v_e^{[h]}(k) - E\{v_e^{[h]}(k)\})$ and $(w^{[h]}(k) - E\{w^{[h]}(k)\})$ (note that these are white sequences). This fact allows to state that the Carleman approximation (2.13) has a bilinear structure with respect to an extended white noise sequence. Moreover, exploiting the same arguments used in [1, 2], it is not difficult, though tedious, to prove that $V^{\mu}\big(k, \widetilde{X}\big)$ and $W^{\mu}\big(k, \widetilde{X}\big)$ are *uncorrelated* sequences of zero mean *uncorrelated* random vectors, and that the extended state $X^{\mu}(k)$ is *uncorrelated* with $W^{\mu}\big(j, \widetilde{X}\big)$ $\forall j$ and with $V^{\mu}\big(j, \widetilde{X}\big)$ for $k \leq j$ (this result is a direct consequence of the fact that the noises $v(k)$ and $w(k)$ in the original system (1.1) are independent and white, and that the original state $x(k)$ is independent of $w(j)$ $\forall j$ and independent of $v(j)$ for $k \leq j$).

In order to ensure that all random vectors in (2.13) ($X^{\mu}(k)$, $Y^{\mu}(k)$, $V^{\mu}\big(k, \widetilde{X}\big)$ and $W^{\mu}\big(k, \widetilde{X}\big)$) have finite means and covariances, it is necessary to assume that the noises and the initial extended state X_0 have finite and available moments up to order 2μ:

$$I\!E\big\{X_0^{[i]}\big\} < \infty, \qquad \begin{aligned} I\!E\big\{v^{[i]}(k)\big\} &= \xi_i^v(k) < \infty, \\ I\!E\big\{w^{[i]}(k)\big\} &= \xi_i^w(k) < \infty, \end{aligned} \qquad i = 1, \ldots, 2\mu; \tag{2.17}$$

as a matter of fact, it comes that the moments of both the initial original state x_0 and the unknown parameters vector have to be finite and available up to order 2μ. Such moments are needed for the recursive computation of the covariances $\Psi^{V^{\mu}}\big(k, \widetilde{X}\big)$ and $\Psi^{W^{\mu}}\big(k, \widetilde{X}\big)$ of the extended noises $V^{\mu}\big(k, \widetilde{X}\big)$ and $W^{\mu}\big(k, \widetilde{X}\big)$. The mean and covariance of the extended state $X^{\mu}(k)$, also needed for the computation of $\Psi^{V^{\mu}}\big(k, \widetilde{X}\big)$ and $\Psi^{W^{\mu}}\big(k, \widetilde{X}\big)$, can be recursively computed using standard formulas for bilinear systems. Details on the computation of the system matrices and sequences involved in (2.12), and of the extended noise statistics are reported in Appendix.

3. The filtering algorithm

The previous section has described the μ-th order Carleman approximation of a stochastic nonlinear system. The result is a bilinear system driven by white noise, given by eq.'s (2.13). For the filter construction it is assumed that the output of the original system (1.1) is generated in fact by the approximate model (2.13), and thus in the filter equations $Y_i^\mu(k)$ will coincide with $y^{[i]}(k)$. For a system of the type (2.13) the optimal linear filter (linear w.r.t. the *extended* measurements Y^μ) provides the optimal μ-degree polynomial filter w.r.t. the *original* measurements, and can be constructed without any further approximation (see [1, 2]). Since the extended noises $V^\mu\big(k,\tilde{x}\big)$ and $W^\mu\big(k,\tilde{x}\big)$ in eq.'s (2.13) are uncorrelated sequences of uncorrelated zero-mean vectors, as discussed in the previous section, the optimal linear filter is implemented by the standard Kalman filter equations. According to the same philosophy of the standard EKF, the system matrices and the covariances needed in the Riccati equations are computed using, at each step, the equations of the Carleman approximation around the current state estimate and prediction. In particular, the state estimate is used instead of $\widetilde{X}$ for the computation of matrices $\mathcal{A}^\mu$, $\mathcal{U}^\mu$ and Ψ^{V^μ}, while the state prediction is used for the computation of matrices $\mathcal{C}^\mu$, Γ^μ and Ψ^{W^μ}, according to the formulas reported in the Appendix. Note that the estimate $\widehat{X}(k)$ and prediction $\widehat{X}(k+1|k)$ of $X(k)$ are computed by selecting from the estimate and prediction of the extended state, $\widehat{X}^\mu(k)$ and $\widehat{X}^\mu(k+1|k)$, respectively, the first n_e components:

$$\widehat{X}(k) = [I_{n_e} \;\; O_{n_e\times(n_\mu-n_e)}]\widehat{X}^\mu(k), \quad \widehat{X}(k+1|k) = [I_{n_e} \;\; O_{n_e\times(n_\mu-n_e)}]\widehat{X}^\mu(k+1|k); \tag{3.1}$$

the original state and parameter estimate, $\hat{x}(k)$ and $\hat{\theta}(k)$ respectively, are then simultaneously achieved by selecting from $\widehat{X}(k)$ the first n and last m components:

$$\hat{x}(k) = [I_n \;\; O_{n\times m}]\widehat{X}(k), \qquad \hat{\theta}(k) = [O_{m\times n} \;\; I_m]\widehat{X}(k). \tag{3.2}$$

