A. Germani, C. Manes, P. Palumbo

POLYNOMIAL EXTENDED KALMAN FILTERING FOR DISCRETE-TIME NONLINEAR STOCHASTIC SYSTEMS

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`.

Abstract

This paper deals with the state estimation problem for a discrete-time nonlinear system driven by additive noise (not necessarily Gaussian). The solution here proposed is a filtering algorithm which is polynomial with respect to the measurements. The first step for the filter derivation is the embedding of the nonlinear system into an infinite-dimensional bilinear system (linear drift and multiplicative noise), following the Carleman approach. Then, the infinite dimensional system is approximated by neglecting all the powers of the state up to a chosen degree μ, and the minimum variance estimate among all the μth-degree polynomial transformations of the measurements is computed. The proposed filter can be considered a Polynomial Extended Kalman Filter (PEKF), because when $\mu = 1$ the classical EKF algorithm is recovered. Numerical simulations support the theoretical results and show the improvements of a quadratic filter with respect to the classical EKF.

Key words: Polynomial filtering, Extended Kalman Filter, Carleman approximation, stochastic systems.

1. Introduction

In the last decades an increasing interest has been devoted to the analysis and control of nonlinear systems, and the filtering problems related to them. It is well known that the minimum variance state estimate of a stochastic system requires to the knowledge of the probability density of the current state conditioned by all the measurements up to the current time. Unfortunately, the computation of the conditional probabilities is a difficult problem and the optimal filter has not, in general, a finite-dimensional representation [2, 5, 6]. Further difficulties occur when the noises driving the system are not Gaussian. However, a great deal of efforts have been made to approximate the infinite-dimensional equations achieving the conditional probabilities or to implement suboptimal filters [12].

According to its superior practical usefulness, the Extended Kalman Filter (EKF) (see, e.g., [1, 9, 15]) is one of the most widely used algorithm for the filtering of nonlinear systems in many frameworks, such as adaptive filtering [11], parameter estimation [10], robust control [17], state observation in zero-noise cases [3, 4], system identification [13], and many others. It is well known that, being based on the linear approximation of the system, the EKF performs well if the initial estimation error and the disturbing noises are small enough. In [16] conditions are given that ensure the boundedness of the state estimation error.

This work deals with the state estimation problem of a discrete-time nonlinear system with additive state and measurements noises (not necessarily Gaussian). The aim is to derive a polynomial filtering algorithm based on the μth order polynomial Carleman approximation of the system [14] (instead of the standard linear one of the EKF). The Carleman approximation allows to approximate a smooth nonlinear stochastic system with a bilinear one (linear drift, multiplicative noise). The filter is obtained by projecting the state of the approximated system onto the Hilbert space of all the μ degree polynomial transformations of the measurements, according to a well known result in the literature concerning suboptimal polynomial estimates of linear and bilinear systems in the discrete-time framework (see [7, 8]).

The paper is organized as follows: the next section deals with the system to be filtered and how to obtain a suitable approximation using the Carleman bilinearization theory; in section three the polynomial minimum variance filter of the finite-dimensional system approximation is derived; section four shows some numerical simulations in order to taste the goodness of the proposed algorithm.

2. The system to be filtered

The class of systems investigated in this paper is described by the following set of discrete-time equations:

$$x(k+1) = f\big(k, x(k)\big) + v(k), \qquad k \geq 0, \qquad x(0) = x_0, \tag{2.1}$$
$$y(k) = h\big(k, x(k)\big) + w(k),$$

where the state $x(k)$ is a stochastic variable in $I\!R^n$, $y(k)$ is the measured output in $I\!R^q$ and $f : I\!N \times I\!R^n \mapsto I\!R^n$, $h : I\!N \times I\!R^n \mapsto I\!R^q$ are time-varying nonlinear maps. Both the state and output noises, $\{v(k)\}$ and $\{w(k)\}$ respectively, are independent sequences of zero-mean independent random vectors, not necessarily Gaussian, with finite and available moments up to order $2\mu^{\text{th}}$, named

$$I\!E\big[v^{[i]}(k)\big] = \xi_i^v(k), \qquad I\!E\big[w^{[i]}(k)\big] = \xi_i^w(k), \qquad i = 1, \ldots, 2\mu, \tag{2.2}$$

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

$$M^{[0]} = 1, \qquad M^{[i]} = M \otimes M^{[i-1]}, \qquad i \geq 1, \tag{2.3}$$

with $\otimes$ the standard Kronecker product. The definition of the Kronecker power and some of its properties are reported in the Appendix (for a quick survey on the Kronecker algebra see [8] and references therein). Throughout the paper only the superscripts in square brackets have to be intended as Kronecker powers.

The initial state x_0 is a random vector with finite and available moments up to order $2\mu^{th}$:

$$I\!E\big[x_0^{[i]}\big] = \zeta_i^0, \qquad i = 1, \ldots, 2\mu. \tag{2.4}$$

Moreover, x_0 is assumed to be independent of both the noise sequences $\{v(k)\}$, $\{w(k)\}$.

Under standard analyticity hypotheses, the nonlinear maps f and h can be written by using the Taylor polynomial expansion around a given state $\tilde{x}$. According to the Kronecker formalism, the system equations in (2.1) can be written as:

$$x(k+1) = \sum_{i=0}^{\infty} F_{1,i}(k, \tilde{x})\big(x(k) - \tilde{x}\big)^{[i]} + v(k),$$
$$\qquad\qquad k \geq 0, \qquad x(0) = x_0, \tag{2.5}$$
$$y(k) = \sum_{i=0}^{\infty} H_{1,i}(k, \tilde{x})\big(x(k) - \tilde{x}\big)^{[i]} + w(k),$$

with:

$$F_{1,i}(k, x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes f\right), \qquad H_{1,i}(k, x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes h\right). \tag{2.6}$$

The operator $\nabla_x^{[i]}\otimes$ applied to a function $\psi = \psi(k, x) : I\!N \times I\!R^n \mapsto I\!R^p$ is defined as follows

$$\nabla_x^{[0]} \otimes \psi = \psi, \qquad \nabla_x^{[i+1]} \otimes \psi = \nabla_x \otimes \nabla_x^{[i]} \otimes \psi, \qquad i \geq 1, \tag{2.7}$$

with $\nabla_x = [\partial/\partial x_1 \ \cdots \ \partial/\partial x_n]$. Note that $\nabla_x \otimes \psi$ is the standard Jacobian of the vector function ψ. Analogously, by taking the mth power of the state and the output vectors:

$$x^{[m]}(k+1) = \sum_{i=0}^{\infty} F_{m,i}(k, \tilde{x})\big(x(k) - \tilde{x}\big)^{[i]} + \varphi_m\big(k, \tilde{x}, v(k), x(k) - \tilde{x}\big), \qquad k \geq 0,$$
$$\tag{2.8}$$
$$y^{[m]}(k) = \sum_{i=0}^{\infty} H_{m,i}(k, \tilde{x})\big(x(k) - \tilde{x}\big)^{[i]} + \vartheta_m\big(k, \tilde{x}, w(k), x(k) - \tilde{x}\big), \qquad m \in I\!N.$$

where

$$\varphi_m\big(k,\tilde{x},v(k),x(k)-\tilde{x}\big) = \sum_{i=0}^{\infty} \varphi_{m,i}(k,\tilde{x},v(k))\big(x(k)-\tilde{x}\big)^{[i]},$$

$$\vartheta_m\big(k,\tilde{x},w(k),x(k)-\tilde{x}\big) = \sum_{i=0}^{\infty} \vartheta_{m,i}(k,\tilde{x},w(k))\big(x(k)-\tilde{x}\big)^{[i]},$$

