Polynomial Extended Kalman Filter

Alfredo Germani, Costanzo Manes, and Pasquale Palumbo

Abstract—This work presents a polynomial version of the well-known extended Kalman filter (EKF) for the state estimation of nonlinear discrete-time stochastic systems. The proposed filter, denoted polynomial EKF (PEKF), consists in the application of the optimal polynomial filter of a chosen degree μ to the Carleman approximation of a nonlinear system. When $\mu = 1$ the PEKF algorithm coincides with the standard EKF. For the filter implementation the moments of the state and output noises up to order 2μ are required. Numerical simulations compare the performances of the PEKF with those of some other existing filters, showing significant improvements.

Index Terms—Extended Kalman filtering, nonlinear stochastic systems, polynomial filtering.

I. INTRODUCTION

This note investigates the problem of state estimation for nonlinear discrete-time stochastic systems of the type:

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

where $x(k) \in I\!R^n$ is the system state, $y(k) \in I\!R^q$ is the measured output, $f : Z^+ \times I\!R^n \mapsto I\!R^n, h : Z^+ \times I\!R^n \mapsto I\!R^q$ are time-varying smooth nonlinear maps, denoted state-transition map and state-output map, respectively. The state and output noises $v(k)$ and $w(k)$ are assumed to be independent white sequences (independent sequences of zero-mean independent random vectors), not necessarily Gaussian. The initial state x_0 is a random vector independent of both the noise sequences.

It is well known that the minimum variance state estimate requires the knowledge of the conditional probability density, whose computation, in the general case, is a difficult infinite-dimensional problem [3], [8], [9]. Only in few cases does the optimal filter have a finite dimension [28]. For this reason, a great deal of work has been made to devise suboptimal implementable filtering algorithms. A way to approach the problem is to find finite-dimensional approximations of the conditional density using, e.g., Gaussian sum approximations as in [1], [16], or discrete distributions as in particle filters [24].

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 (see, e.g., [2], [5], [6], [13]–[15], [17], [19], [21], and [23]). Because of its *local* nature, the EKF performs well if the initial estimation error and the disturbing noises are small enough (in [22] conditions are given for the boundedness of the error variance). Improved versions of the EKF are the iterated EKF and the second order EKF (see [14], [17]). An effective modification of the EKF is the unscented Kalman filter (UKF) [18], that uses the so-called *un-*

scented transform for the state and output prediction steps in the EKF equations.

The filter here proposed is a *polynomial* extension of the EKF (denoted PEKF throughout the note) and belongs to the last group of methods. The PEKF is obtained by the application of the optimal polynomial filter of [10], [11] to the Carleman approximation of a nonlinear system (see [20], [25]), 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 [10] and [11]). 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 EKF (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.

It is important to stress that the Carleman approximation of continuous time nonlinear systems is a bilinear system w.r.t. the input, whereas discrete-time systems are approximated by means of bilinear system w.r.t. the input and some of its powers. For the continuous-time case the tool of Carleman bilinearization has found some applications in problems of systems approximation [26], [27], [29], since there are many reasons for finding a bilinear approximation of a nonlinear system (see [7]). In recent times, such method has been successfully used in the problem of reduction of large scale systems [4]. On the other hand, the Carleman approach has never been used for the approximation of discrete-time systems, mainly because of the presence of the powers of the input in the approximate model, that makes not useful the approach for control applications. However, in the filtering framework the presence of the powers of the input noise does not constitute a significative limitation: the Carleman approximation for system (1) provides a stochastic bilinear system with respect to the extended state and an extended input noise. For this class of systems the optimal polynomial filter is already available in literature without any further approximation [10], [11]. Surprisingly, to the authors' knowledge, this technique has not been used so far for the construction of suboptimal filters. Preliminary results on this field have been presented by the authors in [12].

The note is organized as follows: the next section presents the Carleman approximation of stochastic nonlinear systems of the type (1); in section three the polynomial minimum variance filter for the Carleman approximation (PEKF) is derived; Section IV displays some numerical results where the performances of the PEKF and of other existing algorithms are compared. An Appendix reports some formulas needed for the implementation of the PEKF.

II. CARLEMAN APPROXIMATION OF STOCHASTIC SYSTEMS

Choose an integer μ and consider the sequences $x^{[m]}(k)$ and $y^{[m]}(k)$ of the Kronecker powers of the states and outputs of system (1) for $m =$

$1, \ldots, \mu$ (here superscripts in square brackets denote the Kronecker powers of vectors and matrices; for a quick survey on the Kronecker algebra see [11]). The update equations for these sequences are

$$x^{[m]}(k+1) = (f(k, x(k)) + v(k))^{[m]}$$
$$y^{[m]}(k) = (h(k, x(k)) + w(k))^{[m]}. \quad (2)$$

