A FAMILY OF POLYNOMIAL FILTERS FOR DISCRETE-TIME NONLINEAR STOCHASTIC SYSTEMS

A. Germani[1,2] **C. Manes**[1,2] **P. Palumbo**[2]

[1]*Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, Monteluco di Roio, 67040 L'Aquila, Italy*
e-mail: germani@ing.univaq.it, manes@ing.univaq.it

[2]*Istituto di Analisi dei Sistemi e Informatica "A. Ruberti" (CNR) Viale Manzoni 30, 00185 Roma, Italy*
e-mail: palumbo@iasi.rm.cnr.it

Abstract: This work presents a family of polynomial filters for discrete-time nonlinear stochastic systems. These filters can be considered the polynomial version of the well known Extended Kalman Filter. The standard EKF consists in the optimal linear filter applied to the linear approximation of systems. The filters presented in this paper are polynomial filters applied to polynomial approximations of nonlinear systems, and therefore each of them is characterized by a pair of integers: the degree of the system approximation and the degree of the filter. The first filter of the family, the one of order (1,1), coincides with the EKF in the standard form. The implementation of the proposed filters does not require the complete knowledge of the noise distributions, but only the moments up to suitable orders. Numerical simulations show the performances of the filters for some values of the degrees.

Keywords: Polynomial methods, stochastic systems, state estimation.

1. INTRODUCTION

This paper consider the filtering problem for nonlinear stochastic systems of the form

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

where $x(k) \in I\!R^n$ is the system state, $y(k) \in I\!R^q$ is the measured output, $f : I\!R^p \times I\!R^n \mapsto I\!R^n$ and $h : I\!R^p \times I\!R^n \mapsto I\!R^q$ are smooth nonlinear maps. $v(k)$ and $w(k)$, the state and output noises, are independent white sequences, independent of the initial state x_0.

It is well known that the minimum variance estimate of the state requires the computation of the conditional density, a difficult infinite-dimensional problem in the general case (Bucy, 1970; Andrade Netto *et al.*, 1978). Approximate filters can be constructed by computing finite dimensional approximations of the conditional density, e.g. through Gaussian sum approximations (Alspace and Sorensons, 1972; Ito and Xiong, 2000), or through point-mass distributions as in particle filters (Arulampalam *et al.*, 2002). An alternative approach consists in finding an approximation of the stochastic system for which known filtering procedures are available. In this framework, the Extended Kalman Filter (EKF) is the most widely used algorithm (Anderson, 1979; Gelb, 1984; Jazwinski, 1970). Improved versions of the EKF are the iterated EKF and the second order EKF (Gelb, 1984; Jazwinski, 1970). An effective modification of the EKF is the Unscented Kalman Filter (UKF) (Julier and Uhlmann, 2004), that uses the so-called *unscented transform* for the state and output prediction steps in the EKF equations.

The family of filters here proposed extends the idea

This work is supported by *CNR* (Italian National Research Council) and by *MIUR* (Italian Ministry for Education, University and Research).

of the EKF using polynomial techniques, both for the system approximation and the degree of the filter, and therefore will be denoted PEKF (Polynomial EKF) throughout the paper. Each filter in the family is characterized by a pair of integers $(\mu_s,\ \mu_o)$: μ_s is the degree of the polynomial approximation of the nonlinear systems while μ_o is the degree of the polynomial filter (Carravetta, *et al.*, 1996; Carravetta, *et al.*, 1997). The procedure for the construction of the approximated system depends on both the indexes $(\mu_s,\ \mu_o)$ and is similar to the Carleman approximation method (Sastry, 1999; Kowalski *et al.*, 1991). When $\mu_o = 1$ the result coincides with the standard Carleman approximation of order μ_s. The construction of the PEKF's is based on the approximated system. The classical EKF is just the filter with $\mu_s = \mu_o = 1$. Preliminary results on the PEKF have been presented in (Germani *et al.*, 2003), where only the simpler case $\mu_o = \mu_s$ has been considered.

The paper is organized as follows: section two presents the Carleman-like approximation for the system (1) required for the filter construction; in section three the polynomial filter is derived; section four shows numerical comparison of the performances of some filters in the family. Conclusion follows.