The steps of the PEKF algorithm are summarized below:

The Polynomial Extended Kalman Filter (PEKF)

I) Computation of the initial conditions of the filter:

$$\begin{aligned}
&\widehat{X}^\mu(0|-1) = I\!E\big\{X^\mu(0)\big\}, && \textit{a priori} \text{ estimate of } X^\mu(0), \text{ from eq. (A.24)}\\
&P_P(0) = \mathrm{Cov}\big(X^\mu(0)\big), && \text{covariance of the } \textit{a priori} \text{ estimate from eq. (A.24)}\\
&k = -1, && \text{inizialization of the counter;}
\end{aligned}$$

II) computation of the matrices of the μ-th degree approximation of the extended output equation around $\widehat{X}(k+1|k) = [I_{n_e} \;\; O_{n_e\times(n_\mu-n_e)}]\widehat{X}^\mu(k+1|k)$:

$$\begin{aligned}
\bar{\mathcal{C}}^\mu(k+1) &= \mathcal{C}^\mu\big(k+1, \widehat{X}^\mu(k+1|k)\big), && \text{from eq.'s (2.16) and (A.5),}\\
\bar{\Gamma}^\mu(k+1) &= \Gamma^\mu\big(k+1, \widehat{X}^\mu(k+1|k)\big), && \text{from eq.'s (2.16) and (A.8),}\\
\bar{\Psi}^{W^\mu}(k+1) &= \Psi^{W^\mu}\big(k+1, \widehat{X}^\mu(k+1|k)\big); && \text{from eq. (A.13),}
\end{aligned} \tag{3.3}$$

III) computation of the prediction of the extended output:

$$\widehat{Y}^\mu(k+1|k) = \bar{\mathcal{C}}^\mu\big(k+1\big)\widehat{X}^\mu(k+1|k) + \bar{\Gamma}^\mu(k+1); \tag{3.4}$$

IV) computation of the Kalman gain:

$$K(k+1) = P_P(k+1)\overline{\mathcal{C}}^{\mu}(k+1)^T\Big(\overline{\mathcal{C}}^{\mu}(k+1)P_P(k+1)\overline{\mathcal{C}}^{\mu}(k+1)^T + \overline{\Psi}^{W^{\mu}}(k+1)\Big)^{\dagger}; \quad (3.5)$$

V) computation of the error covariance matrix:

$$P(k+1) = \Big(I_{n_\mu} - K(k+1)\overline{\mathcal{C}}^{\mu}(k+1)\Big)P_P(k+1); \quad (3.6)$$

VI) computation of the extended state estimates $\widehat{X}^{\mu}(k+1)$ and $\widehat{X}(k+1)$ and of the estimates $\hat{x}(k+1)$, $\hat{\theta}(k+1)$ of the original state and of the unknown parameter:

$$\begin{aligned} \widehat{X}^{\mu}(k+1) &= \widehat{X}^{\mu}(k+1|k) + K(k+1)\Big(Y^{\mu}(k+1) - \widehat{Y}^{\mu}(k+1|k)\Big), \\ \widehat{X}(k+1) &= [I_{n_e} \;\; O_{n_e\times(n_\mu-n_e)}]\widehat{X}^{\mu}(k+1), \\ \hat{x}(k) = [I_n \;\; O_{n\times m}]\widehat{X}(k), &\qquad \hat{\theta}(k) = [O_{m\times n} \;\; I_m]\widehat{X}(k); \end{aligned} \quad (3.7)$$

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

VIII) computation of the matrices of the μ-th degree approximation of the extended state equation around $\widehat{X}(k)$:

$$\begin{aligned} \overline{\mathcal{A}}^{\mu}(k) &= \mathcal{A}^{\mu}\big(k, \widehat{X}(k)\big), \quad \text{from eq.'s (2.15) and (A.4)}, \\ \overline{\mathcal{U}}^{\mu}(k) &= \mathcal{U}^{\mu}\big(k, \widehat{X}(k)\big), \quad \text{from eq.'s (2.15) and (A.8)}, \\ \overline{\Psi}^{V^{\mu}}(k) &= \Psi^{V^{\mu}}\big(k, \widehat{X}(k)\big); \quad \text{from eq. (A.12)}, \end{aligned} \quad (3.8)$$

IX) computation of the extended state prediction:

$$\widehat{X}^{\mu}(k+1|k) = \overline{\mathcal{A}}^{\mu}(k)\widehat{X}^{\mu}(k) + \overline{\mathcal{U}}^{\mu}(k); \quad (3.9)$$

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

$$P_P(k+1) = \overline{\mathcal{A}}^{\mu}(k)P(k)\overline{\mathcal{A}}^{\mu}(k)^T + \overline{\Psi}^{V^{\mu}}(k); \quad (3.10)$$

XI) GOTO STEP **II**.

