Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Polynomial Filtering and Identification of Discrete-Time Nonlinear Uncertain Stochastic Systems

Alfredo Germani Costanzo Manes Pasquale Palumbo

Abstract— **This paper deals with the problem of system identification and state estimation for nonlinear uncertain stochastic systems, in the discrete-time framework. By suitably extending the state space with the inclusion of the unknown vector of parameters, the filtering and identification problems are simultaneously solved. The algorithm here proposed applies the optimal polynomial filter of a chosen degree μ to the Carleman approximation of the same degree of the extended nonlinear system. Simulations support theoretical results.**

Index Terms—**Polynomial methods, Filtering problems, Parameter estimation, System identification.**

I. INTRODUCTION

This work investigates the problem of simultaneous filtering and parameters identification for nonlinear uncertain stochastic system of the type:

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

where $x(k) \in \mathbb{R}^n$ is the system state, $y(k) \in \mathbb{R}^q$ is the measured output, $\theta \in \mathbb{R}^m$ is the vector of unknown parameters, $\{v(k)\}$ and $\{w(k)\}$ are sequences of zero-mean, auto and mutually independent random vectors, x_0 is a random variable, independent of both the sequences $\{v(k)\}$ and $\{w(k)\}$. The state and output transition maps $f : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}^n$ and $h : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}^q$ are nonlinear maps, smooth with respect to both the parameter and the state vectors.

The most popular real time algorithm for simultaneous state and parameter estimation is the Extended Kalman Filter (EKF) applied to the extended system, whose state is made of the original state and the parameter vector [4]. The diffusion of the EKF is due to its simplicity and to the fact that in many applications it provides good estimates. The EKF applies the optimal linear estimate to the linear approximation of a nonlinear system around the current estimate, and therefore it performs well in those cases in which the initial state estimate is good and the noises have low variance and approximately gaussian distribution. In such cases the state estimate remains close to the true state and the first-order Taylor expansion around such estimate remains a good approximation of the system dynamics. However, in the presence of high level non gaussian noises the state estimate deteriorates and the first-order approximation is no more a good model for the nonlinear system. Other estimation algorithms exist in the literature, most of them based on clever applications of the Recursive Least Squares (see e.g. [8]). Most works deal with the estimation problem for the class of nonlinear systems with the transition and output maps that are affine functions of the unknown parameters [8, 11].

The starting point of the estimation algorithm here proposed is the same of the EKF: the original state is extended by adding the unknown parameters as a new component, i.e. $X(k) = \begin{bmatrix} x^T(k) & \theta^T(k) \end{bmatrix}^T$. Being constant the parameter θ, the update equation considered is $\theta(k+1) = \theta(k)$. Next, the Polynomial Extended Kalman Filter (PEKF) presented in [6, 7] is adopted for the estimation of the extended state. The PEKF is obtained by the application of the optimal polynomial filter [2, 3] to the Carleman approximation of a nonlinear system [9]. In the stochastic discrete-time framework the μ-degree Carleman approximated system consists of a bilinear system (linear drift/multiplicative noise) with respect to an augmented state made up of all the Kronecker of the original state powers up to the degree μ. Defining in the same way an augmented output, it turns out that this can be expressed as a linear function of the augmented state, corrupted by multiplicative noise. Once the approximation is obtained, the recursive equations of the optimal polynomial filter of order μ can be constructed with no further approximations. When $\mu = 1$ the PEKF reduces to the classical EKF. As in the case of the classical EKF, the 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 because a higher degree of approximation of the nonlinear system is adopted and the optimal polynomial estimate is implemented, instead of the linear estimate of the EKF.

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

A. Germani and C. Manes are with the Dipartimento di Ingegneria Elettrica, Università degli Studi dell'Aquila, Poggio di Roio, 67040 L'Aquila, Italy, germani@ing.univaq.it, manes@ing.univaq.it.

P. Palumbo is with Istituto di Analisi ed Informatica "A. Ruberti", IASI-CNR (National Research Council of Italy), Viale Manzoni 30, 00185 Roma, Italy, palumbo@iasi.rm.cnr.it.