Under standard analyticity hypotheses the nonlinear functions $(f + v)^{[m]}$ and $(h + w)^{[m]}$ can be approximated in a suitable neighborhood of a given point $\tilde{x}$ using Taylor polynomials of degree μ

$$(f(k, x(k)) + v(k))^{[m]}$$
$$\approx \sum_{i=0}^{\mu} F_{m,i}(k, \tilde{x})(x(k) - \tilde{x})^{[i]}$$
$$+ \sum_{i=0}^{\mu} \varphi_{m,i}(k, \tilde{x}, v(k))(x(k) - \tilde{x})^{[i]} \quad (3)$$
$$(h(k, x(k)) + w(k))^{[m]}$$
$$\approx \sum_{i=0}^{\mu} H_{m,i}(k, \tilde{x})(x(k) - \tilde{x})^{[i]}$$
$$+ \sum_{i=0}^{\mu} \vartheta_{m,i}(k, \tilde{x}, w(k))(x(k) - \tilde{x})^{[i]} \quad (4)$$

where $\varphi_{m,i}(k, \tilde{x}, v(k))$ and $\vartheta_{m,i}(k, \tilde{x}, v(k))$ are suitably defined polynomials of $v(k)$ and $w(k)$ (see [12]), and

$$F_{m,i}(k, x) = \frac{1}{i!} \left(\nabla_x^{[i]} \otimes f^{[m]} \right)$$
$$H_{m,i}(k, x) = \frac{1}{i!} \left(\nabla_x^{[i]} \otimes h^{[m]} \right) \quad (5)$$

where the operator $\nabla_x^{[i]} \otimes$ applied to a function $\psi = \psi(k, x) : Z^+ \times \mathbb{R}^n \mapsto \mathbb{R}^p$ is defined as

$$\nabla_x^{[0]} \otimes \psi = \psi$$
$$\nabla_x^{[i+1]} \otimes \psi = \nabla_x \otimes \nabla_x^{[i]} \otimes \psi, \qquad i \geq 1 \quad (6)$$

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.

The expansion of the powers of the binomials in the summations in (3) and (4) allow to write these as polynomials of $x(k)$ of degree μ (see [12]). The substitution of the ith power of $x(k)$ in the summations with a vector $X_i^\mu(k)$ of the same dimension (recall that $x^{[i]}(k) \in \mathbb{R}^{n^i}$), and of the mth power of $y(k)$ with a vector $Y_m^\mu(k) \in \mathbb{R}^{q^m}$ in the output equations, yield the recursive equations of the Carleman approximation of order μ around $\tilde{x}$

$$X_m^\mu(k+1) = \sum_{i=1}^{\mu} A_{m,i}^\mu(k, \tilde{x}) X_i^\mu(k) + u_m^\mu(k, \tilde{x}) + v_m^\mu(k, \tilde{x})$$
$$Y_m^\mu(k) = \sum_{i=1}^{\mu} C_{m,i}^\mu(k, \tilde{x}) X_i^\mu(k) + \gamma_m^\mu(k, \tilde{x}) + w_m^\mu(k, \tilde{x}) \quad (7)$$

with $m = 1, \ldots, \mu$ and $X_m^\mu(0) = x_0^{[m]}$. The expressions of the matrices $A_{m,i}^\mu(k, \tilde{x})$, $C_{m,i}^\mu(k, \tilde{x})$, of the deterministic sequences $u_m^\mu(k, \tilde{x})$, $\gamma_m^\mu(k, \tilde{x})$ and of the stochastic sequences $v_m^\mu(k, \tilde{x})$, $w_m^\mu(k, \tilde{x})$ are quite long, and are reported in the Appendix for the reader's convenience. The 2μ (7) of the Carleman

approximation of system (1) can be put in the following compact form:

$$X^\mu(k+1) = \mathcal{A}^\mu(k, \tilde{x}) X^\mu(k) + \mathcal{U}^\mu(k, \tilde{x}) + V^\mu(k, \tilde{x})$$
$$Y^\mu(k) = \mathcal{C}^\mu(k, \tilde{x}) X^\mu(k) + \Gamma^\mu(k, \tilde{x}) + W^\mu(k, \tilde{x}) \quad (8)$$

where

$$X^\mu(k) = \begin{bmatrix} X_1^\mu(k) \\ \vdots \\ X_\mu^\mu(k) \end{bmatrix} \in \mathbb{R}^{n_\mu}$$
$$Y^\mu(k) = \begin{bmatrix} Y_1^\mu(k) \\ \vdots \\ Y_\mu^\mu(k) \end{bmatrix} \in \mathbb{R}^{q_\mu} \quad (9)$$

with $n_\mu = n + n^2 + \cdots + n^\mu$ and $q_\mu = q + q^2 + \cdots + q^\mu$, and