2. CARLEMAN-LIKE APPROXIMATION

According to the approach of Carravetta *et al.*(1997), the construction of a polynomial filter of a chosen degree μ_o requires the definition of a polynomial extended system that produces as output the Kronecker powers of the original system up to the order μ_o. The extension of this method to nonlinear systems can be made through the suitable definition of a bilinear system that approximates the original one and produces the required polynomial output. The degree of approximation, indicated with μ_s, can be chosen independently of μ_o. The construction of the approximated system, in a suitable neighborhood of a point $\tilde{x}$, is based on the approximation of the maps f and h by means of the Taylor polynomials of degree μ_s:

$$\begin{aligned} f_{\mu_s}(x,\tilde{x},u) &= \sum_{i=0}^{\mu_s} F_{1,i}(\tilde{x},u)\big(x-\tilde{x}\big)^{[i]} \\ h_{\mu_s}(x,\tilde{x},u) &= \sum_{i=0}^{\mu_s} H_{1,i}(\tilde{x},u)\big(x-\tilde{x}\big)^{[i]} \end{aligned} \tag{2}$$

$$F_{1,i}(x,u) = \frac{1}{i!}\nabla_x^{[i]} \otimes f, \quad H_{1,i}(x,u) = \frac{1}{i!}\nabla_x^{[i]} \otimes h. \tag{3}$$

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

$$\begin{aligned} &\nabla_x^{[0]} \otimes \psi = \psi, \\ &\nabla_x^{[i+1]} \otimes \psi = \nabla_x \otimes \nabla_x^{[i]} \otimes \psi, \quad i \geq 1, \end{aligned} \tag{4}$$

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.

Let $\mu = (\mu_s, \mu_o)$ and $\bar{\mu} = \max(\mu_s, \mu_o)$. For the construction of the approximating system the following moments must exist and be known

$$\begin{aligned} I\!E\{x_0^{[i]}\} &= \zeta_i^0 < \infty, \\ I\!E\{v^{[s]}(k)\} &= \xi_s^v < \infty, \\ I\!E\{w^{[m]}(k)\} &= \xi_m^w < \infty, \end{aligned} \quad \begin{aligned} s &= 1,\dots,2\bar{\mu}, \\ m &= 1,\dots,2\mu_o. \end{aligned} \tag{5}$$

The sequences of the Kronecker powers of the state and output considered for the approximation are $x^{[s]}(k)$ for $s = 1,\dots,\bar{\mu}$ and $y^{[m]}(k)$ for $m = 1,\dots,\mu_o$. The update equations

$$\begin{aligned} x^{[s]}(k+1) &= \Big(f\big(x(k),u(k)\big) + v(k)\Big)^{[s]}, \\ y^{[m]}(k) &= \Big(h\big(x(k),u(k)\big) + w(k)\Big)^{[m]}, \end{aligned} \tag{6}$$

are approximated, in a neighborhood of $\tilde{x}$, as

$$\begin{aligned} x^{[s]}(k+1) &\approx \Big[\Big(f_{\mu_s}\big(x(k),\tilde{x},u(k)\big) + v(k)\Big)^{[s]}\Big]_{\bar{\mu}} \\ y^{[m]}(k) &\approx \Big[\Big(h_{\mu_s}\big(x(k),\tilde{x},u(k)\big) + w(k)\Big)^{[m]}\Big]_{\bar{\mu}} \end{aligned} \tag{7}$$

where the subscript $\bar{\mu}$ denotes the truncation of polynomials of $x(k) - \tilde{x}$ to the degree $\bar{\mu}$. The truncated polynomials can be written as follows

$$\begin{aligned} \Big[\big(f_{\mu_s}(x,\tilde{x},u)+v\big)^{[s]}\Big]_{\bar{\mu}} &= \sum_{i=0}^{(s\cdot\mu_s)\wedge\bar{\mu}} F_{s,i}^{\mu}(\tilde{x},u,v)\big(x-\tilde{x}\big)^{[i]} \\ \Big[\big(h_{\mu_s}(x,\tilde{x},u)+w\big)^{[m]}\Big]_{\mu} &= \sum_{i=0}^{(m\cdot\mu_s)\wedge\bar{\mu}} H_{m,i}^{\mu}(\tilde{x},u,w)\big(x-\tilde{x}\big)^{[i]} \end{aligned} \tag{8}$$

where the matrix coefficients $F_{s,i}^{\mu}(\tilde{x},u,v)$ and $H_{m,i}^{\mu}(\tilde{x},u,w)$ are polynomials of v and w of degree s and m, respectively, of the form:

$$F_{s,i}^{\mu}(\tilde{x},u,v) = \sum_{j=0}^{s} \widetilde{F}_{s,i,j}^{\mu}(\tilde{x},u)\Big(I_{n^i} \otimes v^{[j]}\Big), \tag{9}$$

$$H_{m,i}^{\mu}(\tilde{x},u,w) = \sum_{j=0}^{m} \widetilde{H}_{m,i,j}^{\mu}(\tilde{x},u)\Big(I_{n^i} \otimes w^{[j]}\Big), \tag{10}$$