0-7803-9568-9/05/$20.00 ©2005 IEEE

II. CARLEMAN APPROXIMATION

As previously mentioned, the unknown parameters vector θ can be treated as a further state component governed by the equation $\theta(k+1)=\theta(k)$, so that system (1) becomes:

$$\begin{aligned} X(k+1) &= f_e\big(X(k)\big)+v_e(k), \\ y(k) &= h_e\big(X(k)\big)+w(k), \end{aligned} \quad X(0)=\begin{bmatrix} x_0 \\ \theta \end{bmatrix}, \quad (2)$$

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

$$\begin{aligned} f_e\big(X\big) &= \begin{bmatrix} f\big(X_2,X_1\big) \\ X_2 \end{bmatrix}, \quad v_e(k)=\begin{bmatrix} v(k) \\ 0 \end{bmatrix} \\ h_e\big(X\big) &= h\big(X_2,X_1\big). \end{aligned} \quad (3)$$

The μ-degree Carleman approximation is applied to the extended system (2). As a first step, consider the sequences of the Kronecker powers $X^{[i]}(k)$ and $y^{[i]}(k)$, for $i=1,\ldots,\mu$ (for a quick survey on the Kronecker algebra see [3]). The update equations for these sequences are

$$\begin{aligned} X^{[i]}(k+1) &= \Big(f_e\big(X(k)\big)+v_e(k)\Big)^{[i]}, \\ y^{[i]}(k) &= \Big(h_e\big(X(k)\big)+w(k)\Big)^{[i]}. \end{aligned} \quad (4)$$

Under standard analyticity hypotheses the nonlinear functions $(f_e+v_e)^{[i]}$ and $(h_e+w)^{[i]}$ are approximated in a suitable neighborhood of a given point $\widetilde{X}$ using Taylor polynomials of degree μ, so that the right hand terms in (4) are replaced by:

$$\begin{aligned} &\sum_{j=0}^{\mu} F_{i,j}(\widetilde{X})\big(X(k)-\widetilde{X}\big)^{[j]}+\sum_{j=0}^{\mu}\varphi_{i,j}\cdot\big(X(k)-\widetilde{X}\big)^{[j]}, \\ &\sum_{j=0}^{\mu} H_{i,j}(\widetilde{X})\big(X(k)-\widetilde{X}\big)^{[j]}+\sum_{j=0}^{\mu}\psi_{i,j}\cdot\big(X(k)-\widetilde{X}\big)^{[j]}, \end{aligned} \quad (5)$$

with $\varphi_{i,j}=\varphi_{i,j}(\widetilde{X},v_e(k))$ and $\psi_{i,j}=\psi_{i,j}(\widetilde{X},w(k))$ are suitably defined polynomials of $v_e(k)$ and $w(k)$, and

$$F_{i,j}(X)=\frac{\big(\nabla_X^{[j]}\otimes f_e^{[i]}\big)}{j!}, \quad H_{i,j}(X)=\frac{\big(\nabla_X^{[j]}\otimes h_e^{[i]}\big)}{j!}; \quad (6)$$

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

$$\nabla_x^{[0]}\otimes\phi=\phi, \quad \nabla_x^{[j+1]}\otimes\phi=\nabla_x\otimes\nabla_x^{[j]}\otimes\phi, \quad (7)$$

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

Remark 1: From a computational point of view, the Carleman coefficients $F_{1,j}$ and $H_{1,j}$ are directly achieved from the original maps:

$$F_{1,j}(X)=\frac{1}{j!}\begin{bmatrix} \nabla_X^{[j]}\otimes f(\theta,x) \\ \nabla_X^{[j]}\otimes\theta \end{bmatrix}_{(\theta,x)=(X_2,X_1)}, \quad (8)$$

$$H_{1,j}(X)=\frac{1}{j!}\Big(\nabla_X^{[j]}\otimes h(\theta,x)\Big)_{(\theta,x)=(X_2,X_1)}, \quad (9)$$

with

$$\nabla_X^{[j]}\otimes\theta=\begin{cases} \theta, & \text{for } j=0, \\ \big[O_{m\times n} \ \ I_m\big], & \text{for } j=1, \\ O_{m\times n_e^j}, & \text{for } j>1. \end{cases} \quad (10)$$

In the Appendix it is shown that all the Carleman coefficients related to the powers of f_e and h_e can be computed from $F_{1,j}$ and $H_{1,j}$, and therefore from (8) and (9). •

The expansion of the powers of binomials in the summations in eq.s (5) allows to write these as polynomials of $X(k)$ of degree μ [6, 7]. The substitution of the j-th power of $X(k)$ in the summations with a vector $X_j^{\mu}(k)$ of the same dimension (recall that $X^{[j]}(k)\in I\!R^{n_e^j}$), and of the i-th power of $y(k)$ with a vector $Y_i^{\mu}(k)\in I\!R^{q^i}$ in the output equations, yield the equations of the Carleman approximation of order μ around $\widetilde{X}$:

$$\begin{aligned} X_i^{\mu}(k+1) &= \sum_{j=1}^{\mu} A_{i,j}^{\mu}(\widetilde{X})X_j^{\mu}(k)+u_i^{\mu}(k)+v_i^{\mu}(k), \\ Y_i^{\mu}(k) &= \sum_{j=1}^{\mu} C_{i,j}^{\mu}(\widetilde{X})X_j^{\mu}(k)+\gamma_i^{\mu}(k)+w_i^{\mu}(k), \end{aligned} \quad (11)$$