$$\mathcal{A}^\mu = \begin{bmatrix} A_{1,1}^\mu & \cdots & A_{1,\mu}^\mu \\ \vdots & \ddots & \vdots \\ A_{\mu,1}^\mu & \cdots & A_{\mu,\mu}^\mu \end{bmatrix} \quad \mathcal{U}^\mu = \begin{bmatrix} u_1^\mu \\ \vdots \\ u_\mu^\mu \end{bmatrix} \quad V^\mu = \begin{bmatrix} v_1^\mu \\ \vdots \\ v_\mu^\mu \end{bmatrix}$$
$$(10)$$
$$\mathcal{C}^\mu = \begin{bmatrix} C_{1,1}^\mu & \cdots & C_{1,\mu}^\mu \\ \vdots & \ddots & \vdots \\ C_{\mu,1}^\mu & \cdots & C_{\mu,\mu}^\mu \end{bmatrix} \quad \Gamma^\mu = \begin{bmatrix} \gamma_1^\mu \\ \vdots \\ \gamma_\mu^\mu \end{bmatrix} \quad W^\mu = \begin{bmatrix} w_1^\mu \\ \vdots \\ w_\mu^\mu \end{bmatrix}.$$
$$(11)$$

From (35), in Appendix, it is clear that the terms $V^\mu(k, \tilde{x})$ and $W^\mu(k, \tilde{x})$ are bilinear functions of the extended state $X^\mu(k)$ and of zero-mean random vectors uncorrelated with $X^\mu(k)$ of the type $(v^{[h]}(k) - E\{v^{[h]}(k)\})$ and $(w^{[h]}(k) - E\{w^{[h]}(k)\})$ (note that these are white sequences). This fact allows to state that the Carleman approximation (8) has a bilinear structure with respect to an extended white noise sequence. Moreover, exploiting the same arguments used in [10], [11], 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, and that the extended state $X^\mu(k)$ is *uncorrelated* with $W^\mu(j, \tilde{x}) \, \forall j$ and with $V^\mu(j, \tilde{x})$ 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) 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 (8) $(X^\mu(k), Y^\mu(k), V^\mu(k, \tilde{x})$ and $W^\mu(k, \tilde{x}))$ have finite means and covariances, it is necessary to assume that the noises and the initial state of the original system have finite moments up to order 2μ

$$\mathbb{E}\left\{x_0^{[i]}\right\} = \zeta_i^0 < \infty \qquad \begin{array}{l} \mathbb{E}\left\{v^{[i]}(k)\right\} = \xi_i^v(k) < \infty \\ \mathbb{E}\left\{w^{[i]}(k)\right\} = \xi_i^w(k) < \infty \end{array} \quad (12)$$

for $i = 1, \ldots, 2\mu$. The moments $\xi_i^v(k), \xi_i^w(k)$, and ζ_i^0 are needed for the recursive computation of the covariances $\Psi^{V^\mu}(k, \tilde{x})$ and $\Psi^{W^\mu}(k, \tilde{x})$ of the extended noises $V^\mu(k, \tilde{x})$ and $W^\mu(k, \tilde{x})$ (see Appendix). The mean and covariance of the extended state $X^\mu(k)$, also needed for the computation of $\Psi^{V^\mu}(k, \tilde{x})$ and $\Psi^{W^\mu}(k, \tilde{x})$, can be recursively computed using standard formulas for bilinear systems (see Appendix and also [11]). It is worthwhile to note that, differently from what may happen in continuous-time, the Carleman Approximation for discrete-time systems never exhibits finite escape time phenomena.

III. FILTERING ALGORITHM

The previous section has described the μth-order Carleman approximation of a stochastic nonlinear system. The result is a bilinear system driven by white noise, given by (8). For the filter construction it is assumed that the output of the original system (1) is generated in fact by the approximate model (8), and thus in the filter equations $Y_m^\mu(k)$ will coincide with $y^{[m]}(k)$. For a system of the type (8) the optimal linear filter (linear w.r.t. the *extended* measurements) provides the optimal μ-degree polynomial filter w.r.t. the *original* measurements, and can be constructed without any further approximation (see [10] and [11]). Since the extended noises $V^\mu(k, \tilde{x})$ and $W^\mu(k, \tilde{x})$ in (8) 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 $\tilde{x}$ 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 $\mathcal{C}^\mu, \Gamma^\mu$ and Ψ^{W^μ}, according to the formulas reported in the Appendix. Note that the estimate $\hat{x}(k)$ and prediction $\hat{x}(k+1|k)$ of the original state are computed by selecting from the estimate and prediction of the extended state, $\hat{X}^\mu(k)$ and $\hat{X}^\mu(k+1|k)$, respectively, the first n components

$$\hat{x}(k) = [I_n \ O_{n \times (n_\mu - n)}]\hat{X}^\mu(k)$$
$$\hat{x}(k+1|k) = [I_n \ O_{n \times (n_\mu - n)}]\hat{X}^\mu(k+1|k). \qquad (13)$$

The steps of the PEKF algorithm are summarized here.