Remark 3.1. For consistency with all the developments made in the paper, the PEKF algorithm has been here presented in a form that is not computationally optimized, in that the Kronecker powers contain redundant components (if $X \in \mathbb{R}^{n_e}$ then $X^{[i]} \in \mathbb{R}^{n_e^i}$, but only $\tilde{n}_i = \binom{n_e+i-1}{i}$ monomials are independent). Such redundancies can be avoided through the definition of *reduced Kronecker powers*, containing the independent components of ordinary Kronecker powers (see [1]). More in detail, denoting with $X^{(i)} \in \mathbb{R}^{n_i}$ the reduced i-th Kronecker power of X, it is always possible to define a selection matrix $T_i(n_e) \in \mathbb{R}^{\tilde{n}_i \times n^i}$ made of 0's and 1's, such that:

$$X^{(i)} = T_i(n_e)X^{[i]} \quad (3.11)$$

(note that the choice of matrix $T_i(n_e)$ is not univocal). Similarly, the ordinary Kronecker power $X^{[i]}$ is recovered from the reduced power $X^{(i)}$ through multiplication with a suitable matrix $\widetilde{T}_i(n_e) \in \mathbb{R}^{n^i \times \tilde{n}_i}$. Straightforward but tedious substitutions in the above PEKF algorithm provide a filter with a reduced computational burden, and this last should be considered for efficient implementations.

4. Simulation results

Some significative results are here reported in order to show the effectiveness of the proposed algorithm. Consider the following nonlinear system:

$$\begin{aligned} x_1(k+1) &= \alpha(\theta)x_1(k) + x_1(k)x_2(k) + 0.1 + 0.01v_1(k), \\ x_2(k+1) &= 1.5x_2(k) - x_1(k)x_2(k) + 0.1 + 0.01v_2(k), \\ y(k) &= x_2(k) + 0.04w(k), \end{aligned} \tag{4.1}$$

with the zero-mean white noises v_1, v_2, w independent and obeying the following discrete distributions:

$$\begin{aligned} &P\{v_1(k) = -1\} = 0.6, \quad P\{v_2(k) = -1\} = 0.8, \quad P\{w(k) = -7\} = 0.3, \\ &P\{v_1(k) = 0\} = 0.2, \quad P\{v_2(k) = 4\} = 0.2, \quad P\{w(k) = 3\} = 0.7. \\ &P\{v_1(k) = 3\} = 0.2, \end{aligned} \tag{4.2}$$

The initial state $x(0)$ is also a random variable, independent of both the state and output noises, with distribution:

$$\begin{aligned} &P\{x_1(0) = 0.4\} = 0.2, \quad P\{x_2(0) = 0.1\} = 0.2, \\ &P\{x_1(0) = 0.8\} = 0.8, \quad P\{x_2(0) = 0.4\} = 0.8. \end{aligned} \tag{4.3}$$

The coefficient α in the first state equation of (4.1) is unknown: it is known that it is close to its nominal value 0.8, so that a parameterization could be as follows:

$$\alpha(\theta) = 0.8\frac{\theta}{\sqrt{1+\theta^2}}, \qquad \theta \in [0, 10], \tag{+}$$

with θ assuming the uniform distribution in its interval constraint (i.e. the unknown coefficient α is internal with respect to the open interval $[0, 0.8)$). Such a parameterization is useful, in that it avoids the filter to overestimates α so reducing the possibility of divergences for the algorithm. A further consequence is that there is a great probability that α is close to the nominal value 0.8: more than 90% to belong to $[0.7, 0.8)$.

Simulations are reported comparing the standard first and second order EKF with the quadratic ($\mu = 2$) and cubic ($\mu = 3$) version of the proposed PEKF.

The sample error variances computed in a typical simulation over a 500 points horizon, are reported in table 4.1.

Table 4.1. Steady state error variances

	EKF	2nd ord. EKF	PEKF$_{\mu=2}$	PEKF$_{\mu=3}$
$\sigma^2_{x_1}$	$9.15\cdot 10^{-3}$	$1.00\cdot 10^{-2}$	$7.45\cdot 10^{-3}$	$4.75\cdot 10^{-3}$
$\sigma^2_{x_2}$	$1.75\cdot 10^{-3}$	$1.76\cdot 10^{-3}$	$1.30\cdot 10^{-3}$	$4.28\cdot 10^{-4}$

In this example the quadratic and cubic PEKF perform better than the EKF and 2nd order EKF, which have a very similar behavior. In particular, the quadratic PEKF achieves 18% and 25% reduction of the error variance of the two state components, respectively, w.r.t. standard EKF, while the cubic PEKF achieves 48% and 75% variance reduction. Figures 4.1 and 4.2 report the true states and their estimates using the 2nd order EKF, the quadratic PEKF and

the cubic PEKF (for the clarity of the representation, only the last 80 time steps is reported). The EKF estimates is not reported in the figures because they are extremely similar to those provided by the 2nd order EKF and can not be distinguished.