with $X_i^{\mu}(0)=X_0^{[i]}$. $\{u_i^{\mu}(k)\}$, $\{\gamma_i^{\mu}(k)\}$ are deterministic sequences while $\{v_i^{\mu}\}$, $\{w_i^{\mu}\}$ are stochastic sequences, all depending on $\widetilde{X}$. The 2μ equations (11) of the Carleman approximation of system (2) can be put in the following compact form

$$\begin{aligned} X^{\mu}(k+1) &= \mathcal{A}^{\mu}(\widetilde{X})X^{\mu}(k)+\mathcal{U}^{\mu}(k)+V^{\mu}(k), \\ Y^{\mu}(k) &= \mathcal{C}^{\mu}(\widetilde{X})X^{\mu}(k)+\Gamma^{\mu}(k)+W^{\mu}(k); \end{aligned} \quad (12)$$

the augmented state $X^{\mu}(k)\in I\!R^{n_{\mu}}$, $n_{\mu}=\sum_{j=1}^{\mu} n_e^j$, the augmented output $Y^{\mu}(k)\in I\!R^{q_{\mu}}$, $q_{\mu}=\sum_{j=1}^{\mu} q^j$, the matrices $\mathcal{A}^{\mu}(\widetilde{X})$, $\mathcal{C}^{\mu}(\widetilde{X})$, the deterministic sequences $\{\mathcal{U}^{\mu}(k)\}$, $\{\Gamma^{\mu}(k)\}$ and the augmented noises $\{V^{\mu}(k)\}$, $\{W^{\mu}(k)\}$ easily come as the aggregate of the terms in (11) for $i=1,\ldots,\mu$.

According to eq.'s (A.7-8) in Appendix, the noises $V^{\mu}(k)$ and $W^{\mu}(k)$ are bilinear functions of the augmented state $X^{\mu}(k)$ and of zero-mean random vectors uncorrelated with $X^{\mu}(k)$ of the type $(v_e^{[h]}(k)-E\{v_e^{[h]}(k)\})$ and $(w^{[h]}(k)-E\{w^{[h]}(k)\})$ (note that these are white sequences). This fact allows to state that the Carleman approximation (12) has a bilinear structure with respect to an augmented white noise sequence. Moreover, exploiting the same arguments used in [2, 3],

it is not difficult, though tedious, to prove that $\{V^\mu(k)\}$ and $\{W^\mu(k)\}$ are *uncorrelated* sequences of zero mean *uncorrelated* random vectors, and that the extended state $X^\mu(k)$ is *uncorrelated* with $W^\mu(j)$ $\forall j$ and with $V^\mu(j)$ 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 $k \leq j$).

To ensure that all random vectors in (12) ($X^\mu(k)$, $Y^\mu(k)$, $V^\mu(k)$ and $W^\mu(k)$) have finite means and covariances, it is necessary to assume that the noises and the initial extended state X_0 have finite and available moments up to order 2μ:

$$I\!E\{X_0^{[i]}\} < \infty, \quad \begin{array}{l} I\!E\{v^{[i]}(k)\} = \xi_i^v < \infty, \\ I\!E\{w^{[i]}(k)\} = \xi_i^w < \infty, \end{array} \tag{13}$$

for $i = 1, \ldots, 2\mu$. It follows that a stochastic characterization should be given also to the vector of parameters θ. A very frequent case found in the literature is the case of *interval systems* [1], [10], i.e. uncertain systems where each component of the vector of unknown parameters belongs to a known interval. In the case of bounded intervals the uniform probability density could be used. If no a priori bound is known for a parameter, a normal distribution can be used for its description, with known mean and variance.

The moments (13) are needed for the recursive computation of the covariances $\Psi^{V^\mu}(k)$ and $\Psi^{W^\mu}(k)$ of the augmented noises. The mean and covariance of the augmented state $X^\mu(k)$, also needed for the computation of $\Psi^{V^\mu}(k)$ and $\Psi^{W^\mu}(k)$, can be recursively computed using standard formulas for bilinear systems [3]. Details on the computation of the system matrices and sequences involved in (11), and of the augmented noise statistics are reported in Appendix.