I) Computation of the initial conditions of the filter (the *a priori* estimate of the initial extended state and its covariance)

$$\hat{X}^\mu(0 \,|\, -1) = I\!\!E\{X^\mu(0)\}$$
$$P_P(0) = \mathrm{Cov}(X^\mu(0))$$
$$k = -1 \quad \text{inizialization of the counter.} \qquad (14)$$

II) Computation of the matrices of the μth degree approximation of the extended output equation around the point $\hat{x}(k+1|k) = [I_n \ O_{n \times (n_\mu - n)}]\hat{X}^\mu(k+1|k)$

$$\bar{\mathcal{C}}^\mu(k+1) = \mathcal{C}^\mu(k+1, \hat{x}(k+1|k))$$
$$\bar{\Gamma}^\mu(k+1) = \Gamma^\mu(k+1, \hat{x}(k+1|k))$$
$$\bar{\Psi}^{W^\mu}(k+1) = \Psi^{W^\mu}(k+1, \hat{x}(k+1|k)). \qquad (15)$$

[the first two equations are obtained from (11) and (31), while the third is obtained from (39)].

III) Computation of the prediction of the extended output

$$\hat{Y}^\mu(k+1|k) = \bar{\mathcal{C}}^\mu(k+1)\hat{X}^\mu(k+1|k) + \bar{\Gamma}^\mu(k+1). \qquad (16)$$

IV) Computation of the Kalman gain

$$K(k+1) = P_P(k+1)\bar{\mathcal{C}}^\mu(k+1)^T$$
$$(\bar{\mathcal{C}}^\mu(k+1)P_P(k+1)\bar{\mathcal{C}}^\mu(k+1)^T + \bar{\Psi}^{W^\mu}(k+1))^\dagger. \qquad (17)$$

V) Computation of the error covariance matrix:

$$P(k+1) = \left(I_{n_\mu} - K(k+1)\bar{\mathcal{C}}^\mu(k+1)\right)P_P(k+1). \qquad (18)$$

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

$$\hat{X}^\mu(k+1) = \hat{X}^\mu(k+1|k)$$
$$\qquad + K(k+1)(Y^\mu(k+1) - \hat{Y}^\mu(k+1|k))$$
$$\hat{x}(k+1) = [I_n \ O_{n \times (n_\mu - n)}]\hat{X}^\mu(k+1). \qquad (19)$$

VII) Increment of the counter: $k = k + 1$.

VIII) Computation of the matrices of the μth degree approximation of the state-transition around the point $\hat{x}(k)$

$$\bar{\mathcal{A}}^\mu(k) = \mathcal{A}^\mu(k, \hat{x}(k)) \quad \text{from eq.'s (10) and (30)}$$
$$\bar{\mathcal{U}}^\mu(k) = \mathcal{U}^\mu(k, \hat{x}(k)) \quad \text{from eq.'s (10) and (34)}$$
$$\bar{\Psi}^{V^\mu}(k) = \Psi^{V^\mu}(k, \hat{x}(k)) \quad \text{from eq. (38).} \qquad (20)$$

IX) Computation of the extended state prediction:

$$\hat{X}^\mu(k+1|k) = \bar{\mathcal{A}}^\mu(k)\hat{X}^\mu(k) + \bar{\mathcal{U}}^\mu(k). \qquad (21)$$

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

$$P_P(k+1) = \bar{\mathcal{A}}^\mu(k)P(k)\bar{\mathcal{A}}^\mu(k)^T + \bar{\Psi}^{V^\mu}(k). \qquad (22)$$

XI) GOTO STEP **II**).

Remark 1: For consistency with all the developments made in the note, the PEKF algorithm has been here presented in a form that is not computationally optimized, in that the Kronecker powers contain redundant components (if $x \in I\!\!R^n$ then $x^{[i]} \in I\!\!R^{n^i}$, but only $\tilde{n}_i = \binom{n+i-1}{i}$ monomials are independent). Such redundancies can be avoided through the definition of *reduced Kronecker powers*, containing the independent components of ordinary Kronecker powers (see [10]). More in detail, denoting with $x^{(i)} \in I\!\!R^{n_i}$ the reduced ith Kronecker power of x, it is always possible to define a selection matrix $T_i(n) \in I\!\!R^{\tilde{n}_i \times n^i}$ made of 0's and 1's, such that:

$$x^{(i)} = T_i(n)x^{[i]} \qquad (23)$$

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

IV. SIMULATION RESULTS

The performances of the PEKF with $\mu = 2$ and $\mu = 3$ (quadratic and cubic PEKF, respectively) have been compared with those of the standard EKF, the second-order PEKF [14] and the UKF in the augmented state version [18] through computer simulations. In most cases the quadratic and cubic PEKF have given better performances, in terms of error variance. This section reports simulation results obtained in