with suitable definitions of the matrix coefficients $\widetilde{F}_{s,i,j}^{\mu}(\tilde{x},u)$ and $\widetilde{H}_{m,i,j}^{\mu}(\tilde{x},u)$ (see Appendix). Expanding in equations (8) the powers of the binomials $(x - \tilde{x})$, as shown in the Appendix, and rearranging the terms gives the structure

$$\Big[\big(f_{\mu_s} + v\big)^{[s]}\Big]_{\bar{\mu}} = \sum_{i=1}^{(s\cdot\mu_s)\wedge\bar{\mu}} \left(\sum_{j=0}^{s} \widetilde{A}_{s,i,j}^{\mu}(\tilde{x},u)\Big(I_{n^i} \otimes v^{[j]}\Big)\right) x^{[i]}$$

$$+\sum_{i=0}^{(s\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{s}\widetilde B^{\mu}_{s,i,j}(\tilde x,u)\Big(I_{n^i}\otimes v^{[j]}\Big)\Big)\tilde x^{[i]}, \quad (11)$$

$$\Big[\big(h_{\mu_s}+w\big)^{[m]}\Big]_{\bar\mu}=$$

$$\sum_{i=1}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde C^{\mu}_{m,i,j}(\tilde x,u)\Big(I_{n^i}\otimes w^{[j]}\Big)\Big)x^{[i]}$$

$$+\sum_{i=0}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde D^{\mu}_{m,i,j}(\tilde x,u)\Big(I_{n^i}\otimes w^{[j]}\Big)\Big)\tilde x^{[i]}. \quad (12)$$

(The expressions for the matrices $\widetilde A^{\mu}_{s,i,j}(\tilde x,u)$, $\widetilde B^{\mu}_{s,i,j}(\tilde x,u)$, $\widetilde C^{\mu}_{s,i,j}(\tilde x,u)$ and $\widetilde D^{\mu}_{s,i,j}(\tilde x,u)$ are not reported here because of their typographical length.) Obviously, using (11) and (12) instead of the exact update equations does not produce the exact sequences $x^{[s]}(k)$ and $y^{[m]}(k)$. Operating the formal substitution of $x^{[s]}(k)$ with $X^{\mu}_s(k)$ and of $y^{[m]}(k)$ with $Y^{\mu}_m(k)$ in the equations (7), and defining the stochastic sequences $V^j(k)=I_{n^i}\otimes v^{[j]}(k)$ and $W^j(k)=I_{n^i}\otimes w^{[j]}(k)$, the following update equations are obtained

$$X^{\mu}_s(k+1)=\sum_{i=1}^{(s\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{s}\widetilde A^{\mu}_{s,i,j}(\tilde x,u(k))V^j(k)\Big)X^{\mu}_i(k)$$
$$+\sum_{i=0}^{(s\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{s}\widetilde B^{\mu}_{s,i,j}(\tilde x,u(k))V^j(k)\Big)\tilde x^{[i]}, \quad (13)$$

$$Y^{\mu}_m(k)=\sum_{i=1}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde C^{\mu}_{m,i,j}(\tilde x,u(k))W^j(k)\Big)X^{\mu}_i(k)$$
$$+\sum_{i=0}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde D^{\mu}_{m,i,j}(\tilde x,u(k))W^j(k)\Big)\tilde x^{[i]}. \quad (14)$$

The initial states are $X^{\mu}_s(0)=x_0^{[s]}$. As usual, $s=1,\dots,\bar\mu$ and $m=1,\dots,\mu_o$. Defining the white sequences

$$V^j_c(k)=I_{n^i}\otimes\big(v^{[j]}(k)-\xi^v_j\big)$$
$$W^j_c(k)=I_{n^i}\otimes\big(w^{[j]}(k)-\xi^w_j\big) \quad (15)$$

the equations (13) and (14) can be put in the form

$$X^{\mu}_s(k+1)=\sum_{i=1}^{(s\cdot\mu_s)\wedge\bar\mu}A^{\mu}_{s,i}X^{\mu}_i(k)+u^{\mu}_s+v^{\mu}_s,$$
$$Y^{\mu}_m(k)=\sum_{i=1}^{(m\cdot\mu_s)\wedge\bar\mu}C^{\mu}_{m,i}X^{\mu}_i(k)+\gamma^{\mu}_m+w^{\mu}_m, \quad (16)$$

where

$$A^{\mu}_{s,i}(k,\tilde x)=\sum_{j=0}^{s}\widetilde A^{\mu}_{s,i,j}(\tilde x,u(k))\big(I_{n^i}\otimes\xi^v_j\big), \quad (17)$$

$$u^{\mu}_s(k,\tilde x)=\sum_{i=0}^{(s\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{s}\widetilde B^{\mu}_{s,i,j}(\tilde x,u(k))\big(I_{n^i}\otimes\xi^v_j\big)\Big)\tilde x^{[i]},$$