(2.9)

and

$$F_{m,i}(k,x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes f^{[m]}\right), \quad \varphi_{m,i}(k,x,v) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes \big[(f+v)^{[m]} - f^{[m]})\big]\right),$$

$$H_{m,i}(k,x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes h^{[m]}\right), \quad \vartheta_{m,i}(k,x,w) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes \big[(h+w)^{[m]} - h^{[m]})\big]\right).$$

(2.10)

Looking at the definitions (2.10) it can be seen that the sequences $v(k)$ and $w(k)$ appear as multiplicative noises in the definitions (2.9) of the sequences $\varphi_m(k)$ and $\vartheta_m(k)$. Taking into account all the state and measurements equations in (2.8) for $m \geq 1$, the nonlinear system (2.5) is embedded into an infinite-dimensional bilinear one (see [14] and references therein for more details). The finite-dimensional approximation of the bilinear embedding is preliminary to the construction of a polynomial filter, and will be the object of this section. The statement and the proof of the Lemmas below require some definitions and results reported in Appendix. An important formula used in the paper is the one that expresses the Kronecker power of a sum of $\nu + 1$ vectors $z_i \in I\!\!R^p$, $i = 0,1,\ldots,\nu$, by using a multiindex $t \in I\!\!N^{\nu+1}$, $t = \{t_0, t_1, \cdots, t_\nu\}$

$$\left(\sum_{s=0}^{\nu} z_s\right)^{[i]} = \sum_{|t|=i} M_t^p \, \bigotimes_{s=0}^{\nu} z_s^{[t_s]}.$$

(2.11)

with a suitable definition of matrices $M_t^p \in I\!\!R^{p^i \times p^i}$ (see Appendix). The symbol $|t|$ denotes the *modulus* of a multiindex, i.e. $|t| = t_0 + \cdots + t_\nu$, and the symbol $\bigotimes$ denotes Kronecker products of indexed vectors.

Given a nonlinear function $g : I\!\!R^n \mapsto I\!\!R^p$ and a random vector x assuming values in $I\!\!R^n$, let the symbol $[g]_\mu(x,\tilde{x})$ denote the polynomial expansion of $g(x)$ around $\tilde{x}$ up to the degree μ. Lemma A.1 in Appendix proves that $[g^{[m]}]_\mu = \big[[g]_\mu^{[m]}\big]_\mu$. Considering a random vector $\eta \in I\!\!R_p$ and the random vector $\beta = g(x) + \eta$, the symbol $[\beta]_\mu$ will denote the μ-th degree approximation $[g]_\mu(x,\tilde{x}) + \eta$. Lemma A.2 proves that $[\beta^{[m]}]_\mu = \big[[\beta]_\mu^{[m]}\big]_\mu$. Thanks to the result of Lemma A.2, throughout the paper the symbol $\beta_\mu^{[m]}$ will be used to denote either $[\beta^{[m]}]_\mu$ or $\big[[\beta]_\mu^{[m]}\big]_\mu$.

Lemma 2.1. *Let $x_\mu(k)$ and $y_\mu(k)$ be the state and output of the system that approximates (2.1) by neglecting in (2.5) all the powers greater than μ, according to the Carleman approximation scheme. Then $x_\mu^{[m]}$ and $y_\mu^{[m]}$, $m \geq 1$, evolve according to the following equations:*

$$x_\mu^{[m]}(k+1) = \sum_{i=1}^{\mu} A_{m,i}^\mu(k,\tilde{x}) x_\mu^{[i]}(k) + u_m^\mu(k,\tilde{x}) + v_m^\mu(k),$$

$$y_\mu^{[m]}(k) = \sum_{i=1}^{\mu} C_{m,i}^\mu(k,\tilde{x}) x_\mu^{[i]}(k) + \gamma_m^\mu(k,\tilde{x}) + w_m^\mu(k),$$

$$k \geq 0, \quad x_\mu(0) = x_0, \qquad (2.12)$$

6.

where

$$A_{ij}^{\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{ij}^{\mu}} M_r^n \overline{F}_r(k, \tilde{x}) \big(M_{\alpha(r)-j,j}^n \otimes \xi_{r_{\mu+1}}^v\big)\big(I_{n^j} \otimes (-\tilde{x})^{[\alpha(r)-j]}\big), \qquad (2.13)$$

$$C_{ij}^{\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{ij}^{\mu}} M_r^q \overline{H}_r(k, \tilde{x}) \big(M_{\alpha(r)-j,j}^n \otimes \xi_{r_{\mu+1}}^w\big)\big(I_{n^j} \otimes (-\tilde{x})^{[\alpha(r)-j]}\big), \qquad (2.14)$$

with $r = \{r_0, \cdots, r_{\mu+1}\}$ *a multi-index in* $\mathbb{N}^{\mu+2}$ *and*

$$\alpha(r) = \sum_{s=1}^{\mu} s\, r_s, \qquad \mathcal{R}_{ij}^{\mu} = \{r \in \mathbb{N}^{\mu+2} : |r| = i, \ j \le \alpha(r) \le \mu\} \qquad (2.15)$$

and the matrices $\overline{F}_r, \overline{H}_r$ *defined as:*

$$\overline{F}_r(k, \tilde{x}) = \left(\boxtimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(k, \tilde{x})\right) \otimes I_{n^{r_{\mu+1}}}, \qquad \overline{H}_r(k, \tilde{x}) = \left(\boxtimes_{s=0}^{\mu} H_{1,s}^{[r_s]}(k, \tilde{x})\right) \otimes I_{q^{r_{\mu+1}}}. \qquad (2.16)$$

The deterministic drifts u_i^{μ}, γ_i^{μ} *and the random sequences* $\{v_i^{\mu}\}$, $\{w_i^{\mu}\}$ *are given by:*

$$u_i^{\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{i0}^{\mu}} M_r^n \overline{F}_r(k, \tilde{x}) \big(\tilde{x}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^v(k)\big),$$

$$\gamma_i^{\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{i0}^{\mu}} M_r^q \overline{H}_r(k, \tilde{x}) \big(\tilde{x}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^w(k)\big), \qquad (2.17)$$

$$v_i^{\mu}(k) = \sum_{r \in \mathcal{R}_{i0}^{\mu}} \sum_{s=0}^{\alpha(r)} \Delta_{i,s}^r(k, \tilde{x}) \Big(x_{\mu}^{[s]}(k) \otimes \big(v^{[r_{\mu+1}]}(k) - \xi_{r_{\mu+1}}^v(k)\big)\Big),$$

$$w_i^{\mu}(k) = \sum_{r \in \mathcal{R}_{i0}^{\mu}} \sum_{s=0}^{\alpha(r)} \Phi_{i,s}^r(k, \tilde{x}) \Big(x_{\mu}^{[s]}(k) \otimes \big(w^{[r_{\mu+1}]}(k) - \xi_{r_{\mu+1}}^w(k)\big)\Big), \qquad (2.18)$$

with

$$\Delta_{i,s}^r(k, \tilde{x}) = M_r^n \overline{F}_r(k, \tilde{x}) \Big(M_{\alpha(r)-s,s}^n \big(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]}\big) \otimes I_{n^{r_{\mu+1}}}\Big),$$

$$\Phi_{i,s}^r(k, \tilde{x}) = M_r^q \overline{H}_r(k, \tilde{x}) \Big(M_{\alpha(r)-s,s}^n \big(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]}\big) \otimes I_{q^{r_{\mu+1}}}\Big). \qquad (2.19)$$