TABLE I
STEADY-STATE ERROR VARIANCES

	$\sigma_{\tilde{x}_1}^2$	$\sigma_{\tilde{x}_2}^2$
EKF	$2.23 \cdot 10^{-3}$	$4.03 \cdot 10^{-4}$
EKF2	$2.23 \cdot 10^{-3}$	$4.03 \cdot 10^{-4}$
UKF	$2.23 \cdot 10^{-3}$	$4.03 \cdot 10^{-4}$
PEKF$_{\mu=2}$	$1.89 \cdot 10^{-3}$	$3.03 \cdot 10^{-4}$
PEKF$_{\mu=3}$	$1.12 \cdot 10^{-3}$	$5.23 \cdot 10^{-6}$

a critical case in which the second-order EKF (EKF2) and the UKF do not give improved estimates w.r.t. the standard EKF. The nonlinear system considered is the following:

$$x_1(k+1) = 0.8x_1(k) + x_1(k)x_2(k) + 0.1 + 0.01v_1(k)$$
$$x_2(k+1) = 1.5x_2(k) - x_1(k)x_2(k) + 0.1 + 0.01v_2(k)$$
$$y(k) = x_2(k) + 0.04w(k). \tag{24}$$

The noise sequences $v_1(k), v_2(k), w(k)$ are zero-mean and independent, and obey the following discrete distributions:

$$
\begin{aligned}
P\{v_1(k) = -1\} &= 0.6 \\
P\{v_1(k) = 0\} &= 0.2 \\
P\{v_1(k) = 3\} &= 0.2
\end{aligned}
\qquad
\begin{aligned}
P\{v_2(k) = -1\} &= 0.8 \\
P\{v_2(k) = 4\} &= 0.2 \\
P\{w(k) = -7\} &= 0.3 \\
P\{w(k) = 3\} &= 0.7.
\end{aligned}
\tag{25}
$$

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

$$
\begin{aligned}
P\{x_1(0) = 0.4\} &= 0.2 & P\{x_2(0) = 0.1\} &= 0.2 \\
P\{x_1(0) = 0.8\} &= 0.8 & P\{x_2(0) = 0.4\} &= 0.8.
\end{aligned}
\tag{26}
$$

Table I reports the sample error variances in a typical simulation over a 1.000 points horizon.

In this example, the EKF, the EKF2, and the UKF have a very similar behavior, while the quadratic and cubic PEKF perform better. In particular, the quadratic PEKF achieves 15% and 24% reduction of the error variance of the two state components, respectively, w.r.t. standard EKF, while the cubic PEKF achieves 50% and 98% variance reduction. Figs. 1 and 2 report the true and estimated states. For the clarity of the representation, only a window of 70 time steps is reported and the EKF and UKF estimates are not reported because they are extremely similar to those provided by the second-order EKF. Note that in Fig. 2 the cubic state estimate and the true state are quite indistinguishable.

The computational times required by the PEKF and by the other algorithms used for comparison have been evaluated in term of CPU time on a PC with 1.8 GHz clock. All algorithms have been implemented in MATLAB. The CPU time required by the standard EKF implementation is $T_{\text{EKF}} = 0.35$ s, the time required by the EKF2 is $T_{\text{EKF2}} = 0.62$ s, the time of the UKF is $T_{\text{UKF}} = 0.69$ s, the times of the PEKF$_2$ and PEKF$_3$ are $T_{\text{PEKF}_2} = 5.89$ s and $T_{\text{PEKF}_3} = 23.34$ s, respectively.

The performances of the PEKF have been compared also with those of the particle filter [24], denoted PF in the following. Both the PEKF and the PF have the same feature of improving the error variance by increasing the algorithm complexity. Note that the complexity of the PEKF depends on the chosen polynomial degree, while the complexity of the PF depends on the number of particles used. A zero-mean Gaussian noise has been added to the measurement equation of the
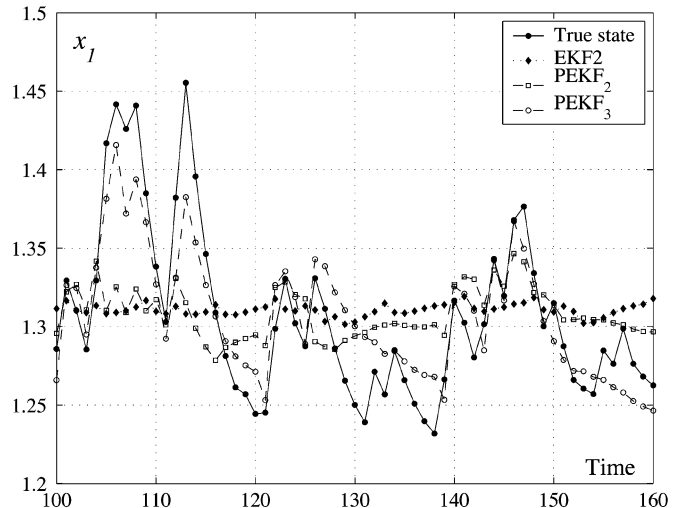