$$v^{\mu}_s(k,\tilde x)=\sum_{i=1}^{(s\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{s}\widetilde A^{\mu}_{s,i,j}(\tilde x,u(k))V^j_c(k)\Big)X^{\mu}_i(k)$$
$$+\sum_{i=0}^{(s\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{s}\widetilde B^{\mu}_{s,i,j}(\tilde x,u(k))V^j_c(k)\Big)\tilde x^{[i]}$$

and

$$C^{\mu}_{m,i}(k,\tilde x)=\sum_{j=0}^{m}\widetilde C^{\mu}_{m,i,j}(\tilde x,u(k))\big(I_{n^i}\otimes\xi^w_j\big), \quad (18)$$

$$\gamma^{\mu}_m(k,\tilde x)=\sum_{i=0}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde D^{\mu}_{m,i,j}(\tilde x,u(k))\big(I_{n^i}\otimes\xi^w_j\big)\Big)\tilde x^{[i]},$$

$$w^{\mu}_m(k,\tilde x)=\sum_{i=1}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde C^{\mu}_{m,i,j}(\tilde x,u(k))W^j_c(k)\Big)X^{\mu}_i(k)$$
$$+\sum_{i=0}^{(m\cdot\mu_s)\wedge\bar\mu}\Big(\sum_{j=0}^{m}\widetilde D^{\mu}_{m,i,j}(\tilde x,u(k))W^j_c(k)\Big)\tilde x^{[i]}$$

Note that $A^{\mu}_{s,i}(k,\tilde x)$, $u^{\mu}_s(k,\tilde x)$, $C^{\mu}_{m,i}(k,\tilde x)$ and $\gamma^{\mu}_m(k,\tilde x)$ are deterministic sequences of matrices and vectors, while $v^{\mu}_s(k,\tilde x)$ and $w^{\mu}_m(k,\tilde x)$ are zero-mean stochastic sequences.

The equations (16) of the Carleman-like approximation of system (1) with polynomial extended output can be put in the following compact form

$$X^{\mu}(k+1)=\mathcal{A}^{\mu}(k,\tilde x)X^{\mu}(k)+\mathcal{U}^{\mu}(k,\tilde x)+V^{\mu}(k,\tilde x),$$
$$Y^{\mu}(k)=\mathcal{C}^{\mu}(k,\tilde x)X^{\mu}(k)+\Gamma^{\mu}(k,\tilde x)+W^{\mu}(k,\tilde x), \quad (19)$$

where

$$X^{\mu}(k)=\begin{bmatrix}X^{\mu}_1(k)\\ \vdots\\ X^{\mu}_{\bar\mu}(k)\end{bmatrix}\in I\!R^{n_{\bar\mu}},\ Y^{\mu}(k)=\begin{bmatrix}Y^{\mu}_1(k)\\ \vdots\\ Y^{\mu}_{\mu_o}(k)\end{bmatrix}\in I\!R^{q_{\mu_o}},$$
$$n_{\bar\mu}=n+n^2+\dots+n^{\bar\mu},\quad q_{\mu_o}=q+q^2+\dots+q^{\mu_o}, \quad (20)$$

$$\mathcal{A}^{\mu}=\begin{bmatrix}A^{\mu}_{1,1}&\cdots&A^{\mu}_{1,\bar\mu}\\ \vdots&\ddots&\vdots\\ A^{\mu}_{\bar\mu,1}&\cdots&A^{\mu}_{\bar\mu,\bar\mu}\end{bmatrix},\quad \mathcal{U}^{\mu}=\begin{bmatrix}u^{\mu}_1\\ \vdots\\ u^{\mu}_{\bar\mu}\end{bmatrix}, \quad (21)$$

$$\mathcal{C}^{\mu}=\begin{bmatrix}C^{\mu}_{1,1}&\cdots&C^{\mu}_{1,\bar\mu}\\ \vdots&\ddots&\vdots\\ C^{\mu}_{\mu_o,1}&\cdots&C^{\mu}_{\mu_o,\bar\mu}\end{bmatrix},\quad \Gamma^{\mu}=\begin{bmatrix}\gamma^{\mu}_1\\ \vdots\\ \gamma^{\mu}_{\mu_o}\end{bmatrix}, \quad (22)$$

$$V^{\mu}=\begin{bmatrix}v^{\mu}_1\\ \vdots\\ v^{\mu}_{\bar\mu}\end{bmatrix},\qquad W^{\mu}=\begin{bmatrix}w^{\mu}_1\\ \vdots\\ w^{\mu}_{\mu_o}\end{bmatrix}, \quad (23)$$