III. THE FILTERING ALGORITHM

As previously mentioned the μ-degree Carleman approximation of a stochastic nonlinear system is a bilinear system driven by white noise, eq.'s (12). Then the optimal linear filter (linear w.r.t. the *augmented* measurements Y^μ) provides the optimal μ-degree polynomial filter w.r.t. the *original* measurements, and is obtained without any further approximation [3]. Since the augmented noises $\{V^\mu(k)\}$ and $\{W^\mu(k)\}$ are sequences of uncorrelated random vectors, as previously discussed, the optimal linear filter is implemented by the standard Kalman filter equations. According to the same philosophy of the 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, instead of $\widetilde{X}$. In particular, the state estimate is used for $\mathcal{A}^\mu$, $\mathcal{U}^\mu$ and Ψ^{V^μ}, while the state prediction is used for $\mathcal{C}^\mu$, Γ^μ and Ψ^{W^μ}, according to the formulas reported in the Appendix. The estimate $\widehat{X}(k)$ and prediction $\widehat{X}(k+1|k)$ of the extended state $X(k)$ are achieved by selecting the first n_e components from the estimate and prediction of the augmented state $\widehat{X}^\mu(k)$:

$$\begin{array}{l} \widehat{X}(k) = [I_{n_e} \;\; O_{n_e \times (n_\mu - n_e)}] \widehat{X}^\mu(k), \\ \widehat{X}(k+1|k) = [I_{n_e} \;\; O_{n_e \times (n_\mu - n_e)}] \widehat{X}^\mu(k+1|k); \end{array} \tag{14}$$

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

$$\hat{x}(k) = [I_n \;\; O_{n \times m}] \widehat{X}(k), \quad \hat{\theta}(k) = [O_{m \times n} \;\; I_m] \widehat{X}(k). \tag{15}$$

The steps of the PEKF are summarized below:

Polynomial Extended Kalman Filter

I) Filter initialization:

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

II) Computation of the matrices of the μ-degree approximation of the augmented output equation around $\widehat{X}(k+1|k)$ (i.e. $\mathcal{C}^\mu(k+1)$, $\Gamma^\mu(k+1)$, $\Psi^{W^\mu}(k+1)$);

III) Prediction of the augmented output:

$$\widehat{Y}^\mu(k+1|k) = \mathcal{C}^\mu(k+1) \widehat{X}^\mu(k+1|k) + \Gamma^\mu(k+1); \tag{16}$$

IV) Computation of the Kalman gain:

$$\begin{array}{l} K(k+1) = P_P(k+1) \mathcal{C}^\mu(k+1)^T \Big(\mathcal{C}^\mu(k+1) \\ \quad \cdot P_P(k+1) \mathcal{C}^\mu(k+1)^T + \Psi^{W^\mu}(k+1) \Big)^\dagger; \end{array} \tag{17}$$

V) Computation of the error covariance matrix:

$$P(k+1) = \Big(I_{n_\mu} - K(k+1) \mathcal{C}^\mu(k+1) \Big) P_P(k+1); \tag{18}$$

VI) Estimation of the augmented state:

$$\begin{array}{l} \widehat{X}^\mu(k+1) = \widehat{X}^\mu(k+1|k) + K(k+1) \\ \quad \cdot \Big(Y^\mu(k+1) - \widehat{Y}^\mu(k+1|k) \Big), \end{array} \tag{19}$$

of $\widehat{X}(k+1)$ and of the estimates $\hat{x}(k+1)$, $\hat{\theta}(k+1)$ of the original state and of the unknown parameter with (14) and (15);

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

VIII) Computation of the matrices of the μ-degree approximation of the augmented state equation around $\widehat{X}(k)$ (i.e. $\mathcal{A}^{\mu}(k)$, $\mathcal{U}^{\mu}(k)$, $\Psi^{V^{\mu}}(k)$);

IX) Prediction of the augmented state:

$$\widehat{X}^{\mu}(k+1|k) = \mathcal{A}^{\mu}(k)\widehat{X}^{\mu}(k) + \mathcal{U}^{\mu}(k); \qquad (20)$$

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

$$P_P(k+1) = \mathcal{A}^{\mu}(k)P(k)\mathcal{A}^{\mu}(k)^T + \Psi^{V^{\mu}}(k); \qquad (21)$$

XI) GOTO STEP **II**.

Remark 2: For consistency with all the developments made in the paper, the PEKF algorithm has been here presented in a form that is not computationally optimized, in that the Kronecker powers contain redundant components (if $X \in I\!R^{n_e}$ then $X^{[i]} \in I\!R^{n_e^i}$, but only $\tilde{n}_i = \binom{n_e+i-1}{i}$ monomials are independent). Such redundancies can be avoided through the definition of *reduced Kronecker powers*, containing the independent components of ordinary Kronecker powers [3]. •

IV. SIMULATION RESULTS

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

$$\begin{aligned} x_1(k+1) &= \alpha x_1(k) + x_1(k)x_2(k) + 0.1 + 0.01v_1(k), \\ x_2(k+1) &= 1.5x_2(k) - x_1(k)x_2(k) + 0.1 + 0.01v_2(k), \\ y(k) &= x_2(k) + 0.04w(k), \end{aligned} \qquad (22)$$

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

$$\begin{array}{lll} P_{v_1}(-1) = 0.6, & P_{v_2}(-1) = 0.8, & P_w(-7) = 0.3, \\ P_{v_1}(0) = 0.2, & P_{v_2}(4) = 0.2, & P_w(3) = 0.7. \\ P_{v_1}(3) = 0.2, & & \end{array} \qquad (23)$$

The initial state $x(0)$ is a Gaussian standard random vector (zero-mean, identity covariance matrix), independent of both the state and output noises.