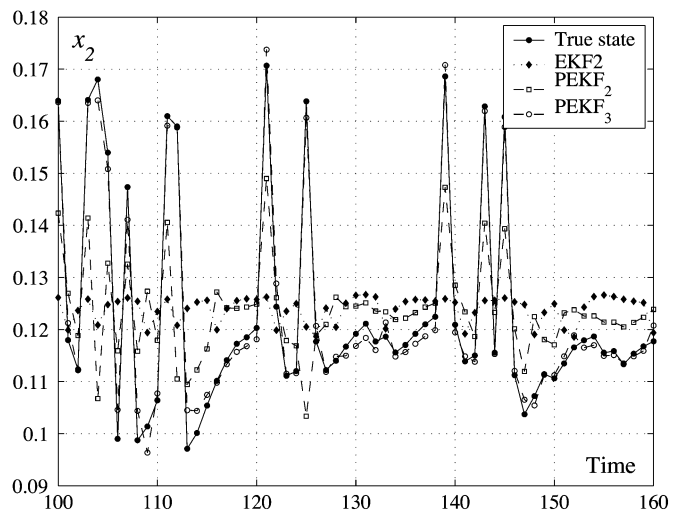Fig. 1.   True and estimated states: the first component.

Fig. 2.   True and estimated states: the second component.

system (24) in order to apply the PF in a standard form (the output noise should not obey a discrete distribution). Many simulations have been performed for different values of the variance of the Gaussian component of the noise and for different numbers of particles. It results that for variances ranging from 0.4 to 1 the PF with about 140 particles (PF_{140}) requires a CPU time similar to the one of the PEKF$_3$. However, the error variance of the PEKF$_3$ is about 20% lower than the error variance of the PF_{140}. The PF with about 250 particles achieves an error variance similar to the one of the PEKF$_3$. However, in this case the CPU time required by the PF_{250} is increased of about 60%. It should be noted that, due to its stochastic nature, the repeated application of a PF to the same sequence of measurements may provide significantly different results, both in term of error variance and CPU time.

V. CONCLUSION

A polynomial extension of the standard EKF for the state estimation of nonlinear discrete-time systems has been proposed in this note. The polynomial algorithm is based on two steps: First the nonlinear system is approximated by a bilinear system using the Carleman approximation of a chosen degree μ; next, the minimum variance filter for the

approximating system among all the μth degree polynomial transformations of the measurements is computed. This step is based on the theory of suboptimal polynomial estimation for linear and bilinear state space representations studied in [10], [11]. When $\mu = 1$, the proposed algorithm gives back the standard EKF. The performances of the proposed filter have been compared with those of other existing filters via computer simulations.

APPENDIX

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 [11] for a quick survey) and take advantage of a formalism that allows to expand Kronecker powers of sums of vectors. Consider a multiindex $t = \{t_0, t_1, \ldots t_\nu\} \in (Z^+)^{\nu+1}$. Its modulus, denoted $|t|$, is defined as the sum of its entries, i.e., $|t| = t_0 + \cdots + t_\nu$. The ith Kronecker power of a sum of $\nu + 1$ vectors $z_i \in \mathbb{R}^p, i = 0, 1, \ldots, \nu$, can be expressed as

$$(z_0 + z_1 + \cdots + z_\nu)^{[i]} = \sum_{|t|=i} M_t^p \left(z_0^{[t_0]} \otimes z_1^{[t_1]} \otimes \cdots \otimes z_\nu^{[t_\nu]} \right) \quad (27)$$

with a suitable definition of matrices $M_t^p \in \mathbb{R}^{p^i \times p^i}$ (see [11]). 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

$$\bigboxtimes_{h=1}^{n} A_h = A_1 \otimes A_2 \otimes \cdots \otimes A_n. \quad (28)$$

With this definition, (27) can be put in the more compact form

$$\left(\sum_{h=0}^{\nu} z_h \right)^{[i]} = \sum_{|t|=i} M_t^p \prod_{h=0}^{\nu} z_h^{[t_h]}. \quad (29)$$

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 $r \cdot c$ into a $r \times c$ matrix.

Lemma 1: The matrices $A_{ij}^\mu(k, \tilde{x})$ and $C_{m,i}^\mu(k, \tilde{x})$ of system (7) are as follows:

$$A_{ij}^\mu(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{ij}^\mu} M_r^n \bar{F}_r(k, \tilde{x})$$
$$\times \left(M_{\alpha(r)-j,j}^n \otimes \xi_{r_{\mu+1}}^v \right) \left(I_{n^j} \otimes (-\tilde{x})^{[\alpha(r)-j]} \right) (30)$$
$$C_{ij}^\mu(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{ij}^\mu} M_r^q \bar{H}_r(k, \tilde{x})$$
$$\times \left(M_{\alpha(r)-j,j}^n \otimes \xi_{r_{\mu+1}}^w \right) \left(I_{n^j} \otimes (-\tilde{x})^{[\alpha(r)-j]} \right) (31)$$

with $r = \{r_0, \ldots, r_{\mu+1}\}$ a multi-index in $(Z^+)^{\mu+2}$ and

$$\alpha(r) = \sum_{s=1}^{\mu} s\, r_s$$
$$\mathcal{R}_{ij}^\mu = \{r \in (Z^+)^{\mu+2} : |r| = i, j \leq \alpha(r) \leq \mu\} \quad (32)$$

the matrices $\bar{F}_r, \bar{H}_r$ in (30), (31) are defined as

$$\bar{F}_r(k, \tilde{x}) = \left(\bigboxtimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(k, \tilde{x}) \right) \otimes I_{n^{r_{\mu+1}}}$$
$$\bar{H}_r(k, \tilde{x}) = \left(\bigboxtimes_{s=0}^{\mu} H_{1,s}^{[r_s]}(k, \tilde{x}) \right) \otimes I_{q^{r_{\mu+1}}}. \quad (33)$$

