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

Filtering of Differential Nonlinear Systems via a Carleman Approximation Approach

44th IEEE Conf. on Decision and Control & European Control Conference (CDC–ECC 2005)

Alfredo Germani Costanzo Manes Pasquale Palumbo

***Abstract*— This paper deals with the state estimation problem for a stochastic nonlinear differential system driven by a standard Wiener process. The solution here proposed is a linear filtering algorithm and is achieved by means of the Carleman approximation scheme applied to both the state and the measurement nonlinear equations. Such a procedure allows to define an approximate representation by means of a suitable bilinear system for which a filtering algorithm is available from literature. Numerical simulations support the theoretical results and show a rather interesting improvement in terms of sampled error covariance of the proposed approach with respect to the classical Kalman-Bucy filter applied to the linearized differential system.**

***Index Terms*—Nonlinear filtering, Extended Kalman Filter, Carleman approximation, Stochastic Systems.**

I. INTRODUCTION

Given the probability triple $(\Omega, \mathcal{F}, \mathcal{P})$, in this paper it will be considered the filtering problem for the following nonlinear stochastic differential system described by the Ito equations:

$$\begin{aligned} dx_t &= \phi(x_t)dt + \overline{F}dW^s(t), \qquad && x_{|t=0} = x_0, \\ dy_t &= h(x_t)dt + \overline{G}dW^o(t), \qquad && y_{|t=0} = 0, \end{aligned} \tag{1}$$

where x_t is the state vector in $I\!R^n$, y_t is the measured output in $I\!R^q$ and $W^s(t) \in I\!R^p$, $W^o(t) \in I\!R^q$ are independent standard Wiener processes with respect to some increasing family of σ-algebras, namely $\{\mathcal{F}_t,\ t \geq 0\}$. $\phi : I\!R^n \mapsto I\!R^n$ and $h : I\!R^n \mapsto I\!R^q$ are smooth nonlinear maps.

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 [4, 20, 21, 22, 29]. Only in few cases the optimal filter has a finite dimension [27]. For this reason a great deal of work has been made to devise suboptimal implementable filtering algorithms [9, 10, 12, 15].

Another approach consists in considering the time discretization of the original system and then to apply suboptimal filtering procedures like the Extended Kalman Filter (EKF), the most widely used algorithm in nonlinear filtering problems (see, e.g., [1, 8, 11, 17]), particle filters [23], Gaussian sum approximations [16] and more. Because of its *local* nature, the EKF performs well if the initial estimation error and the disturbing noises are small enough. An effective modification of the EKF is the Unscented Kalman Filter (UKF) [18], that uses the so-called *unscented transform* for the state and output prediction steps in the EKF scheme. More recently, in [13] has been proposed a *polynomial* extension of the EKF (denoted PEKF) which is based on the application of the optimal polynomial filter of [5, 6] to the Carleman approximation of the nonlinear system (see [19, 24]).

The aim of this paper is to overcome the drawbacks of the time discretization errors by means of the use of the Carleman approximation on the original stochastic differential system. The Carleman approximation of order ν 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 ν. The result is a bilinear system (linear drift and multiplicative noise) with respect to the extended state.

The tool of Carleman bilinearization has found some applications in problems of systems approximation [25, 26, 28], since there are many reasons for finding a bilinear approximation of a nonlinear system (see [3]). In recent times such method has been successfully used in the problem of reduction of large scale systems [2].

Once the approximation is obtained, the recursive equations of the optimal linear filter for bilinear stochastic differential systems are available and can be applied with no further approximations [7]. When $\nu = 1$ the proposed filtering algorithm reduces to the classical Kalman-Bucy filter applied to the linearized differential system. With respect to this classical approach the method here proposed performs a better behavior by increasing the approximation degree of the nonlinear system.

The paper is organized as follows: the next section

This work is supported 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

presents the Carleman approximation of stochastic nonlinear differential systems of the type (1); in section three the linear minimum variance filter for the Carleman approximation is derived; section four displays some numerical results where the performances of the proposed algorithm are compared with those of the classical approach (the case of $\nu = 1$).