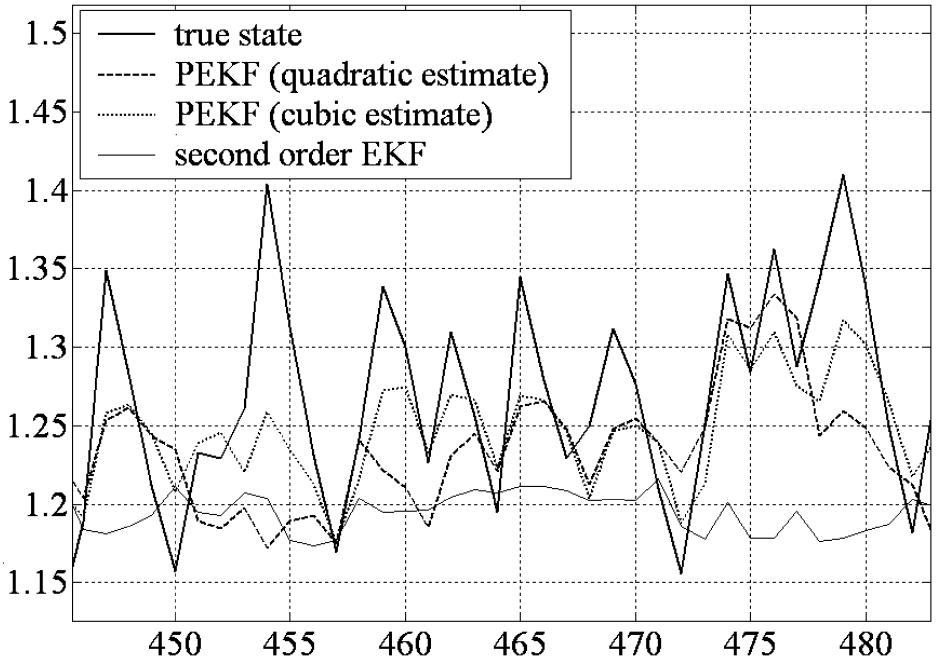

Fig. 4.1 – True and estimated state: the first component.

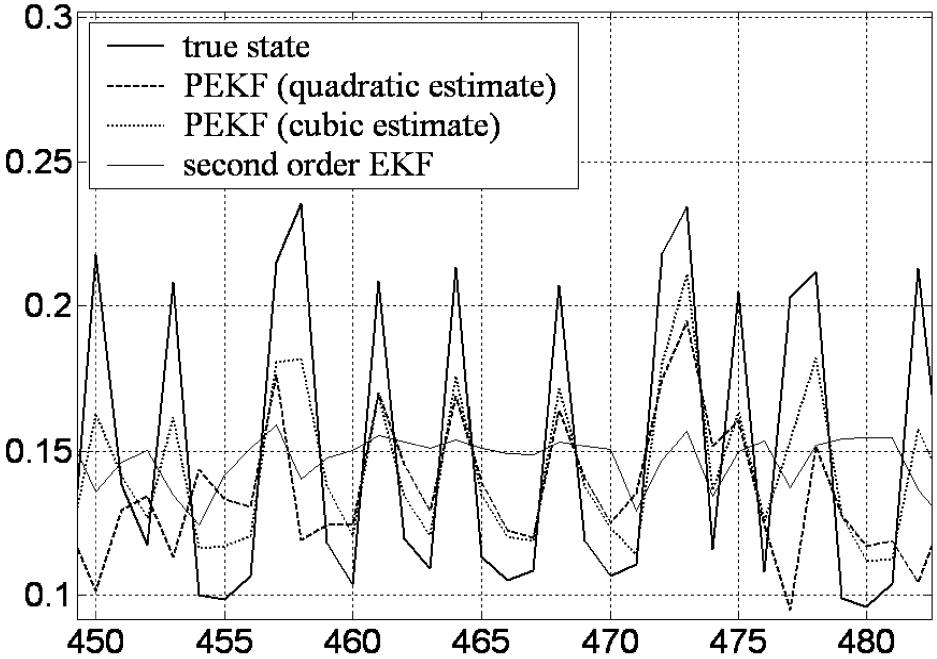

Fig. 4.2 – True and estimated state: the second component.