α is the unknown parameter of system (22) to be estimated. The only a priori knowledge assumed on α is its definition interval

$$\alpha \in [-0.8, 0.8], \qquad (24)$$

that guarantees stability of the system (22). In order to develop a robust parameter estimator, such that at each time step the estimated parameter remains inside its definition interval, it is convenient to consider α as a bounded function of an auxiliary variable θ:

$$\alpha(\theta) = 0.8\frac{\theta}{\sqrt{1+\theta^2}}, \qquad \theta \in [-\theta_M, \theta_M]. \qquad (25)$$

with θ_M such that $\theta_M/\sqrt{1+\theta_M^2} \approx 1$. The estimation of θ instead of α leads to more robust filters, in that the structure of (25) is such that whatever the estimate of θ, $\alpha(\theta)$ always remains inside its definition interval (24). This reduces the possibility of instability of the filtering algorithm. In the simulations presented the true value of the parameter θ is 4.2, while the value $\theta_M = 10$.

For the purpose of providing the filtering algorithm with all the moments required, θ is assumed uniformly distributed in the interval $[-\theta_M, \theta_M]$.

Simulations are reported comparing the standard first and second order EKF [5] with the quadratic ($\mu = 2$) and cubic ($\mu = 3$) version of the proposed PEKF. The sample error variances computed in a typical simulation over a 1000 points horizon, are the following:

$$\begin{array}{rll} EKF: & \sigma^2_{x_1} = 9.15 \cdot 10^{-3} & \sigma^2_{x_2} = 1.75 \cdot 10^{-3} \\ 2\text{nd}EKF: & \sigma^2_{x_1} = 1.00 \cdot 10^{-2} & \sigma^2_{x_2} = 1.76 \cdot 10^{-3} \\ PEKF_{\mu=2}: & \sigma^2_{x_1} = 7.45 \cdot 10^{-3} & \sigma^2_{x_2} = 1.30 \cdot 10^{-3} \\ PEKF_{\mu=3}: & \sigma^2_{x_1} = 4.75 \cdot 10^{-3} & \sigma^2_{x_2} = 4.28 \cdot 10^{-4} \end{array}$$

In this example the quadratic and cubic PEKF perform better than the standard EKF and its second order version, which have a very similar behavior. In particular, the quadratic PEKF achieves 18% and 25% reduction of the error variance of the two state components w.r.t. standard EKF, while the cubic PEKF achieves 48% and 75% variance reduction. Fig.'s 1 and 2 report the true states and their estimates using the second order EKF, the quadratic and the cubic PEKF (for the clarity of the representation, only a window of 30 time steps is reported).

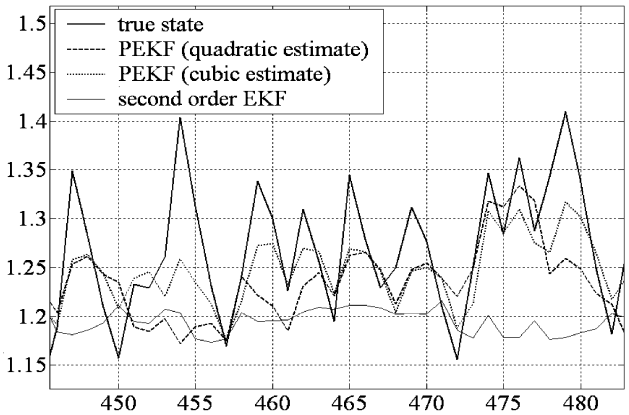

Fig. 1 – True and estimated state: x_1.

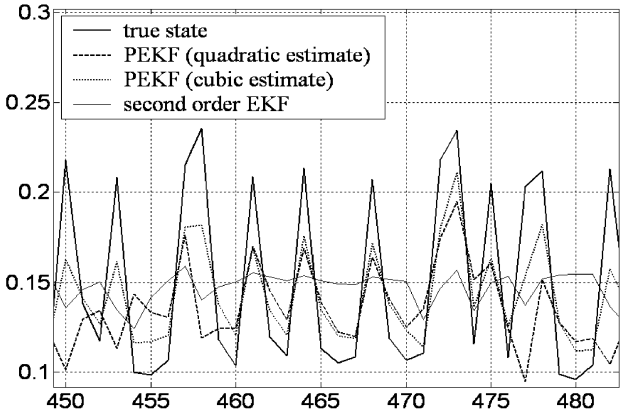

Fig. 2 – True and estimated state: x_2.