Proof. The proof is readily obtained through the application of the result stated in Lemma A.2 (see appendix). ∎

System (2.12) can be put in a more compact form by defining the extended vectors:

$$X^{\mu}(k) = \begin{pmatrix} x_{\mu}(k) \\ \vdots \\ x_{\mu}^{[\mu]}(k) \end{pmatrix} \in \mathbb{R}^{n_{\mu}}, \qquad Y^{\mu}(k) = \begin{pmatrix} y_{\mu}(k) \\ \vdots \\ y_{\mu}^{[\mu]}(k) \end{pmatrix} \in \mathbb{R}^{q_{\mu}}, \qquad (2.20)$$

with $n_{\mu} = n + n^2 + \cdots + n^{\mu}$, $q_{\mu} = q + q^2 + \cdots + q^{\mu}$, so that:

$$X^{\mu}(k+1) = A^{\mu}(k, \tilde{x})X^{\mu}(k) + \mathcal{U}^{\mu}(k, \tilde{x}) + V^{\mu}(k),$$

$$Y^{\mu}(k) = C^{\mu}(k, \tilde{x})X^{\mu}(k) + \Gamma^{\mu}(k, \tilde{x}) + W^{\mu}(k). \qquad (2.21)$$

with the obvious definitions of matrices and vectors, inherited by equation (2.12).

Remark 2.2. From the expressions (2.18) for $v_i^\mu(k)$ and $w_i^\mu(k)$ it can be seen that the approximating system (2.21) is affected by multiplicative noise. Moreover, it is easily proved that the sequences $v_i^\mu(k)$ and $w_i^\mu(k)$ are zero-mean, because the state $x_\mu(k)$ is independent of the noises $v(k)$ and $w(k)$ at the same instant.

Lemma 2.3. *The noises* $\{V^\mu\}$, $\{W^\mu\}$ *are sequences of uncorrelated random vectors (white noise sequences). Moreover, named* Ψ^{V^μ}, Ψ^{W^μ} *their covariance matrices, with:*

$$\Psi_{ij}^{V^\mu}(k, \tilde{x}) = I\!E\left[v_i^\mu(k)v_j^\mu(k)^T\right], \qquad \Psi_{ij}^{W^\mu}(k, \tilde{x}) = I\!E\left[w_i^\mu(k)w_j^\mu(k)^T\right], \qquad i, j = 1, \ldots, \mu, \quad (2.22)$$

the block matrices composing them, then

$$\Psi_{ij}^{V^\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \mathrm{st}_{n^i, n^j}^{-1}\left(\left(\Delta_{j,s}^r(k, \tilde{x}) \otimes \Delta_{i,l}^t(k, \tilde{x})\right)\right.$$
$$\left. \cdot \left(I_{n^s} \otimes C_{n^{l+t_{\mu+1}}, n^{r_{\mu+1}}}^T\right)\left(Z_{s+l}^\mu(k) \otimes \left(\xi_{t_{\mu+1}+r_{\mu+1}}^v(k) - \xi_{t_{\mu+1}}^v(k) \otimes \xi_{r_{\mu+1}}^v(k)\right)\right)\right), \tag{2.23}$$

$$\Psi_{ij}^{W^\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \mathrm{st}_{q^i, q^j}^{-1}\left(\left(\Phi_{j,s}^r(k, \tilde{x}) \otimes \Phi_{i,l}^t(k, \tilde{x})\right)\right.$$
$$\left. \cdot \left(I_{n^s} \otimes C_{n^l q^{t_{\mu+1}}, q^{r_{\mu+1}}}^T\right)\left(Z_{s+l}^\mu(k) \otimes \left(\xi_{t_{\mu+1}+r_{\mu+1}}^w(k) - \xi_{t_{\mu+1}}^w(k) \otimes \xi_{r_{\mu+1}}^w(k)\right)\right)\right). \tag{2.24}$$

where $Z_i^\mu(k)$, $i = 0, \ldots, 2\mu$, *are the expected values* $I\!E\left[x_\mu^{[i]}(k)\right]$, *computed as*

$$Z_i^\mu(k+1) = \sum_{j=1}^\mu A_{ij}^\mu(k, \tilde{x})Z_j^\mu(k) + u_i^\mu(k, \tilde{x}), \qquad k \geq 0, \qquad\qquad i = 1, \ldots, 2\mu, \tag{2.25}$$
$$Z_i^\mu(0) = \zeta_i^0$$

Matrices $C_{a,b}^T$ *in (2.23) and (2.24) are the commutation matrices of Kronecker products, while* $\mathrm{st}_{a,b}^{-1}$ *is the inverse of the stack operator giving matrices in* $I\!R^{a \times b}$ *(see Appendix and [8]).*

Proof. In order to show that the extended noises are both sequences of uncorrelated random variables, let $k > h$; by using the Kronecker product properties [8]:

$$I\!E\left[v_j^\mu(h) \otimes v_i^\mu(k)\right] = \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \left(\Delta_{j,s}^r(h, \tilde{x}) \otimes \Delta_{i,l}^t(k, \tilde{x})\right)$$
$$\cdot I\!E\left[x_\mu^{[s]}(h) \otimes \left(v^{[r_{\mu+1}]}(h) - \xi_{r_{\mu+1}}^v(h)\right) \otimes x_\mu^{[l]}(k) \otimes \left(v^{[t_{\mu+1}]}(k) - \xi_{t_{\mu+1}}^v(k)\right)\right]. \tag{2.26}$$

Now $v(k)$ is independent of $v(h)$, as $\{v(\tau), \ \tau \in I\!N\}$ is a sequence of independent random vectors. Moreover, $x_\mu(k)$ only depends on the noise sequence $\{v(\tau), \ \tau < k\}$, so that, the expectation in (2.26) gives:

$$I\!E\left[x_\mu^{[s]}(h) \otimes \left(v^{[r_{\mu+1}]}(h) - \xi_{r_{\mu+1}}^v(h)\right) \otimes x_\mu^{[l]}(k)\right] \otimes I\!E\left[\left(v^{[t_{\mu+1}]}(k) - \xi_{t_{\mu+1}}^v(k)\right)\right] = 0. \tag{2.27}$$

8.