The model (19) is a bilinear system with respect to the white sequences (15) (see the definitions of the components of $V^{\mu}(k,\tilde x)$ and $W^{\mu}(k,\tilde x)$, eq.'s (17) and (18)). Exploiting the same arguments used in (Germani *et al.*, 1996 and 1997) it is not difficult, though tedious, to prove that $V^{\mu}(k,\tilde x)$ and $W^{\mu}(k,\tilde x)$ are *uncorrelated* sequences of zero mean *uncorrelated* random vectors. Moreover, the assumptions (5) ensure that all random vectors in (19) have finite means and covariances. The state and output components of the system (19) are aimed to approximate the Kronecker powers of the state and output of the original system (1), i.e. $X_s^\mu(k) \approx x^{[s]}(k)$ and $Y_m^\mu(k) \approx y^{[m]}(k)$.

3. A FAMILY OF POLYNOMIAL FILTERS

This section shows how the approximate system (19) is used for the filter construction: the optimal linear filter for (19) (linear w.r.t. the polynomial extended output $Y^\mu(k)$), provides the μ_o-degree polynomial filter (polynomial w.r.t. the original output $y(k)$) for the μ_s-degree approximation of the original system. Since the extended noises $V^\mu(k,\tilde{x})$ and $W^\mu(k,\tilde{x})$ in eq.'s (19) 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. As in the standard EKF, the system matrices and the covariances needed in the Riccati equations are not evaluated at a fixed $\tilde{x}$ but are computed at each step at the current state estimate and prediction. In particular, the state estimate is used instead of $\tilde{x}$ 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}. Note that the estimate $\hat{x}(k)$ and prediction $\hat{x}(k+1|k)$ of the original state are computed as

$$\begin{aligned} \hat{x}(k) &= [I_n \;\; O_{n\times(n_{\bar{\mu}}-n)}]\widehat{X}^\mu(k), \\ \hat{x}(k+1|k) &= [I_n \;\; O_{n\times(n_{\bar{\mu}}-n)}]\widehat{X}^\mu(k+1|k). \end{aligned} \tag{24}$$

The steps of the polynomial filtering algorithm are summarized below:

POLYNOMIAL FILTER $\mathcal{P}^{\mu_s,\mu_o}$
FOR NONLINEAR SYSTEMS

I) Filter initialization: $k=-1$

$$\widehat{X}^\mu(0|-1) = I\!E\{X^\mu(0)\}, \;\; P_P(0) = \mathrm{Cov}(X^\mu(0))$$

II) computation of the matrices of the output equation of system (19) at $\hat{x}(k+1|k)$

$$\begin{aligned} \overline{\mathcal{C}}^\mu(k+1) &= \mathcal{C}^\mu\big(k+1, \hat{x}(k+1|k)\big), \\ \overline{\Gamma}^\mu(k+1) &= \Gamma^\mu\big(k+1, \hat{x}(k+1|k)\big), \\ \overline{\Psi}^{W^\mu}(k+1) &= \Psi^{W^\mu}\big(k+1, \hat{x}(k+1|k)\big); \end{aligned} \tag{25}$$

III) computation of the prediction of the extended output:

$$\widehat{Y}^\mu(k+1|k) = \overline{\mathcal{C}}^\mu(k+1)\widehat{X}^\mu(k+1|k) + \overline{\Gamma}^\mu(k+1); \tag{26}$$

IV) computation of the Kalman gain:

$$\begin{aligned} K(k+1) &= P_P(k+1)\overline{\mathcal{C}}^\mu(k+1)^T \\ &\cdot\Big(\overline{\mathcal{C}}^\mu(k+1)P_P(k+1)\overline{\mathcal{C}}^\mu(k+1)^T + \overline{\Psi}^{W^\mu}(k+1)\Big)^\dagger \end{aligned} \tag{27}$$

V) computation of the error covariance matrix:

$$P(k+1) = \Big(I_{n_{\bar{\mu}}} - K(k+1)\overline{\mathcal{C}}^\mu(k+1)\Big)P_P(k+1); \tag{28}$$

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

$$\begin{aligned} \widehat{X}^\mu(k+1) &= \widehat{X}^\mu(k+1|k) + K(k+1) \\ &\quad\cdot\Big(Y^\mu(k+1) - \widehat{Y}^\mu(k+1|k)\Big), \\ \hat{x}(k+1) &= [I_n \;\; O_{n\times(n_{\bar{\mu}}-n)}]\widehat{X}^\mu(k+1); \end{aligned} \tag{29}$$

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

VIII) computation of the matrices of the state equation of system (19) at $\hat{x}(k)$:

$$\begin{aligned} \overline{\mathcal{A}}^\mu(k) &= \mathcal{A}^\mu\big(k, \hat{x}(k)\big), \\ \overline{\mathcal{U}}^\mu(k) &= \mathcal{U}^\mu\big(k, \hat{x}(k)\big), \\ \overline{\Psi}^{V^\mu}(k) &= \Psi^{V^\mu}\big(k, \hat{x}(k)\big); \end{aligned} \tag{30}$$