As far as what concerns the parameter identification, figure 4.3 shows the improvements from the 2nd order EKF up to the quadratic and cubic PEKF:

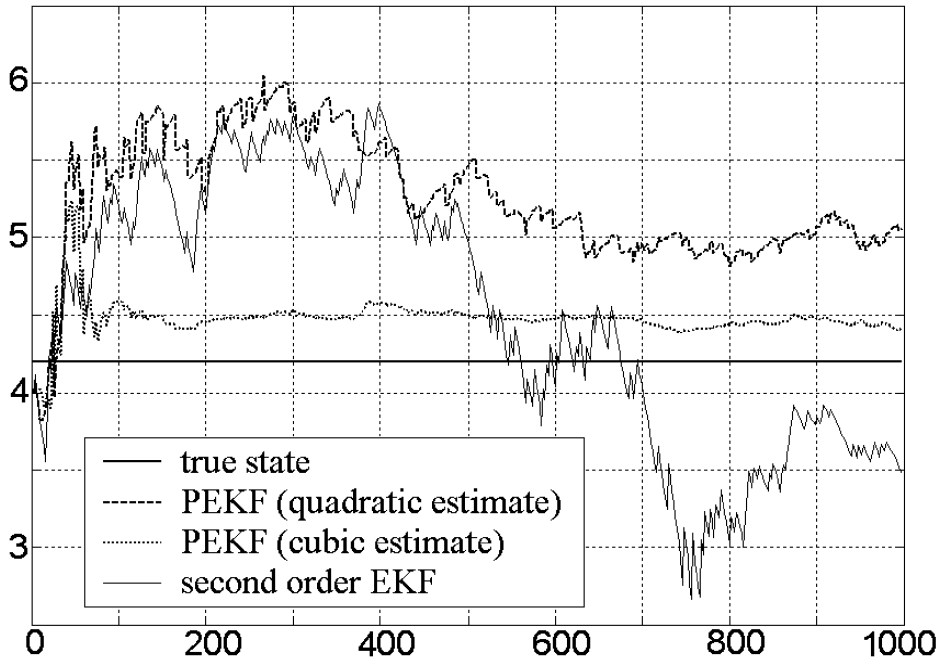

Fig. 4.3 – True and estimated parameter θ.

5. Conclusions

The problem of state estimation for a nonlinear system 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 nonlinear system is approximated using the Carleman bilinearization approach, taking into account all the powers of the series expansion up to a fixed degree μ; next, the minimum variance filter of the approximating system in the Hilbert space of all the μ^{th}-degree polynomial transformations of the measurements is computed. This step is based on a well known literature concerning suboptimal polynomial estimates for linear and bilinear state space representations [1, 2]. When $\mu = 1$, the proposed algorithm gives back the standard Extended Kalman Filter.

Appendix: the Kronecker algebra

This Appendix reports the expressions of all the terms needed for the PEKF implementation. The derivation of these equations exploits the rules of the Kronecker algebra (see [2] for a quick survey) and take advantage of a formalism that allows to expand Kronecker powers of sums of vectors. Consider a multiindex $t = \{t_0, t_1, \cdots\ t_\nu\} \in \left(\mathbb{Z}^+\right)^{\nu+1}$. Its modulus, denoted $|t|$, is defined as the sum of its entries, i.e. $|t| = t_0 + \cdots + t_\nu$. The i-th Kronecker power of a sum of $\nu+1$ vectors $z_i \in \mathbb{R}^p$, $i = 0, 1, \ldots, \nu$, can be expressed as

$$(z_0 + z_1 + ... + z_\nu)^{[i]} = \sum_{|t|=i} M_t^p \left(z_0^{[t_0]} \otimes z_1^{[t_1]} \otimes \cdots \otimes z_\nu^{[t_\nu]} \right), \tag{A.1}$$

with a suitable definition of the matricial coefficients $M_t^p \in \mathbb{R}^{p^i \times p^i}$ (see [2]). Note that for $t \in (Z^+)^2$ it is $M_{t_0,t_1}^1 = \binom{t_0+t_1}{t_0}$. The Kronecker product of n matrices A_h, $h = 1, \ldots n$, is denoted as

$$\bigotimes_{h=1}^{n} A_h = A_1 \otimes A_2 \otimes \cdots \otimes A_n. \tag{A.2}$$

With this definition, equation (A.1) can be put in the more compact form

$$\Big(\sum_{h=0}^{\nu} z_h \Big)^{[i]} = \sum_{|t|=i} M_t^p \bigotimes_{h=0}^{\nu} z_h^{[t_h]}. \tag{A.3}$$

Let us recall that the stack of a matrix $A \in \mathbb{R}^{r \times c}$ is the vector in $\mathbb{R}^{r \cdot c}$ that piles up all the columns of matrix A, and is denoted $\mathrm{st}(A)$. The inverse operation is denoted $\mathrm{st}_{r,c}^{-1}(\cdot)$, and transforms a vector of size $r \cdot c$ in a $r \times c$ matrix.