Analogously, it comes that also $\{W^\mu\}$ is a sequence of uncorrelated random variables. Moreover $\{V^\mu\}$ is uncorrelated with $\{W^\mu\}$, in that, from (2.26),

$$\mathbb{E}\big[w_j^\mu(h) \otimes v_i^\mu(k)\big] = \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \Big(\Phi_{j,s}^r(h,\tilde{x}) \otimes \Delta_{i,l}^t(k,\tilde{x})\Big)$$

$$\cdot \, \mathbb{E}\Big[x_\mu^{[s]}(h) \otimes \big(w^{[r_{\mu+1}]}(h) - \xi_{r_{\mu+1}}^w(h)\big) \otimes x_\mu^{[l]}(k) \otimes \big(v^{[t_{\mu+1}]}(k) - \xi_{t_{\mu+1}}^v(k)\big)\Big].$$

$$(2.28)$$

According to the independence of the state and measurements noise sequences, then $\forall k, h \in \mathbb{N}$, the last expectation is:

$$\mathbb{E}\Big[x_\mu^{[s]}(h) \otimes \mathbb{E}\big[w^{[r_{\mu+1}]}(h) - \xi_{r_{\mu+1}}^w(h)\big] \otimes x_\mu^{[l]}(k) \otimes \big(v^{[t_{\mu+1}]}(k) - \xi_{t_{\mu+1}}^v(k)\big)\Big] = 0. \qquad (2.29)$$

At last, the block matrices building the covariances for the extended state noises are

$$\Psi_{ij}^{V^\mu}(k,\tilde{x}) = \mathbb{E}\Big[v_i^\mu(k)v_j^\mu(k)^T\Big] = \mathrm{st}_{n^i,n^j}^{-1}\Big(\mathbb{E}\big[v_j^\mu(k) \otimes v_i^\mu(k)\big]\Big)$$

$$= \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \mathrm{st}_{n^i,n^j}^{-1}\Big(\Big(\Delta_{j,s}^r(k,\tilde{x}) \otimes \Delta_{i,l}^t(k,\tilde{x})\Big)$$

$$\cdot \, \mathbb{E}\Big[x_\mu^{[s]}(k) \otimes \big(v^{[r_{\mu+1}]}(k) - \xi_{r_{\mu+1}}^v(k)\big) \otimes x_\mu^{[l]}(k) \otimes \big(v^{[t_{\mu+1}]}(k) - \xi_{t_{\mu+1}}^v(k)\big)\Big]\Big)$$

$$= \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \mathrm{st}_{n^i,n^j}^{-1}\Big(\Big(\Delta_{j,s}^r(k,\tilde{x}) \otimes \Delta_{i,l}^t(k,\tilde{x})\Big)$$

$$\cdot \, \big(I_{n^s} \otimes C_{n^{l+t_{\mu+1}},n^{r_{\mu+1}}}^T\big)\Big(\mathbb{E}\big[x_\mu^{[s+l]}(k)\big]$$

$$\otimes \mathbb{E}\Big[\big(v^{[t_{\mu+1}]}(k) - \xi_{t_{\mu+1}}^v(k)\big) \otimes \big(v^{[r_{\mu+1}]}(k) - \xi_{r_{\mu+1}}^v(k)\big)\Big]\Big)\Big).$$

$$(2.30)$$

Defining $Z_i^\mu(k) = \mathbb{E}\big[x_\mu^{[i]}(k)\big]$, $i = 0, 1, \ldots, 2\mu$, equation (2.30) gives back (2.23). Similar computation provide equation (2.24). The recursive equation (2.25) is obtained through the expectation of equation (2.12). ∎

3. The filtering algorithm

As previously mentioned, the proposed algorithm aims to extend and improve the applicability and the results of the Extended Kalman Filter in all those cases where the standard linear approximation of the system is not satisfactory for the solution of state estimation problem. A μ-th degree polynomial filter can be computed by using the μ-th order Carleman approximation described in Lemma 2.1. More in details, at each step k, the extended state and output equations are obtained as the embedding of the μth-degree polynomial approximation of the state and measurements vectors around the estimate $\hat{x}(k)$ and the prediction $\hat{x}(k+1|k)$ respectively. Then, following [8], the μth-degree polynomial minimum variance filter for the approximated system (the projection onto the Hilbert space of all the μ-th order polynomial transformations of the measurements) is achieved applying the standard Kalman equations to the bilinear extended system. In the resulting Riccati equations the covariances of the extended noises are needed, and these can be computed using the recursive equations presented in Lemma 2.3.

As in the case of the classical EKF, the Polynomial Extended Kalman Filter (PEKF) 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 nonlinear system is adopted. Here follows the steps of the PEKF algorithm.

The Polynomial Extended Kalman Filter (PEKF)

I) Computation of the initial conditions of the filter:

$$\widehat{X}^{\mu}(0|-1) = I\!E\left[X^{\mu}(0)\right] = \begin{bmatrix} \zeta_1^0 \\ \vdots \\ \zeta_{\mu}^0 \end{bmatrix}, \qquad \begin{array}{l} a\ priori\ \text{estimate} \\ \text{of the initial extended state;} \end{array}$$

$$P_P(0) = \text{Cov}\left(X^{\mu}(0)\right), \quad \text{Covariance of the } a\ priori \text{ estimate;}$$

$$Z_i^{\mu}(0) = \zeta_i^0, \quad i = 1, \ldots, 2\mu, \text{ initialization of (2.25);}$$

$$k = -1, \quad \text{inizialization of the counter;}$$

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

$$\overline{C}^{\mu}(k+1) = C^{\mu}\left(k+1, \hat{x}(k+1|k)\right), \qquad \text{using (2.14)}$$

$$\overline{\Gamma}^{\mu}(k+1) = \Gamma^{\mu}\left(k+1, \hat{x}(k+1|k)\right), \qquad \text{using (2.17)} \qquad (3.1)$$

$$\overline{\Psi}^{W^{\mu}}(k+1) = \Psi^{W^{\mu}}\left(k+1, \hat{x}(k+1|k)\right) \qquad \text{using (2.24);}$$

(note that $\Psi^{W^{\mu}}(k+1)$ requires $Z_i^{\mu}(k+1)$, $i = 1, \ldots, 2\mu$).

III) computation of the extended output prediction:

$$\widehat{Y}^{\mu}(k+1|k) = \overline{C}^{\mu}(k+1)\widehat{X}^{\mu}(k+1|k) + \overline{\Gamma}^{\mu}(k+1); \qquad (3.2)$$

IV) computation of the Kalman gain:

$$K(k+1) = P_P(k+1)\overline{C}^{\mu}(k+1)^T\left(\overline{C}^{\mu}(k+1)P_P(k+1)\overline{C}^{\mu}(k+1)^T + \overline{\Psi}^{W^{\mu}}(k+1)\right)^{\dagger}; \qquad (3.3)$$

V) computation of the error covariance matrix:

$$P(k+1) = \left(I_{n_\mu} - K(k+1)\overline{C}^\mu(k+1)\right)P_P(k+1); \qquad (3.4)$$

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

$$\widehat{X}^\mu(k+1) = \widehat{X}^\mu(k+1|k) + K(k+1)\left(Y^\mu(k+1) - \widehat{Y}^\mu(k+1|k)\right),$$
$$\hat{x}(k+1) = [I_n \ O_{n\times(n_\mu-n)}]\widehat{X}^\mu(k+1); \qquad (3.5)$$

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

VIII) computation of the μ-th degree approximation of the extended state equation around $\hat{x}(k)$

$$\overline{A}^\mu(k) = A^\mu\big(k, \hat{x}(k)\big), \qquad \text{using (2.13)}$$
$$\overline{\mathcal{U}}^\mu(k) = \mathcal{U}^\mu\big(k, \hat{x}(k)\big), \qquad \text{using (2.17)} \qquad (3.1)$$
$$\overline{\Psi}^{V^\mu}(k) = \Psi^{V^\mu}\big(k, \hat{x}(k)\big), \qquad \text{using (2.23)};$$

(also in this case, $\overline{\Psi}^{V^\mu}(k)$ requires $Z_i^\mu(k)$, $i = 0, 1, \ldots, 2\mu$);

IX) computation of the extended state prediction:

$$\widehat{X}^\mu(k+1|k) = \overline{A}^\mu(k)\widehat{X}^\mu(k) + \overline{\mathcal{U}}^\mu(k); \qquad (3.7)$$

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

$$P_P(k+1) = \overline{A}^\mu(k)P(k)\overline{A}^\mu(k)^T + \overline{\Psi}^{V^\mu}(k); \qquad (3.8)$$

XI) computation of $Z_i^\mu(k+1)$, $i = 1, \ldots, 2\mu$, by using (2.25). 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 nonlinear system:

$$
\begin{aligned}
x_1(k+1) &= 0.8x_1(k) + x_1(k)x_2(k) + 0.1 + \alpha v_1(k), \qquad \alpha = 0.01, \\
x_2(k+1) &= 1.5x_2(k) - x_1(k)x_2(k) + 0.1 + \alpha v_2(k), \\
y(k) &= x_2(k) + \alpha w(k),
\end{aligned} \tag{4.1}
$$

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

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

According also to the nature of the original nonlinear maps, a second order filter has been here proposed. In the following plots, the estimates obtained with the proposed filtering algorithm are compared to those obtained with the standard EKF.