As for what concerns the parameter identification, figure 3 shows the improvements of the quadratic and cubic PEKF with respect to the 2nd order EKF over all 1.000 simulation steps.

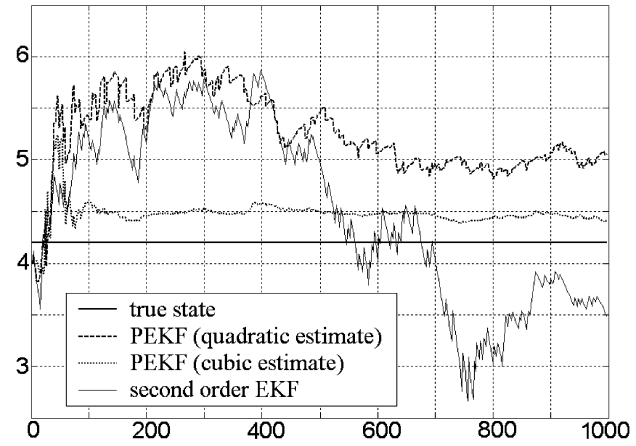

Fig. 3 – True and estimated parameter θ.

V. CONCLUSIONS

The problem of simultaneous state estimation and parameter identification for a nonlinear stochastic system has been solved by means of the Polynomial Extended Kalman Filter [6, 7], applied to the extended state containing the uncertain parameter among its components. According to the polynomial filtering approach, by increasing the order μ of the filter a better behavior is expected. Numerical simulations confirm this feature.

APPENDIX

This Appendix reports the expressions of the main terms needed for the PEKF implementation. The derivation of these equations takes advantage of a formalism that allows to expand the Kronecker powers of a sums of $\nu+1$ vectors $z_i \in \mathbb{R}^p$, $i = 0, 1, \ldots, \nu$, by using a multi-index $t = \{t_0, t_1, \cdots\ t_\nu\} \in \left(\mathbb{Z}^+\right)^{\nu+1}$:

$$\left(\sum_{i=0}^{\nu} z_i\right)^{[j]} = \sum_{|t|=j} M_t^p \boxtimes_{i=0}^{\nu} z_i^{[t_i]}, \tag{A.1}$$

with a suitable definition of the matrix coefficients $M_t^p \in \mathbb{R}^{p^j \times p^j}$ [3]. The symbol $|t|$ denotes the *modulus* of a multi-index (i.e. $|t| = t_0 + t_1 + \cdots t_\nu$) while $\boxtimes$ denotes the Kronecker products of indexed vectors [6, 7].

Lemma A.1: *The matrix $A_{ij}^{\mu}(\widetilde{X})$ in* (11) *can be computed as follows:*

$$\sum_{r \in \mathcal{R}_{ij}^{\mu}} M_r^{n_e} \mathcal{F}_r \left(M_{\alpha(r)-j,j}^{n_e} \otimes \xi_{r_{\mu+1}}^{v}\right)\left(I_{n_e^j} \otimes (-\widetilde{X})^{[\alpha(r)-j]}\right), \tag{A.2}$$

while matrix $C_{ij}^{\mu}(\widetilde{X})$ is computed as:

$$\sum_{r \in \mathcal{R}_{ij}^{\mu}} M_r^{q} \overline{H}_r \left(M_{\alpha(r)-j,j}^{n_e} \otimes \xi_{r_{\mu+1}}^{w}\right)\left(I_{n_e^j} \otimes (-\widetilde{X})^{[\alpha(r)-j]}\right), \tag{A.3}$$

with $r = \left\{r_0, \cdots, r_{\mu+1}\right\}$ a multi-index in $\left(\mathbb{Z}^+\right)^{\mu+2}$, $\alpha(r) = \sum_{s=1}^{\mu} s\, r_s$ and $\mathcal{R}_{ij}^{\mu} = \left\{r \in \left(\mathbb{Z}^+\right)^{\mu+2} : |r| = i,\ j \le \alpha(r) \le \mu\right\}$; the matrices $\mathcal{F}_r$, $\overline{H}_r$ in (A.2), (A.3) *are:*

$$\begin{aligned} \mathcal{F}_r &= \mathcal{F}_r(\widetilde{X}) = \left(\boxtimes_{s=0}^{\mu} F_{1,s}^{[r_s]}(\widetilde{X})\right) \otimes \begin{bmatrix} I_n \\ O_{m\times n} \end{bmatrix}^{[r_{\mu+1}]}, \\ \overline{H}_r &= \overline{H}_r(\widetilde{X}) = \left(\boxtimes_{s=0}^{\mu} H_{1,s}^{[r_s]}(\widetilde{X})\right) \otimes I_{q^{r_{\mu+1}}}. \end{aligned} \tag{A.4}$$