IX) computation of the extended state prediction:

$$\widehat{X}^\mu(k+1|k) = \overline{\mathcal{A}}^\mu(k)\widehat{X}^\mu(k) + \overline{\mathcal{U}}^\mu(k); \tag{31}$$

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

$$P_P(k+1) = \overline{\mathcal{A}}^\mu(k)P(k)\overline{\mathcal{A}}^\mu(k)^T + \overline{\Psi}^{V^\mu}(k); \tag{32}$$

XI) GOTO STEP **II**.

4. SIMULATION RESULTS

This section reports some simulation results that show the improvement of the state estimate when the degree μ_o of the polynomial filter $\mathcal{P}^{\mu_s,\mu_o}$ is increased while the approximation degree μ_s remains fixed. The nonlinear system considered is

$$\begin{aligned} x_1(k+1) &= \big(0.8 + x_2(k)\big)x_1(k) + 0.1 + 0.01v_1(k), \\ x_2(k+1) &= \big(1.5 - x_1(k)\big)x_2(k) + 0.1 + 0.01v_2(k), \\ y(k) &= x_2(k) + 0.04w(k). \end{aligned} \tag{33}$$

The noise sequences $v_1(k)$, $v_2(k)$, $w(k)$ obey the following discrete distributions:

$$\begin{array}{lll} P_{v_1}(-1) = 0.6, & P_{v_2}(-1) = 0.8, & P_w(-7) = 0.3, \\ P_{v_1}(0) = 0.2, & P_{v_2}(4) = 0.2, & P_w(3) = 0.7. \\ P_{v_1}(3) = 0.2, & & \end{array} \tag{34}$$

The filters $\mathcal{P}^{2,1}$, $\mathcal{P}^{2,2}$ and $\mathcal{P}^{2,3}$ of the family of PEKF's have been implemented and compared (the degree of the polynomial filter is increased ($\mu_o = 1, 2, 3$) while the approximation degree is kept constant ($\mu_s = 2$)). For comparison, also the EKF, the 2nd order EKF (Gelb, 1984) and the UKF (Julier and Uhlmann, 2004) have been implemented. However, in this example all these filters gave similar estimates, and therefore only the results obtained using the 2nd order EKF have been reported. The sample error variances computed in a typical simulation over a 1.000 points horizon, are reported in table 1.

Table 1. Steady state error variances

	$\mathcal{P}^{2,1}$	$\mathcal{P}^{2,2}$	$\mathcal{P}^{2,3}$
$\sigma^2_{\tilde{x}_1}$	$2.410 \cdot 10^{-3}$	$2.112 \cdot 10^{-3}$	$9.690 \cdot 10^{-4}$
$\sigma^2_{\tilde{x}_2}$	$4.229 \cdot 10^{-4}$	$3.310 \cdot 10^{-4}$	$5.320 \cdot 10^{-6}$

The error variances obtained with the 2nd order EKF are $\sigma^2_{\tilde{x}_1} = 2.409 \cdot 10^{-3}$ and $\sigma^2_{\tilde{x}_2} = 4.228 \cdot 10^{-4}$, very similar to those obtained with the filter $\mathcal{P}^{2,1}$. The $\mathcal{P}^{2,2}$ achieves a 12% and 22% reduction of the error variances, w.r.t. the filter $\mathcal{P}^{2,1}$, while the filter $\mathcal{P}^{2,3}$ provides error variances with 60% and 98% reduction. Figures 1 reports the true states and their estimates (for the clarity of the representation, only a window of 50 time steps is plotted).