The improvements of the PEKF over the EKF can be recognized by comparing the sampling variances of the estimation errors:

$$
\begin{aligned}
\sigma_1^2(EKF) &= 2.52 \cdot 10^{-3}, & \sigma_1^2(PEKF_{\mu=2}) &= 1.86 \cdot 10^{-3}, \\
\sigma_2^2(EKF) &= 3.66 \cdot 10^{-4}, & \sigma_2^2(PEKF_{\mu=2}) &= 2.44 \cdot 10^{-4}.
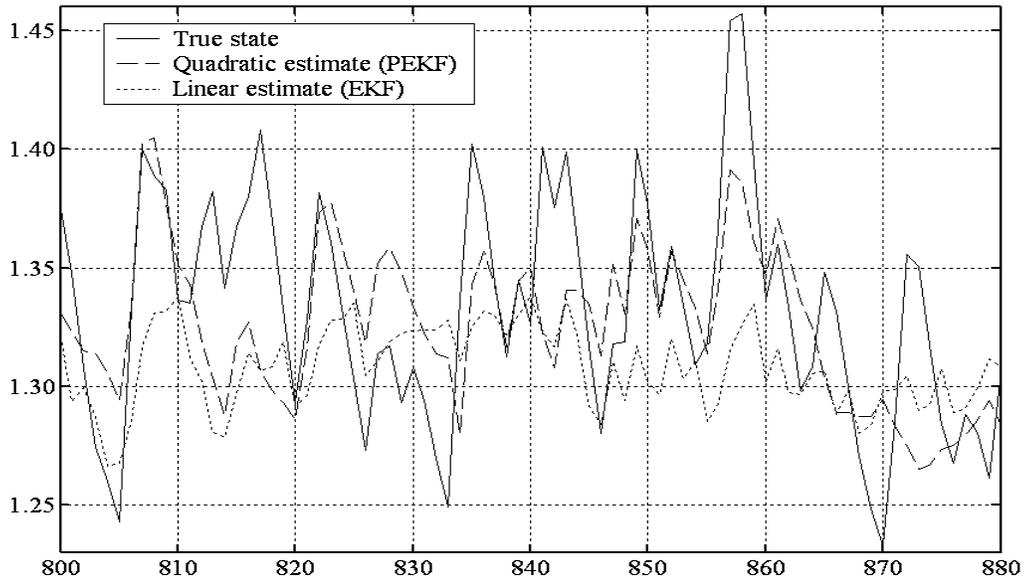\end{aligned} \tag{4.3}
$$

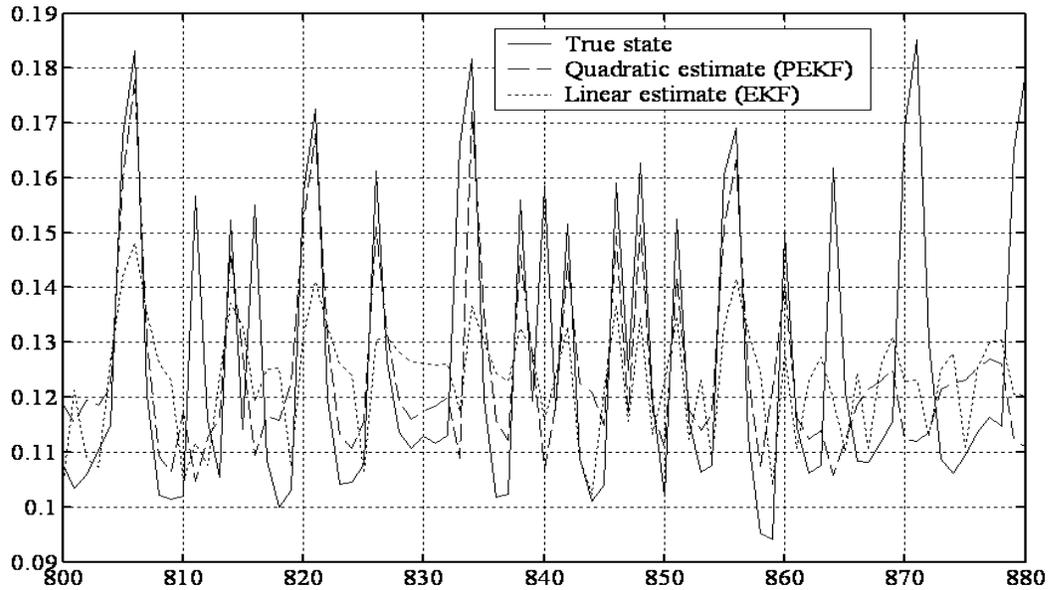Fig. 4.1 – True and estimated state: the first component.

12.

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

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 μth-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 [7, 8]. When $\mu = 1$, the proposed algorithm gives back the standard Extended Kalman Filter.

Appendix: the Kronecker algebra

This appendix reports some definitions and results concerning the Kronecker products and powers that are used in the paper. For a quick survey on Kronecker algebra see [8] and references therein.

Given two matrices $A \in I\!\!R^{r_a \times c_a}$ and $B \in I\!\!R^{r_b \times c_b}$, the Kronecker product $A \otimes B$ is defined as the $(r_a \cdot r_b) \times (c_a \cdot c_b)$ matrix

$$A \otimes B = \begin{bmatrix} a_{1,1}B & \dots & a_{1,c_a}B \\ \vdots & \ddots & \vdots \\ a_{r_a,1}B & \dots & a_{r_a,c_a}B \end{bmatrix}, \tag{A.1}$$

where a_{ij} are the entries of A. The i-th Kronecker power of A is defined as

$$\begin{aligned} A^{[0]} &= 1 \in I\!\!R, \\ A^{[i]} &= A \otimes A^{[i-1]} \quad i \geq 1. \end{aligned} \tag{A.2}$$

Throughout the paper only the superscripts in square brackets have to be intended as Kronecker powers.

The stack of a matrix A is the vector in $I\!\!R^{r_a \cdot c_a}$ 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. When written without any subscript, the inverse stack operator should be intended to generate a square matrix, so that if A is a square matrix then $\mathrm{st}^{-1}(\mathrm{st}(A)) = A$.