Lemma A.1. *The matrices $A_{ij}^\mu\big(k, \widetilde{X}\big)$ and $C_{m,i}^\mu\big(k, \widetilde{X}\big)$ of system (2.12) are as follows:*

$$A_{ij}^\mu\big(k, \widetilde{X}\big) = \sum_{r \in \mathcal{R}_{ij}^\mu} M_r^{n_e} \mathcal{F}_r(\widetilde{X}) \big(M_{\alpha(r)-j,j}^{n_e} \otimes \xi_{r_{\mu+1}}^v(k) \big) \big(I_{n_e^j} \otimes (-\widetilde{X})^{[\alpha(r)-j]} \big), \tag{A.4}$$

$$C_{ij}^\mu\big(k, \widetilde{X}\big) = \sum_{r \in \mathcal{R}_{ij}^\mu} M_r^q \overline{H}_r(\widetilde{X}) \big(M_{\alpha(r)-j,j}^{n_e} \otimes \xi_{r_{\mu+1}}^w(k) \big) \big(I_{n_e^j} \otimes (-\widetilde{X})^{[\alpha(r)-j]} \big), \tag{A.5}$$

with $r = \big\{r_0, \cdots, r_{\mu+1}\big\}$ a multi-index in $\left(\mathbb{Z}^+\right)^{\mu+2}$ and:

$$\alpha(r) = \sum_{s=1}^{\mu} s\, r_s, \qquad \mathcal{R}_{ij}^\mu = \big\{ r \in \left(\mathbb{Z}^+\right)^{\mu+2} : |r| = i,\ j \le \alpha(r) \le \mu \big\}; \tag{A.6}$$

the matrices $\mathcal{F}_r$, $\overline{H}_r$ in (A.4), (A.5) are defined as:

$$\mathcal{F}_r(\widetilde{X}) = \left(\bigotimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(\widetilde{X}) \right) \otimes \begin{bmatrix} I_n \\ O_{m \times n} \end{bmatrix}^{[r_{\mu+1}]}, \qquad \overline{H}_r(\widetilde{X}) = \left(\bigotimes_{s=0}^{\mu} H_{1,s}^{[r_s]}(\widetilde{X}) \right) \otimes I_{q^{r_{\mu+1}}}. \tag{A.7}$$

Moreover, the deterministic drifts u_i^μ, γ_i^μ are computed as

$$u_i^\mu\big(k, \widetilde{X}\big) = \sum_{r \in \mathcal{R}_{i0}^\mu} M_r^{n_e} \mathcal{F}_r(\widetilde{X}) \big(\widetilde{X}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^v(k) \big), \tag{A.8}$$

$$\gamma_i^{\mu}\big(k,\widetilde{X}\big) = \sum_{r\in\mathcal{R}_{i0}^{\mu}} M_r^q\,\overline{H}_r(\widetilde{X})\big(\widetilde{X}^{[\alpha(r)]}\otimes\xi_{r_{\mu+1}}^{w}(k)\big), \tag{A.9}$$

and the random sequences $\big\{v_i^{\mu}\big\}$, $\underset{\alpha(r)}{\big\{w_i^{\mu}\big\}}$ *are given by*

$$v_i^{\mu}\big(k,\widetilde{X}\big) = \sum_{r\in\mathcal{R}_{i0}^{\mu}}\sum_{s=0}^{} \Delta_{i,s}^{r}(\widetilde{X})\Big(X_s^{\mu}(k)\otimes\big(v^{[r_{\mu+1}]}(k)-\xi_{r_{\mu+1}}^{v}(k)\big)\Big), \tag{A.10}$$

$$w_i^{\mu}\big(k,\widetilde{X}\big) = \sum_{r\in\mathcal{R}_{i0}^{\mu}}\sum_{s=0}^{\alpha(r)} \Phi_{i,s}^{r}(\widetilde{X})\Big(X_s^{\mu}(k)\otimes\big(w^{[r_{\mu+1}]}(k)-\xi_{r_{\mu+1}}^{w}(k)\big)\Big), \tag{A.11}$$

with

$$\Delta_{i,s}^{r}(\widetilde{X}) = M_r^{n_e}\mathcal{F}_r(\widetilde{X})\Big(M_{\alpha(r)-s,s}^{n_e}\big(I_{n_e^s}\otimes(-\widetilde{X})^{[\alpha(r)-s]}\big)\otimes I_{n_e^{r_{\mu+1}}}\Big), \tag{A.12}$$

$$\Phi_{i,s}^{r}(\widetilde{X}) = M_r^q\,\overline{H}_r(\widetilde{X})\Big(M_{\alpha(r)-s,s}^{n_e}\big(I_{n_e^s}\otimes(-\widetilde{X})^{[\alpha(r)-s]}\big)\otimes I_{q^{r_{\mu+1}}}\Big). \tag{A.13}$$