The deterministic drifts u_i^{μ}, γ_i^{μ} in (11) *are:*

$$u_i^{\mu} = \sum_{r \in \mathcal{R}_{i0}^{\mu}} M_r^{n_e} \mathcal{F}_r(\widetilde{X})\left(\widetilde{X}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^{v}\right), \tag{A.5}$$

$$\gamma_i^{\mu} = \sum_{r \in \mathcal{R}_{i0}^{\mu}} M_r^{q} \overline{H}_r(\widetilde{X})\left(\widetilde{X}^{[\alpha(r)]} \otimes \xi_{r_{\mu+1}}^{w}\right), \tag{A.6}$$

and the random sequences $\left\{v_i^{\mu}\right\}$, $\left\{w_i^{\mu}\right\}$ are:

$$\sum_{r \in \mathcal{R}_{i0}^{\mu}} \sum_{s=0}^{\alpha(r)} \Delta_{i,s}^{r}(\widetilde{X})\Big(X_s^{\mu} \otimes \left(v^{[r_{\mu+1}]} - \xi_{r_{\mu+1}}^{v}\right)\Big), \tag{A.7}$$

$$\sum_{r \in \mathcal{R}_{i0}^{\mu}} \sum_{s=0}^{\alpha(r)} \Phi_{i,s}^{r}(\widetilde{X})\Big(X_s^{\mu} \otimes \left(w^{[r_{\mu+1}]} - \xi_{r_{\mu+1}}^{w}\right)\Big), \tag{A.8}$$

respectively, with $\Delta_{i,s}^{r}(\widetilde{X})$ defined as:

$$M_r^{n_e} \mathcal{F}_r\Big(M_{\alpha(r)-s,s}^{n_e}\left(I_{n_e^s} \otimes (-\widetilde{X})^{[\alpha(r)-s]}\right) \otimes I_{n_e^{r_{\mu+1}}}\Big), \tag{A.9}$$

and $\Phi^r_{i,s}(\widetilde{X})$ *defined as:*

$$M^q_r \overline{H}_r \Big(M^{n_e}_{\alpha(r)-s,s} \big(I_{n_e^s} \otimes (-\widetilde{X})^{[\alpha(r)-s]} \big) \otimes I_{q^{r_{\mu+1}}} \Big). \tag{A.10}$$

Proof. The Carleman approximation of a generic nonlinear stochastic system has been presented in [6, 7]. Here it is reported the way to define such a derivation for the nonlinear functions f_e and h_e. According to the extended noise v_e defined in (3), from [6, 7]:

$$A^\mu_{ij}(\widetilde{X}) = \sum_{r \in \mathcal{R}^\mu_{ij}} M^{n_e}_r \overline{F}_r(\widetilde{X}) \Bigg(M^{n_e}_{\alpha(r)-j,j} \otimes \Bigg(\begin{bmatrix} I_n \\ O_{m\times n} \end{bmatrix}^{[r_{\mu+1}]} \cdot \xi^v_{r_{\mu+1}} \Bigg) \Bigg) \big(I_{n_e^j} \otimes (-\widetilde{X})^{[\alpha(r)-j]} \big), \tag{A.11}$$

with $\overline{F}_r(\widetilde{X}) = \Big(\boxtimes_{s=0}^{\mu} F^{[r_s]}_{1,s}(\widetilde{X}) \Big) \otimes I_{n_e^{r_{\mu+1}}}$. By substituting in $\overline{F}_r(\widetilde{X})$ in (A.11) and taking into account the Kronecker product properties, eq. (A.2) comes with $\mathcal{F}_r(\widetilde{X})$ as in (A.4). Analogously, eq.'s (A.5), (A.7), with (A.9), readily come. The approximate output equation, i.e. the terms involved in the eq.'s (A.3), (A.4), (A.6), (A.8), (A.10), are straightforward from [6, 7]. □

The following proposition reports the expressions of the covariance matrices Ψ^{V^μ} and Ψ^{W^μ} of the random vectors V^μ and W^μ affecting the augmented system (12). Computations, although easy to obtain, are omitted for the sake of brevity.