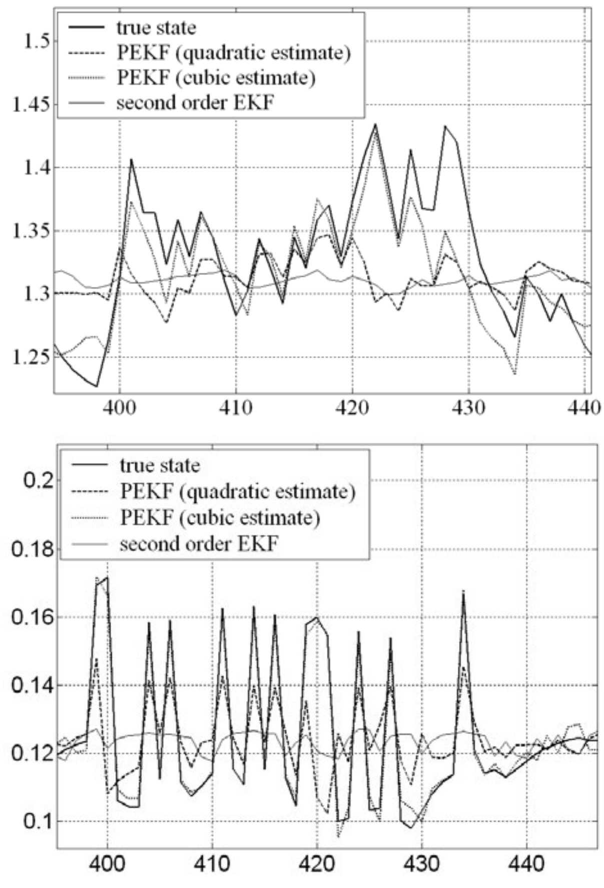

Fig. 1. True and estimated states $x_1(k)$ and $x_2(k)$.

5. CONCLUSION

The application of the polynomial approach for solving the filtering problem of nonlinear stochastic systems has been investigated in this work, and a family of polynomial filters is presented. Each filter in the family is identified by a pair of integers (μ_s, μ_o): the first is the degree of the approximated system used for the filter construction; the second is the degree of the polynomial filter. The two indexes can be independently chosen. The filter with $\mu_s = \mu_o = 1$ coincides with the Extended Kalman Filter in the standard form. Numerical simulations have shown the improvement of the estimate using filters of increasing degree. In particular, a significant reduction of the estimation error variance is achieved by increasing the index μ_o.

APPENDIX

This appendix reports some formulas useful for the computation of matrices and vectors involved in the construction of the approximating system (19) and in the filter equations. A very important formula is the expansion of the Kronecker powers of summations of vectors. For the purposes of this paper, it is useful to consider a multiindex $t = \{t_0, t_1, \cdots\ t_\nu\} \in \left(Z^+\right)^{\nu+1}$. The modulus of a multiindex, denoted $|t|$, is 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

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

with a suitable definition of matrices $M_t^p \in I\!R^{p^i \times p^i}$ (see (Carravetta *et al.*, 1997) for the case $\nu = 1$). Whenever required, we will refer to M_t^p as $M^p_{t_0,\cdots,t_\nu}$. Note that it is $M^1_{k,n-k} = \binom{n}{k}$.

Equation (A.1) can be put in the compact form

$$\left(\sum_{h=0}^{\nu} z_h \right)^{[i]} = \sum_{|t|=i} M_t^p \bigotimes_{h=0}^{\nu} z_h^{[t_h]}. \quad \text{(A.2)}$$

A repeated use of the following property

$$(A \cdot C) \otimes (B \cdot D) = (A \otimes B) \cdot (C \otimes D) \quad \text{(A.3)}$$

allows to work out the identity

$$\left(\sum_{h=0}^{\nu} A_h z_h \right)^{[i]} = \sum_{|t|=i} M_t^p \Big(\bigotimes_{h=0}^{\nu} A_h^{[t_h]} \Big) \Big(\bigotimes_{h=0}^{\nu} z_h^{[t_h]} \Big). \quad \text{(A.4)}$$

Consider now the Taylor approximation of degree μ_s of $f(x, u)$ given by the first of (2). By using some properties of Kronecker product and the multiindex $r = \{r_0, \dots, r_{\mu_s+1}\} \in \left(Z^+\right)^{\mu_s+2}$ the following passages can be made

$$
\begin{aligned}
\Big(f_{\mu_s}(x,\tilde{x},u)+v\Big)^{[s]} &= \sum_{|r|=s} M_r^n \\
&\cdot\left(\Big(\mathop{\boxtimes}_{i=0}^{\mu_s} F_{1,i}^{[r_i]}(\tilde{x},u) \mathop{\boxtimes}_{i=0}^{\mu_s}(x-\tilde{x})^{[ir_i]}\Big)\otimes v^{[r_{\mu_s+1}]}\right), \\
&= \sum_{|r|=s} M_r^n \left(\Big(\overline{F}_r^{\mu_s}(\tilde{x},u)(x-\tilde{x})^{[\alpha(r)]}\Big)\otimes v^{[r_{\mu+1}]}\right) \\
&= \sum_{|r|=s} M_r^n \left(\overline{F}_r^{\mu_s}(\tilde{x},u)\otimes I_{n^{r_{\mu_s+1}}}\right) \qquad \text{(A.5)} \\
&\qquad \cdot\left(I_{n^{\alpha(r)}}\otimes v^{[r_{\mu_s+1}]}\right)(x-\tilde{x})^{[\alpha(r)]}
\end{aligned}
$$

where the function $\alpha(r)$ and the matrix $\overline{F}_r^{\mu_s}(\tilde{x},u)$ are defined as follows

$$
\alpha(r) = \sum_{i=1}^{\mu_s} ir_i, \qquad \overline{F}_r^{\mu_s}(\tilde{x},u) = \mathop{\boxtimes}_{i=0}^{\mu_s} F_{1,i}^{[r_i]}(\tilde{x},u). \tag{A.6}
$$