Some useful properties of the Kronecker product and stack operation, used in the paper, are the following

$$\begin{aligned} (A + B) \otimes (C + D) &= A \otimes C + A \otimes D + B \otimes C + B \otimes D & \text{(A.3a)} \\ A \otimes (B \otimes C) &= (A \otimes B) \otimes C & \text{(A.3b)} \\ (A \cdot C) \otimes (B \cdot D) &= (A \otimes B) \cdot (C \otimes D) & \text{(A.3c)} \\ (A \otimes B)^T &= A^T \otimes B^T & \text{(A.3d)} \\ u \otimes v &= \mathrm{st}(v \cdot u^T) & \text{(A.3e)} \end{aligned}$$

Given two vectors $x \in I\!\!R^n$ and $z \in I\!\!R^p$, the products $x \otimes z$ and $z \otimes x$ have the same entries in a different order. A commutation matrix, denoted $C_{n,p}^T$, is a square matrix in $\{0,1\}^{np \times np}$ such that

$$z \otimes x = C_{n,p}^T (x \otimes z). \tag{A.4}$$

The Kronecker powers of a sum of vectors can be expanded by using multiindexes and suitably defined matrices. For the purposes of this paper, it is useful to consider a multiindex $t \in I\!\!N^{\nu+1}$ whose entries are numbered from 0 to ν, i.e. $t = \{t_0, t_1, \cdots t_\nu\}$. The modulus of a multiindex, 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 $z_i \in I\!\!R^p$, $i = 0, 1, \ldots, \nu$, can be expressed as

$$(z_0 + z_1 + ... + z_\nu)^{[i]} = \sum_{|t|=i} M_t^p \left(z_0^{[t_0]} \otimes z_1^{[t_1]} \otimes \cdots \otimes z_\nu^{[t_\nu]} \right). \tag{A.5}$$

with a suitable definition of matrices $M_t^p \in \mathbb{R}^{p^i \times p^i}$ (see [8]). Whenever required, we will refer to M_t^p as $M_{t_0,\cdots,t_\nu}^p$. Note that for $k \leq n$, it is $M_{k,n-k}^1 = \binom{n}{k}$.

It is useful to define the (ordered) Kronecker product of n matrices A_h, $h = 1, \ldots n$, with the symbol $\boxtimes$, so that

$$\overset{n}{\underset{h=1}{\boxtimes}} A_h = A_1 \otimes A_2 \otimes \cdots \otimes A_n. \tag{A.6}$$

With this definition, equation (A.5) can be put in compact form as

$$\left(\sum_{h=0}^{\nu} z_h\right)^{[i]} = \sum_{|t|=i} M_t^p \overset{\nu}{\underset{h=0}{\boxtimes}} z_h^{[t_h]}. \tag{A.7}$$

Using the properties (A.3) the following computations can be done

$$\left(\sum_{h=0}^{\nu} A_h z_h\right)^{[i]} = \sum_{|t|=i} M_t^p \overset{\nu}{\underset{h=0}{\boxtimes}} \left(A_h z_h\right)^{[t_h]}$$

$$= \sum_{|t|=i} M_t^p \overset{\nu}{\underset{h=0}{\boxtimes}} A_h^{[t_h]} z_h^{[t_h]} \tag{A.8}$$

$$= \sum_{|t|=i} M_t^p \overset{\nu}{\underset{h=0}{\boxtimes}} A_h^{[t_h]} \overset{\nu}{\underset{h=0}{\boxtimes}} z_h^{[t_h]}.$$

Consider an analytical function $g : \mathbb{R}^n \mapsto \mathbb{R}^p$. Using the Kronecker formalism, the polynomial expansion of $g(x)$ around a point $\tilde{x}$ can be written as

$$g(x) = \sum_{i=0}^{\infty} G_{1,i}(\tilde{x})(x - \tilde{x})^{[i]}, \quad \text{with} \quad G_{1,i}(x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes g\right), \tag{A.9}$$

where $\nabla_x = [\partial/\partial x_1 \ \cdots \ \partial/\partial x_n]$. The operator $\nabla_x^{[i]} \otimes$ is defined as

$$\nabla_x^{[0]} \otimes g(x) = g(x),$$
$$\nabla_x^{[i+1]} \otimes g(x) = \nabla_x \otimes \nabla_x^{[i]} \otimes g(x), \qquad i \geq 1. \tag{A.10}$$

Note that $\nabla_x \otimes g(x)$ is the standard Jacobian of the vector function $g(x)$. Let $[g]_\mu(x, \tilde{x})$ denote the μ-th degree polynomial approximation of $g(x)$, obtained neglecting all the powers greater then μ in the series (A.9). Also the Kronecker powers of $g(x)$ allow a polynomial expansion

$$g^{[m]}(x) = \sum_{i=0}^{\infty} G_{m,i}(\tilde{x})(x - \tilde{x})^{[i]}, \quad \text{with} \quad G_{m,i}(x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes g^{[m]}\right). \tag{A.11}$$

Let $\left[g^{[m]}\right]_\mu(x, \tilde{x})$ denote the μ-th degree polynomial approximation of $g^{[m]}(x)$.

The following function of a multiindex $r \in \mathbb{N}^{\mu+2}$

$$\alpha(r) = \sum_{i=1}^{\mu} i r_i, \tag{A.12}$$

and the following subsets of $I\!N^{\mu+2}$

$$\mathcal{R}_{i,j}^{\mu} = \left\{ r \in I\!N^{\mu+2} : |r| = i, \ j \leq \alpha(r) \leq \mu \right\}, \quad \overline{\mathcal{R}}_{i,j}^{\mu} = \left\{ r \in \mathcal{R}_{i,j}^{\mu} : r_{\mu+1} = 0 \right\} \tag{A.13}$$

are used throughout the paper and in the following Lemma.

Lemma A.1. *The following equalities hold:*

$$\left[g^{[m]} \right]_{\mu}(x, \tilde{x}) = \left[[g]_{\mu}^{[m]} \right]_{\mu}(x, \tilde{x}) = \sum_{r \in \overline{\mathcal{R}}_{m,0}^{\mu}} M_r^p \overline{G}_r(\tilde{x})(x - \tilde{x})^{[\alpha(r)]}, \tag{A.14}$$

where

$$\overline{G}_r(\tilde{x}) = \bigboxtimes_{i=0}^{\mu} G_{1,i}^{[r_i]}(\tilde{x}). \tag{A.15}$$

Proof. By definition

$$[g]_{\mu}(x, \tilde{x}) = \sum_{i=0}^{\mu} G_{1,i}(\tilde{x})(x - \tilde{x})^{[i]}. \tag{A.16}$$

Let $\eta \in I\!R^p$. By using some of the properties (A.3) and the multiindex $r \in I\!N^{\mu+2}$ the following computations can be done

$$
\begin{aligned}
\left([g]_{\mu}(x, \tilde{x}) + \eta \right)^{[m]} &= \left(\sum_{i=0}^{\mu} G_{1,i}(\tilde{x})(x - \tilde{x})^{[i]} + \eta \right)^{[m]} \\
&= \sum_{|r|=m} M_r^p \left[\left(\bigboxtimes_{i=0}^{\mu} \left(G_{1,i}(\tilde{x})(x - \tilde{x})^{[i]} \right)^{[r_i]} \right) \otimes \eta^{[r_{\mu+1}]} \right] \\
&= \sum_{|r|=m} M_r^p \left[\left(\bigboxtimes_{i=0}^{\mu} G_{1,i}^{[r_i]}(\tilde{x})(x - \tilde{x})^{[ir_i]} \right) \otimes \eta^{[r_{\mu+1}]} \right] \\
&= \sum_{|r|=m} M_r^p \left[\left(\bigboxtimes_{i=0}^{\mu} G_{1,i}^{[r_i]}(\tilde{x}) \bigboxtimes_{i=0}^{\mu} (x - \tilde{x})^{[ir_i]} \right) \otimes \eta^{[r_{\mu+1}]} \right].
\end{aligned}
\tag{A.17}
$$