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

$$u_i^\mu(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{i0}^\mu} M_r^n \bar{F}_r(k, \tilde{x}) \left(\tilde{x}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^v(k) \right)$$
$$\gamma_i^\mu(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{i0}^\mu} M_r^q \bar{H}_r(k, \tilde{x}) \left(\tilde{x}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^w(k) \right) \quad (34)$$

and the random sequences $\{v_i^\mu\}, \{w_i^\mu\}$ are given by

$$v_i^\mu(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \Delta_{i,s}^r(k, \tilde{x})$$
$$\times \left(X_s^\mu(k) \otimes \left(v^{[r_{\mu+1}]}(k) - \xi_{r_{\mu+1}}^v(k) \right) \right)$$
$$w_i^\mu(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \Phi_{i,s}^r(k, \tilde{x})$$
$$\times \left(X_s^\mu(k) \otimes \left(w^{[r_{\mu+1}]}(k) - \xi_{r_{\mu+1}}^w(k) \right) \right) \quad (35)$$

with

$$\Delta_{i,s}^r(k, \tilde{x}) = M_r^n \bar{F}_r(k, \tilde{x})$$
$$\times \left(M_{\alpha(r)-s,s}^n \left(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]} \right) \otimes I_{n^{r_{\mu+1}}} \right)$$
$$\Phi_{i,s}^r(k, \tilde{x}) = M_r^q \bar{H}_r(k, \tilde{x})$$
$$\times \left(M_{\alpha(r)-s,s}^n \left(I_{n^s} \otimes (-\tilde{x})^{[\alpha(r)-s]} \right) \otimes I_{q^{r_{\mu+1}}} \right). \quad (36)$$

Lemma 2: Consider Ψ^{V^μ} and Ψ^{W^μ}, the covariances of the random vectors V^μ and W^μ defined in (10) and (11) whose entries, by definition, are

$$\Psi_{ij}^{V^\mu}(k, \tilde{x}) = \mathbb{E} \left\{ v_i^\mu(k) v_j^\mu(k)^T \right\}$$
$$\Psi_{ij}^{W^\mu}(k, \tilde{x}) = \mathbb{E} \left\{ w_i^\mu(k) w_j^\mu(k)^T \right\}. \quad (37)$$

These can be computed as follows:

$$\Psi_{ij}^{V^\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \Delta_{i,s}^r(k, \tilde{x})$$
$$\times \left(\Psi_{s,l}^{X^\mu}(k) \otimes \Xi_{i,j,t,r}^v(k) \right) \Delta_{j,l}^t(k, \tilde{x})^T \quad (38)$$

$$\Psi_{ij}^{W^\mu}(k, \tilde{x}) = \sum_{r \in \mathcal{R}_{j0}^\mu} \sum_{t \in \mathcal{R}_{i0}^\mu} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \Phi_{i,s}^r(k, \tilde{x})$$
$$\times \left(\Psi_{s,l}^{X^\mu}(k) \otimes \Xi_{i,j,t,r}^w(k) \right) \Phi_{j,l}^t(k, \tilde{x})^T \quad (39)$$

where

$$\Xi_{i,j,t,r}^{v}(k) = \mathrm{st}_{n^{i},n^{j}}^{-1}\left(\xi_{t_{\mu+1}+r_{\mu+1}}^{v}(k) - \xi_{t_{\mu+1}}^{v}(k) \otimes \xi_{r_{\mu+1}}^{v}(k)\right) \tag{40}$$

$$\Xi_{i,j,t,r}^{w}(k) = \mathrm{st}_{q^{i},q^{j}}^{-1}\left(\xi_{t_{\mu+1}+r_{\mu+1}}^{w}(k) - \xi_{t_{\mu+1}}^{w}(k) \otimes \xi_{r_{\mu+1}}^{w}(k)\right) \tag{41}$$

and $\Psi_{ij}^{X^{\mu}}(k,\tilde{x}) = E\{X_{i}^{\mu}(k)X_{j}^{\mu}(k)^{T}\}$ are the blocks of the matrix of second-order moments of the extended state, $\Psi^{X^{\mu}}(k,\tilde{x}) = E\{X^{\mu}(k)X^{\mu}(k)^{T}\}$. This can be computed by the recursive equation

$$\begin{aligned}\Psi^{X^{\mu}}(k+1,\tilde{x}) &= \mathcal{A}^{\mu}(k,\tilde{x})\Psi^{X^{\mu}}(k,\tilde{x})\mathcal{A}^{\mu}(k,\tilde{x})^{T}\\&\quad + \mathcal{A}^{\mu}(k,\tilde{x})Z^{\mu}(k)\mathcal{U}^{\mu}(k,\tilde{x})^{T}\\&\quad + \mathcal{U}^{\mu}(k,\tilde{x})Z^{\mu}(k)^{T}\mathcal{A}^{\mu}(k,\tilde{x})^{T}\\&\quad + \mathcal{U}^{\mu}(k,\tilde{x})\mathcal{U}^{\mu}(k,\tilde{x})^{T} + \Psi^{V^{\mu}}(k,\tilde{x})\end{aligned} \tag{42}$$