Proof. The derivation of the Carleman approximation of a generic nonlinear stochastic system, endowed by the powers of the output equation, in the discrete-time framework, has been presented in [5]. Here it is reported the way to define such a derivation for the nonlinear function f_e and h_e defined in (2.2). Taking into account the extended noise v_e defined in (2.2), from [5] and according to (A.6), it comes that:

$$A_{ij}^{\mu}\big(k,\widetilde{X}\big)=\sum_{r\in\mathcal{R}_{ij}^{\mu}} M_r^{n_e}\,\overline{F}_r(\widetilde{X})\left(M_{\alpha(r)-j,j}^{n_e}\otimes\left(\begin{bmatrix}I_n\\ O_{m\times n}\end{bmatrix}^{[r_{\mu+1}]}\cdot\xi_{r_{\mu+1}}^{v}(k)\right)\right)\big(I_{n_e^j}\otimes(-\widetilde{X})^{[\alpha(r)-j]}\big), \tag{A.14}$$

with

$$\overline{F}_r(\widetilde{X}) = \left(\bigotimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(\widetilde{X})\right)\otimes I_{n_e^{r_{\mu+1}}}. \tag{A.15}$$

Taking into account that given four matrices suitably dimensioned, the following identity holds

$$\big(A\cdot B\big)\otimes\big(C\cdot D\big) = \big(A\otimes C\big)\cdot\big(B\otimes D\big), \tag{A.16}$$

then:

$$\begin{aligned}
&\overline{F}_r(\widetilde{X})\left(M_{\alpha(r)-j,j}^{n_e}\otimes\left(\begin{bmatrix}I_n\\ O_{m\times n}\end{bmatrix}^{[r_{\mu+1}]}\cdot\xi_{r_{\mu+1}}^{v}(k)\right)\right)\\
&\quad=\left(\left(\bigotimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(\widetilde{X})\right)\otimes I_{n_e^{r_{\mu+1}}}\right)\cdot\left(\left(I_{n_e^i}\cdot M_{\alpha(r)-j,j}^{n_e}\right)\otimes\left(\begin{bmatrix}I_n\\ O_{m\times n}\end{bmatrix}^{[r_{\mu+1}]}\cdot\xi_{r_{\mu+1}}^{v}(k)\right)\right)\\
&\quad=\left(\left(\bigotimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(\widetilde{X})\right)\otimes I_{n_e^{r_{\mu+1}}}\right)\cdot\left(I_{n_e^i}\otimes\begin{bmatrix}I_n\\ O_{m\times n}\end{bmatrix}^{[r_{\mu+1}]}\right)\cdot\Big(M_{\alpha(r)-j,j}^{n_e}\otimes\xi_{r_{\mu+1}}^{v}(k)\Big)\\
&\quad=\left(\left(\bigotimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(\widetilde{X})\right)\otimes\begin{bmatrix}I_n\\ O_{m\times n}\end{bmatrix}^{[r_{\mu+1}]}\right)\cdot\Big(M_{\alpha(r)-j,j}^{n_e}\otimes\xi_{r_{\mu+1}}^{v}(k)\Big),
\end{aligned} \tag{A.17}$$

so that eq. (A.4) comes with $\mathcal{F}_r(\widetilde{X})$ as in (A.7). Analogously, eq.'s (A.8), (A.10), with (A.12), readily come. The approximate output equation (i.e. its term involved in the eq.'s (A.5), (A.7), (A.9), (A.11), (A.13) are straightforward from [5]. ∎

Lemma A.2. *Consider Ψ^{V^μ} and Ψ^{W^μ}, the covariances of the random vectors V^μ and W^μ defined in eq.'s* (2.15) *and* (2.16)*, whose entries, by definition, are:*

$$\Psi_{ij}^{V^\mu}\big(k,\widetilde{X}\big)=I\!E\Big\{v_i^\mu\big(k,\widetilde{X}\big)v_j^\mu\big(k,\widetilde{X}\big)^T\Big\},\qquad \Psi_{ij}^{W^\mu}\big(k,\widetilde{X}\big)=I\!E\Big\{w_i^\mu\big(k,\widetilde{X}\big)w_j^\mu\big(k,\widetilde{X}\big)^T\Big\}. \quad (A.18)$$

These can be computed as follows:

$$\begin{aligned}\Psi_{ij}^{V^\mu}(k,\widetilde{X})=&\sum_{r\in\mathcal{R}_{j0}^\mu}\sum_{t\in\mathcal{R}_{i0}^\mu}\sum_{s=0}^{\alpha(r)}\sum_{l=0}^{\alpha(t)}\Delta_{i,s}^r(\widetilde{X})\\&\cdot\left(\big(\Psi_{s,l}^{X^\mu}\big(k,\widetilde{X}\big)\otimes \mathrm{st}^{-1}_{n_e^i,n_e^j}\Big(\big(\xi^v_{t_{\mu+1}+r_{\mu+1}}(k)-\xi^v_{t_{\mu+1}}(k)\otimes\xi^v_{r_{\mu+1}}(k)\big)\Big)\right)\Delta_{j,l}^t(\widetilde{X})^T,\end{aligned} \quad (A.19)$$