Thanks to the definition (A.12) it is

$$\bigboxtimes_{i=0}^{\mu} (x - \tilde{x})^{[ir_i]} = (x - \tilde{x})^{[\alpha(r)]}, \tag{A.18}$$

and taking into account the definition of matrix $\overline{G}_r(\tilde{x})$ in (A.15), the following equation is obtained

$$\left([g]_{\mu}(x, \tilde{x}) + \eta \right)^{[m]} = \sum_{|r|=m} M_r^p \left[\left(\overline{G}_r(\tilde{x})(x - \tilde{x})^{[\alpha(r)]} \right) \otimes \eta^{[r_{\mu+1}]} \right], \tag{A.19}$$

which is a polynomial of degree $m\mu$ of $x - \tilde{x}$. Taking the summation in the polynomial (A.19) over all the multiindexes r in the set $\mathcal{R}_{m,0}^{\mu}$ defines its truncation to the degree μ:

$$\left[\left([g]_{\mu}(x, \tilde{x}) + \eta \right)^{[m]} \right]_{\mu} = \sum_{r \in \mathcal{R}_{m,0}^{\mu}} M_r^p \left(\overline{G}_r(\tilde{x})(x - \tilde{x})^{[\alpha(r)]} \otimes \eta^{[r_{\mu+1}]} \right). \tag{A.20}$$

16.

Moreover, it is readily verified that setting $\eta = 0$ the summation can be restricted over $r \in \overline{\mathcal{R}}^{\mu}_{m,0}$, obtaining

$$\left[[g]^{[m]}_{\mu}\right]_{\mu}(x,\tilde{x}) = \sum_{r \in \overline{\mathcal{R}}^{\mu}_{m,0}} M^p_r \overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}. \tag{A.21}$$

On the other hand, defining $\rho_{\mu}(x,\tilde{x}) = g(x) - [g]_{\mu}(x,\tilde{x})$ and substituting $\eta = \rho_{\mu}(x,\tilde{x})$ in (A.20) yields

$$g^{[m]}(x,\tilde{x}) = \left([g]_{\mu}(x,\tilde{x}) + \rho_{\mu}(x,\tilde{x})\right)^{[m]}$$
$$= \sum_{|r|=m} M^p_r \left[\left(\overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}\right) \otimes \rho_{\mu}^{[r_{\mu+1}]}(x,\tilde{x})\right]. \tag{A.22}$$

Observing that $\rho_{\mu}(x,\tilde{x})$ is a summation of factors $(x-\tilde{x})^{[i]}$ with $i > \mu$, it is clear that $\left[g^{[m]}\right]_{\mu}(x,\tilde{x})$ can be obtained by taking the summation in (A.22) for all $r \in \overline{\mathcal{R}}^{\mu}_{m,0}$, so that $r_{\mu+1} = 0$ and $\alpha(r) \leq \mu$, obtaining

$$\left[g^{[m]}\right]_{\mu}(x,\tilde{x}) = \sum_{r \in \overline{\mathcal{R}}^{\mu}_{m,0}} M^p_r \overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}. \tag{A.23}$$

Equalities (A.21) and (A.23) give the thesis. ∎

Lemma A.2. *Let $g : I\!\!R^n \mapsto I\!\!R^p$ be an analytic nonlinear map. Let x and η be independent random vectors assuming values in $I\!\!R^n$ and $I\!\!R^p$, respectively. η is assumed to have zero mean value and finite moments up to the μth order:*

$$I\!\!E\left[\eta^{[i]}\right] = \xi_i, \qquad i = 1, \ldots, \mu. \tag{A.24}$$

Consider the random vector

$$\beta = g(x) + \eta. \tag{A.25}$$

Let $[\beta]_{\mu}$ be the random vector obtained considering the μth-degree polynomial expansion of $g(\cdot)$ around a given point $\tilde{x} \in I\!\!R^n$:

$$[\beta]_{\mu}(k) = [g]_{\mu}(x,\tilde{x}) + \eta = \sum_{i=0}^{\mu} G_{1,i}(\tilde{x})(x-\tilde{x})^{[i]} + \eta. \tag{A.26}$$

It is:

$$\left[\beta^{[m]}\right]_{\mu} = \left[[\beta]^{[m]}_{\mu}\right]_{\mu} = \sum_{s=1}^{\mu} \Theta^{\mu}_{m,s}(\tilde{x})x^{[s]} + \Theta^{\mu}_{m,0}(\tilde{x}) + \theta_{\mu,m}(x,\tilde{x},\eta), \tag{A.27}$$

where the matrices $\Theta^{\mu}_{m,s}(\tilde{x})$, $s = 0, \ldots, m$, are defined as

$$\Theta^{\mu}_{m,s}(\tilde{x}) = \sum_{r \in \mathcal{R}^{\mu}_{m,s}} M^p_r \left(\left(\overline{G}_r(\tilde{x})M^n_{\alpha(r)-s,s}\right) \otimes \xi_{r_{\mu+1}}\right)\left(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]}\right), \tag{A.28}$$

with $r \in I\!\!N^{\mu+1}$ a multiindex $r = \{r_0, \cdots, r_{\mu+1}\}$, $\alpha(r)$ defined in (A.12), $\mathcal{R}^{\mu}_{ij}$ defined in (A.13) and $\theta_{\mu,m}(x,\tilde{x},\eta)$ a zero-mean random variable defined as follows:

$$\theta_{\mu,m}(x,\tilde{x},\eta) = \sum_{r \in \mathcal{R}^{\mu}_{m,0}} \sum_{s=0}^{\alpha(r)} \Xi^r_{m,s}(\tilde{x})\left(x^{[s]} \otimes \left(\eta^{[r_{\mu+1}]} - \xi_{r_{\mu+1}}\right)\right), \tag{A.29}$$

with the matrices $\Xi^r_{m,s}(\tilde{x})$ given by

$$\Xi^r_{m,s}(\tilde{x}) = M^p_r\left(\left(\overline{G}_r(\tilde{x})M^p_{\alpha(r)-s,s}\left(I_{n^s}\otimes(-\tilde{x})^{[\alpha(r^{\mu+1})-s]}\right)\right)\otimes I_{p^{r_{\mu+1}}}\right). \tag{A.30}$$

Proof. From (A.20) it is

$$\left[[\beta]^{[m]}_\mu\right]_\mu = \left[\left([g]_\mu(x,\tilde{x})+\eta\right)^{[m]}\right]_\mu = \sum_{r\in\mathcal{R}^\mu_{m,0}} M^p_r\left[\left(\overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}\right)\otimes\eta^{[r_{\mu+1}]}\right]. \tag{A.31}$$

On the other hand, defining $\rho_\mu(x,\tilde{x}) = g(x) - [g]_\mu(x,\tilde{x})$ it is

$$\begin{aligned}\beta^{[m]} &= \left[\left([g]_\mu(x,\tilde{x})+\rho_\mu(x,\tilde{x})+\eta\right)^{[m]}\right]_\mu \\ &= \sum_{|r|=m} M^p_r\left[\left(\overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}\right)\otimes\left(\rho_\mu(x,\tilde{x})+\eta\right)^{[r_{\mu+1}]}\right].\end{aligned} \tag{A.32}$$

Since $\rho_\mu(x,\tilde{x})$ is a sum of factors $(x-\tilde{x})^{[i]}$ for $i>\mu$, the truncation of the polynomial (A.32) to the power μ can be operated by setting $\rho_\mu = 0$ and by restricting $\alpha(r)\leq\mu$, thus obtaining the same summation in (A.31). This proves that $\left[\beta^{[m]}\right]_\mu = \left[[\beta]^{[m]}_\mu\right]_\mu$.