The truncation of the polynomial (A.5) to the degree $\bar{\mu}$ is obtained through the definition of the following subset of $\left(Z^+\right)^{\mu_s+2}$

$$
\mathcal{S}_{s,i}^{\bar{\mu}} = \left\{r \in \left(Z^+\right)^{\mu_s+2} : |r| = s, \;\; \alpha(r) = i\right\}, \tag{A.7}
$$

(note that if $|r| = s$ then $\alpha(r) \le s\mu_s$). Thus

$$
\left[\left(f_{\mu_s}(x,\tilde{x},u)+v\right)^{[s]}\right]_{\bar{\mu}} = \sum_{i=0}^{(s\mu_s)\wedge\bar{\mu}} F_{s,i}^{\mu}(\tilde{x},u,v)(x-\tilde{x})^{[i]} \tag{A.8}
$$

with $F_{s,i}^{\mu}(\tilde{x},u,v) =$

$$
\sum_{r\in\mathcal{S}_{s,i}^{\bar{\mu}}} M_r^n\left(\overline{F}_r^{\mu_s}(\tilde{x},u)\otimes I_{n^{r_{\mu_s+1}}}\right)\left(I_{n^i}\otimes v^{[r_{\mu_s+1}]}\right) \tag{A.9}
$$

polynomials of degree s of v. Finally, defining the set

$$
\widetilde{\mathcal{S}}_{s,i,j}^{\bar{\mu}} = \left\{r \in \mathcal{S}_{s,i}^{\bar{\mu}}, r_{\mu_s+1} = j\right\} \tag{A.10}
$$

the expressions (8) are obtained with

$$
\widetilde{F}_{s,i,j}^{\mu}(\tilde{x},u) = \sum_{r\in\widetilde{\mathcal{S}}_{s,i,j}^{\bar{\mu}}} M_r^n\left(\overline{F}_r^{\mu_s}(\tilde{x},u)\otimes I_{n^j}\right). \tag{A.11}
$$

Similar computations lead to the expressions (8), (10) with

$$
\widetilde{H}_{m,i,j}^{\mu}(\tilde{x},u) = \sum_{r\in\widetilde{\mathcal{S}}_{m,i,j}^{\bar{\mu}}} M_r^q\left(\overline{H}_r^{\mu_s}(\tilde{x},u)\otimes I_{q^j}\right). \tag{A.12}
$$

From (8) the equations (11), (12) can be obtained expanding the Kronecker powers of the binomials $x - \tilde{x}$ using the following formula

$$
(x-\tilde{x})^{[i]} = \sum_{t_0+t_1=i} M_t^n\left(x^{[t_0]}\otimes(-\tilde{x})^{[t_1]}\right). \tag{A.13}
$$

and rearranging the terms.

The covariances $\Psi^{V^\mu}(k)$ and $\Psi^{W^\mu}(k)$, needed in the filtering algorithm, can be computed following the same lines of (Carravetta *et al.*, 1997; Germani *et al.*, 2003).

REFERENCES

Alspace, D.L. and H.W. Sorenson (1972). Nonlinear Bayesian estimation using Gaussian sum approximation. *IEEE Trans. Automat. Contr.*, **Vol. 17**, pp. 439–448.

Anderson, B.D.O. and J. B. Moore (1979) *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ.

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

Bucy, R. (1970). Linear and nonlinear filtering. *Proc. of IEEE*, **Vol. 58**, No. 6.

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

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

Gelb, A. (1984). *Applied Optimal Estimation*, MIT Press, Cambridge, MA.

Germani, A., C. Manes and P. Palumbo (2003). Polynomial Extended Kalman Filtering for Discrete-Time Nonlinear Stochastic Systems. Proc. of *42nd IEEE Conf. on Decision and Control* (CDC'03), Maui, Hawaii, **Vol. 1**, pp. 886–891.

Ito, K. and K. Xiong (2000). Gaussian filters for nonlinear filtering problems. *IEEE Trans. Automat. Contr.*, **Vol. 45**, pp. 910–927.

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

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

Kowalski, K. and W.H. Steeb (1991). *Nonlinear Dynamical Systems and Carleman Linearization*, World Scientific, Singapore.

Arulampalam, M.S., S. Maskell, N. Gordon and T. Clapp (2002). A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking. *IEEE Trans. Sign. Proc.*, **Vol. 50**, pp. 174–188.

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