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

$$Z^{\mu}(k+1) = \mathcal{A}^{\mu}(k,\tilde{x})Z^{\mu}(k) + \mathcal{U}^{\mu}(k,\tilde{x}). \tag{43}$$

The initialization of (42) and (43) are as follows:

$$Z_{i}^{\mu}(0) = \zeta_{i}^{0}$$
$$\Psi_{ij}^{X^{\mu}}(0,\tilde{x}) = E\left\{x_{0}^{[i]}\left(x_{0}^{[j]}\right)^{T}\right\} = \mathrm{st}_{n^{i},n^{j}}^{-1}\left(\zeta_{i+j}^{0}\right). \tag{44}$$

The proofs of Lemmas 1 and 2, long but not difficult, are straight-forward applications of the rules of the Kronecker algebra.

REFERENCES

[1] D. L. Alspace and H. W. Sorenson, "Nonlinear Bayesian estimation using Gaussian sum approximation," *IEEE Trans. Autom. Control*, vol. AC-17, no. 4, pp. 439–448, Aug. 1972.

[2] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*.   Upper Saddle River, NJ: Prentice-Hall, 1979.

[3] M. L. A. Netto, L. Gimeno, and M. J. Mendes, "On the optimal and suboptimal nonlinear filtering problem for discrete-time systems," *IEEE Trans. Autom. Control*, vol. AC-23, no. 6, pp. 1062–1067, Dec. 1978.

[4] Z. Bai, "Krylov subspace techniques for reduced-order modeling of large-scale dynamical systems," *Appl. Numer. Math.*, vol. 43, pp. 9–44, 2002.

[5] 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.

[6] 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. Control*, vol. 42, no. 4, pp. 581–586, Apr. 1997.

[7] C. Bruni, G. Di Pillo, and G. Koch, "Bilinear systems: An appealing class of *nearly linear* systems in theory and applications," *IEEE Trans. Autom. Control*, vol. AC-19, no. 4, pp. 334–348, Aug. 1974.

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

[9] R. Bucy, "Linear and nonlinear filtering," *Proc. IEEE*, vol. 58, no. 6, Jun. 1970.

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

[11] ——, "Polynomial filtering of discrete-time stochastic linear systems with multiplicative state noise," *IEEE Trans. Autom. Control*, vol. 42, no. 8, pp. 1106–1126, Aug. 1997.

[12] A. Germani, C. Manes, and P. Palumbo, "Polynomial extended Kalman filtering for discrete-time nonlinear stochastic systems," in *Proc. 42nd IEEE Conf. Decision and Control*, vol. 1, Maui, HI, Dec. 2003, pp. 886–891.

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

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

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

[16] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Control*, vol. 45, no. 5, pp. 910–927, May 2000.

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

[18] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[19] Y. Kleyman and I. Mochalov, "Identification of nonstationary systems," *Autom. Rem. Control*, vol. 55, pp. 149–163, 1994.

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

[21] L. Ljung, "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," *IEEE Trans. Autom. Control*, vol. AC-24, no. 1, pp. 36–50, Jan. 1979.

[22] K. Reif, S. Günther, E. Yaz, and R. Unbehauen, "Stochastic stability of the discrete-time extended Kalman filter," *IEEE Trans. Autom. Control*, vol. 44, no. 4, pp. 714–728, Apr. 1999.

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

[24] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[25] S. Sastry, *Nonlinear Systems: Analysis, Stability and Control*.   New York: Springer-Verlag, 1999.

[26] S. Svoronos, G. Stephanopoulos, and R. Aris, "Bilinear approximation of general nonlinear dynamic systems with linear inputs," *Int. J. Control*, vol. 31, no. 1, pp. 109–126, 1980.

[27] J. T.-H. Lo, "Global bilinearization of systems with control appearing linearly," *SIAM J. Control Optim.*, vol. 13, no. 4, pp. 879–885, 1975.

[28] W. S. Wong and S. T. Yau, "The estimation algebra of nonlinear filtering systems," in *Mathematical Control Theory*, J. Baillieul and J. C. Willems, Eds.   New York: Springer-Verlag, 1999, pp. 33–65.

[29] N. van de Wouw, H. Nijmeijer, and D. H. van Campen, "Statistical bilinearization in stochastic nonlinear dynamics," in *Proc. 2nd Int. Conf. on Control of Oscillations and Chaos*, vol. 3, 2000, pp. 394–399.