Subtracting and adding the mean values $\xi_{r_{\mu+1}}$ one has

$$\begin{aligned}\left[[\beta]^{[m]}_\mu\right]_\mu &= \sum_{r\in\mathcal{R}^\mu_{m,0}} M^p_r\left[\left(\overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}\right)\otimes\xi_{r_{\mu+1}}\right] \\ &\quad + \sum_{r\in\mathcal{R}^\mu_{m,0}} M^p_r\left[\left(\overline{G}_r(\tilde{x})(x-\tilde{x})^{[\alpha(r)]}\right)\otimes\left(\eta^{[r_{\mu+1}]}-\xi_{r_{\mu+1}}\right)\right].\end{aligned} \tag{A.33}$$

Considering that

$$(x-\tilde{x})^{[\alpha(r)]} = \sum_{s=0}^{\alpha(r)} M^n_{\alpha(r)-s,s}\left(x^{[s]}\otimes(-\tilde{x})^{[\alpha(r)-s]}\right) \tag{A.34}$$

and that $x^{[s]}\otimes(-\tilde{x})^{[\alpha(r)-s]} = \left(I_{n^s}\otimes(-\tilde{x})^{[\alpha(r)-s]}\right)x^{[s]}$, the first summation can be written as

$$\begin{aligned}&\sum_{r\in\mathcal{R}^\mu_{m,0}} M^p_r\left[\left(\overline{G}_r(\tilde{x})\sum_{s=0}^{\alpha(r)} M^n_{\alpha(r)-s,s}\left(x^{[s]}\otimes(-\tilde{x})^{[\alpha(r)-s]}\right)\right)\otimes\xi_{r_{\mu+1}}\right] \\ &= \sum_{r\in\mathcal{R}^\mu_{m,0}}\sum_{s=0}^{\alpha(r)} M^p_r\left(\left(\overline{G}_r(\tilde{x})M^n_{\alpha(r)-s,s}\right)\otimes\xi_{r_{\mu+1}}\right)\left(x^{[s]}\otimes(-\tilde{x})^{[\alpha(r)-s]}\right) \\ &= \sum_{r\in\mathcal{R}^\mu_{m,0}}\sum_{s=0}^{\alpha(r)} M^p_r\left(\left(\overline{G}_r(\tilde{x})M^n_{\alpha(r)-s,s}\right)\otimes\xi_{r_{\mu+1}}\right)\left(I_{n^s}\otimes(-\tilde{x})^{[\alpha(r)-s]}\right)x^{[s]} \\ &= \sum_{s=0}^{\mu}\left(\sum_{r\in\mathcal{R}^\mu_{m,s}} M^p_r\left(\left(\overline{G}_r(\tilde{x})M^n_{\alpha(r)-s,s}\right)\otimes\xi_{r_{\mu+1}}\right)\left(I_{n^s}\otimes(-\tilde{x})^{[\alpha(r)-s]}\right)\right)x^{[s]} \\ &= \Theta^\mu_{m,0}(\tilde{x})+\sum_{s=1}^{\mu}\Theta^\mu_{m,s}(\tilde{x})x^{[s]}.\end{aligned} \tag{A.35}$$

18.

where the terms $\Theta^{\mu}_{m,s}(\tilde{x})$ are defined in (A.28). The second term in (A.33) is a zero-mean random variable, thanks to the independence of the pair (x, η), and following the computations similar to those in (A.35) it can be written as:

$$
\sum_{r \in \mathcal{R}^{\mu}_{m,0}} M^p_r \left[\left(\overline{G}_r(\tilde{x}) \sum_{s=0}^{\alpha(r)} M^n_{\alpha(r)-s,s} \big(x^{[s]} \otimes (-\tilde{x})^{[\alpha(r)-s]} \big) \right) \otimes \big(\eta^{[r_{\mu+1}]} - \xi_{r_{\mu+1}} \big) \right]
$$

$$
= \sum_{r \in \mathcal{R}^{\mu}_{m,0}} \sum_{s=0}^{\alpha(r)} M^p_r \left(\left(\overline{G}_r(\tilde{x}) M^n_{\alpha(r)-s,s} \big(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]} \big) x^{[s]} \right) \otimes \big(\eta^{[r_{\mu+1}]} - \xi_{r_{\mu+1}} \big) \right)
$$

$$
= \sum_{r \in \mathcal{R}^{\mu}_{m,0}} \sum_{s=0}^{\alpha_r \mu} M^p_r \left(\left(\overline{G}_r(\tilde{x}) M^n_{\alpha(r)-s,s} \big(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]} \big) \right) \otimes I_{p^{r_{\mu+1}}} \right)
$$

$$
\cdot \left(x^{[s]} \otimes \big(\eta^{[r_{\mu+1}]} - \xi_{r_{\mu+1}} \big) \right),
$$
(A.36)

that is the random variable $\theta_{\mu,m}(x, \tilde{x}, \eta)$ defined in (A.29), with $\Xi^r_{m,s}(\tilde{x})$ given by (A.30). Substitution of (A.35) and (A.36) in (A.33) gives equation (A.27). ∎

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Englewood Cliffs, NJ: Prentice-Hall, 1979.

[2] M. L. Andrade Netto, L. Gimeno, and M. J. Mendes, "On the optimal and suboptimal nonlinear filtering problem for discrete-time systems," *IEEE Trans. Autom. Contr.*, Vol. 23, pp. 1062–1067, 1978.

[3] J. Baras, A. Bensoussan, and M. R. James, "Dynamic observers as asymptotic limits of recursive filters: Special cases," *SIAM J. Appl. Math.*, Vol. 48, pp. 1147–1158, 1993.

[4] M. Bontayeb, H. Rafarlaky, and M. Darouach, "Covergence analysis of the extended Kalman filter used as an observer for nonlinear discrete-time systems," *IEEE Trans. Autom. Contr.*, Vol. 42, pp. 581–586, 1997.

[5] R. Bucy and P. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, New York: Wiley, 1968.

[6] R. Bucy, "Linear and nonlinear filtering," *Proc. of IEEE*, Vol. 58, No. 6, June 1970.

[7] 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.

[8] F. Carravetta, A. Germani, and M. Raimondi, "Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise," *IEEE Trans. Autom. Contr.*, Vol. 42, pp. 1106–1126, 1997.

[9] C. K. Chui and G. Chen, *Kalman Filtering with Real-Time Applications*, New York: Springer-Verlag, 1987.

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

[11] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[12] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*, New York: Academic, 1970.

[13] Y. Kleyman and I. Mochalov, "Identification of nonstationary systems," *Autom. Rem. Contr.*, Vol. 55, pp. 149–163, 1994.

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

[15] L. Ljung, "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," *IEEE Trans. Autom. Contr.*, Vol. 24, pp. 36–50, 1979.

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

[17] M. G. Safonov, *Stability and Robustness of Multivariable Feedback Systems*, Cambridge MA: MIT Press, 1980.