Proposition A.2: *Let* $\Psi^{V^\mu}_{ij}(k) = I\!E\{ v^\mu_i(k) v^\mu_j(k)^T \}$ *and* $\Psi^{W^\mu}_{ij}(k) = I\!E\{ w^\mu_i(k) w^\mu_j(k)^T \}$, *the entries of* Ψ^{V^μ} *and* Ψ^{W^μ}. *Then:*

$$\Psi^{V^\mu}_{ij}(k) = \sum_{r \in \mathcal{R}^\mu_{j0}} \sum_{t \in \mathcal{R}^\mu_{i0}} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \Delta^r_{i,s}(\widetilde{X}) \Bigg(\big(\Psi^{X^\mu}_{s,l}(k) \otimes \mathrm{st}^{-1}_{n_e^i, n_e^j} \Big(\big(\xi^v_{t_{\mu+1}+r_{\mu+1}} - \xi^v_{t_{\mu+1}} \otimes \xi^v_{r_{\mu+1}} \big) \Big) \Bigg) \Delta^t_{j,l}(\widetilde{X})^T, \tag{A.12}$$

$$\Psi^{W^\mu}_{ij}(k) = \sum_{r \in \mathcal{R}^\mu_{j0}} \sum_{t \in \mathcal{R}^\mu_{i0}} \sum_{s=0}^{\alpha(r)} \sum_{l=0}^{\alpha(t)} \Phi^r_{i,s}(\widetilde{X}) \Bigg(\big(\Psi^{X^\mu}_{s,l}(k) \otimes \mathrm{st}^{-1}_{q^i, q^j} \Big(\big(\xi^w_{t_{\mu+1}+r_{\mu+1}} - \xi^w_{t_{\mu+1}} \otimes \xi^w_{r_{\mu+1}} \big) \Big) \Bigg) \Phi^t_{j,l}(\widetilde{X})^T, \tag{A.13}$$

where $\mathrm{st}^{-1}_{a,b}$ *is the inverse of the stack operator, giving matrices in* $I\!R^{a\times b}$. $\Psi^{X^\mu}_{ij}(k) = I\!E\{ X^\mu_i(k) X^\mu_j(k)^T \}$ *are the blocks of the second order moments matrix of the augmented state, computed by the recursive equation:*

$$\Psi^{X^\mu}(k+1) = \mathcal{A}^\mu \Psi^{X^\mu}(k) \mathcal{A}^{\mu T} + \Psi^{V^\mu}(k) + \mathcal{A}^\mu Z^\mu(k) \mathcal{U}^{\mu T} + \mathcal{U}^\mu Z^\mu(k)^T \mathcal{A}^{\mu T} + \mathcal{U}^\mu \mathcal{U}^{\mu T}, \tag{A.14}$$

with $Z^\mu(k) = I\!E\{ X^\mu(k) \}$ *the mean value of the augmented state, computed as:*

$$Z^\mu(k+1) = \mathcal{A}^\mu(k) Z^\mu(k) + \mathcal{U}^\mu(k). \tag{A.15}$$

The initialization of (A.14) *and* (A.15) *are:*

$$\Psi^{X^\mu}_{ij}(0) = E\{ X^{[i]}_0 (X^{[j]}_0)^T \} = \mathrm{st}^{-1}_{n_e^i, n_e^j} \Big(I\!E\{ X^{[i+j]}_0 \} \Big), \tag{A.16}$$

with $Z^\mu_i(0) = I\!E\{ X^{[i]}_0 \}$ *given by*

$$\sum_{j=0}^{i} M^2_{j,i-j} \Bigg(\begin{bmatrix} I_n \\ O_{m\times n} \end{bmatrix}^{[j]} \otimes \begin{bmatrix} O_{n\times m} \\ I_m \end{bmatrix}^{[i-j]} \Bigg) \big(\zeta^0_j \otimes \zeta^\theta_{i-j} \big), \tag{A.17}$$

where $\zeta^0_j = I\!E\{ x^{[i]}_0 \}$ *and* $\zeta^\theta_j = I\!E\{ \theta^{[i]} \}$.

REFERENCES

[1] A.C. Bartlett, A. Tesi and A. Vicino, "Frequency response of uncertain systems with interval plants," *IEEE Trans. Automat. Contr.*, Vol. 38, No. 6, pp. 929–933, 1993.

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

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

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

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

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

[7] Germani, A., C. Manes and P. Palumbo, "Polynomial Extended Kalman Filter", *IEEE Trans. Automat. Contr.*, November 2005.

[8] C. Knapp and P. Pal, "Parameter identification in a class of nonlinear systems," *IEEE Trans. Automat. Contr.*, Vol. 28, No. 4, pp. 497–503, 1983.

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

[10] H. Sanquing and J. Wang, "On stabilization of a new class of linear time-invariant interval systems via constant state feedback control," *IEEE Trans. Automat. Contr.*, Vol. 45, No. 11, pp. 2106–2111, 2000.

[11] J. Zhao and I. Kanellakopoulos, "Active identification for discrete-time nonlinear control. I. Output-feedback systems," *IEEE Trans. Automat. Contr.*, Vol. 47, No. 2, pp. 210–224, 2002.