$$\begin{aligned}\Psi_{ij}^{W^\mu}(k,\widetilde{X})=&\sum_{r\in\mathcal{R}_{j0}^\mu}\sum_{t\in\mathcal{R}_{i0}^\mu}\sum_{s=0}^{\alpha(r)}\sum_{l=0}^{\alpha(t)}\Phi_{i,s}^r(\widetilde{X})\\&\cdot\left(\big(\Psi_{s,l}^{X^\mu}\big(k,\widetilde{X}\big)\otimes \mathrm{st}^{-1}_{q^i,q^j}\Big(\big(\xi^w_{t_{\mu+1}+r_{\mu+1}}(k)-\xi^w_{t_{\mu+1}}(k)\otimes\xi^w_{r_{\mu+1}}(k)\big)\Big)\right)\Phi_{j,l}^t(\widetilde{X})^T,\end{aligned} \quad (A.20)$$

with $\Psi_{ij}^{X^\mu}\big(k,\widetilde{X}\big)=I\!E\Big\{X_i^\mu(k)X_j^\mu(k)^T\Big\}$ the blocks of the matrix of second order moments of the extended state, $\Psi^{X^\mu}\big(k,\widetilde{X}\big)=I\!E\Big\{X^\mu(k)X^\mu(k)^T\Big\}$, computed by the recursive equation

$$\begin{aligned}\Psi^{X^\mu}(k+1,\widetilde{X})=&\mathcal{A}^\mu(k,\widetilde{X})\Psi^{X^\mu}(k,\widetilde{X})\mathcal{A}^\mu(k,\widetilde{X})^T+\mathcal{U}^\mu(k,\widetilde{X})\mathcal{U}^\mu(k,\widetilde{X})^T+\Psi^{V^\mu}(k,\widetilde{X})\\&+\mathcal{A}^\mu(k,\widetilde{X})Z^\mu(k)\mathcal{U}^\mu(k,\widetilde{X})^T+\mathcal{U}^\mu(k,\widetilde{X})Z^\mu(k)^T\mathcal{A}^\mu(k,\widetilde{X})^T,\end{aligned} \quad (A.21)$$

where $Z^\mu(k)=I\!E\big\{X^\mu(k)\big\}$ is the mean value of the extended state, computed as:

$$Z^\mu(k+1)=\mathcal{A}^\mu(k,\widetilde{X})Z^\mu(k)+\mathcal{U}^\mu(k,\widetilde{X}). \quad (A.22)$$

The initialization of (A.21) *and* (A.22) *are as follows*

$$\Psi_{ij}^{X^\mu}(0,\widetilde{X})=E\left\{X_0^{[i]}\big(X_0^{[j]}\big)^T\right\}=\mathrm{st}^{-1}_{n_e^i,n_e^j}\Big(I\!E\big\{X_0^{[i+j]}\big\}\Big),\qquad Z_i^\mu(0)=I\!E\big\{X_0^{[i]}\big\}. \quad (A.23)$$

with

$$I\!E\big\{X_0^{[i]}\big\}=\sum_{j=0}^{i}M_{j,i-j}^2\left(\begin{bmatrix}I_n\\O_{m\times n}\end{bmatrix}^{[j]}\otimes\begin{bmatrix}O_{n\times m}\\I_m\end{bmatrix}^{[i-j]}\right)(\zeta_j^0\otimes\zeta_{i-j}^\theta),\qquad \begin{matrix}\zeta_j^0=I\!E\big\{X_0^{[i]}\big\},\\ \zeta_j^\theta=I\!E\big\{\theta^{[i]}\big\}.\end{matrix} \quad (A.24)$$

References

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

[2] 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, pp. 1106–1126, 1997.

[3] H. Cox, "On the estimation of state variables and parameters for noisy dynamic systems," *IEEE Trans. Automat. Contr.*, Vol. 9, pp. 5-12, 1964.

[4] A. Gelb, *Applied Optimal Estimation*, MIT Press, MA: Cambridge, 1984.

[5] Germani, A., C. Manes and P. Palumbo, "Polynomial Extended Kalman Filtering for Discrete-Time Nonlinear Stochastic Systems", *Proc. of 42nd IEEE Conf. on Decision and Control*, Maui, Hawaii, pp. 886–891, 2003.

[6] Jazwinski, A. H. (1970). *Stochastic Processes and Filtering Theory*, Academic, New York.

[7] Julier, S. J. and J. K. Uhlmann (2004). "Unscented filtering and nonlinear estimation," *Proc. of IEEE*, **Vol. 92**, pp. 401–422.

[8] R.E. Kopp and R. J. Oxford, "Linear regression applied to system identification for adaptive control systems," *AIAA Journal*, Vol. 1, No. 10, pp. 2300-2306, 1963.

[9] K. Kowalski and W. H. Steeb, *Nonlinear Dynamical Systems and Carleman Linearization*, World Scientific, Singapore, 1991.

[10] Reif, K., S. Günther, E. Yaz and R. Unbehauen (1999). "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Trans. Autom. Contr.*, **Vol. 44**, pp. 714–728.

[11] Sastry, S. (1999). *Nonlinear systems: Analysis, Stability and Control*, Springer-Verlag, New-York.