II. CARLEMAN APPROXIMATION

Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\{\mathcal{F}_t\}$ with $t \in [0,\ T]$ be a family of nondecreasing σ-algebras of $\mathcal{F}$ and $\big(W^s(t), \mathcal{F}_t\big)$, $\big(W^o(t), \mathcal{F}_t\big)$ be standard Wiener processes taking values in $I\!R^p$ and $I\!R^q$ respectively. Consider the stochastic nonlinear differential system described by (1), where x_t is the state vector in $I\!R^n$, y_t is the measured output in $I\!R^q$ and $\phi : I\!R^n \mapsto I\!R^n$, $h : I\!R^n \mapsto I\!R^q$ are smooth nonlinear maps. By defining the standard Wiener process $W = \begin{bmatrix} W^{sT} & W^{oT} \end{bmatrix}^T$ taking values in $I\!R^b$, $b = p + q$, system (1) may be rewritten as:

$$
\begin{aligned}
dx_t &= \phi(x_t)dt + \sum_{i=1}^{b} F_i dW_i(t), \qquad x_{|t=0} = x_0, \\
dy_t &= h(x_t)dt + \sum_{i=1}^{b} G_i dW_i(t), \qquad y_{|t=0} = 0,
\end{aligned} \tag{2}
$$

with $F_i, G_i, i = 1, \ldots, b$ columns of matrices $F = \begin{bmatrix} \bar{F} & 0 \end{bmatrix}$, $G = \begin{bmatrix} 0 & \bar{G} \end{bmatrix}$.

The initial state x_0 is an $\mathcal{F}_0$-measurable random variable, independent of W, with finite and available moments up to degree 2ν, namely:

$$
\zeta_i = I\!E\big[x_0^{[i]}\big], \qquad \|\zeta_i\| < +\infty, \qquad i = 1, \ldots, 2\nu, \tag{3}
$$

with the square brackets denoting the Kronecker powers (see [6] for a quick survey on the Kronecker product and its main properties).

Under standard analyticity hypotheses, both the state and the output equations can be written by using the Taylor polynomial expansion around a given state $\tilde{x}$. According to the Kronecker formalism, the differential system in (2) becomes:

$$
\begin{aligned}
dx_t &= \sum_{i=0}^{\infty} \Phi_i(\tilde{x})\big(x_t - \tilde{x}\big)^{[i]} dt + \sum_{j=1}^{b} F_j dW_j(t), \\
dy_t &= \sum_{i=0}^{\infty} H_i(\tilde{x})\big(x_t - \tilde{x}\big)^{[i]} dt + \sum_{j=1}^{b} G_j dW_j(t),
\end{aligned} \tag{4}
$$

with:

$$
\Phi_i(x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes \phi\right), \quad H_i(x) = \frac{1}{i!}\left(\nabla_x^{[i]} \otimes h\right). \tag{5}
$$

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

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

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 idea of the paper is to use the ν-degree Carleman approximation applied to system (2). Such an approach allows to obtain a bilinear differential system, which is then filtered according to corresponding optimal linear filter [7]. For the reader's convenience, below are reported some useful propositions, concerning the Kronecker product properties and the differential operator (6). Recall that the Kronecker product is not commutative: given a pair of integers (a, b), the symbol $C_{a,b}$ denotes a *commutation matrix*, that is a matrix in $\{0, 1\}^{ab \times ab}$ such that, given any two matrices $A \in I\!R^{r_a \times c_a}$ and $B \in I\!R^{r_b \times c_b}$

$$
B \otimes A = C_{r_a, r_b}^T (A \otimes B) C_{c_a, c_b}, \tag{7}
$$

where C_{r_a,r_b}, C_{c_a,c_b} are defined so that, denoted $[C_{u,v}]_{h,l}$ their (h, l) entries:

$$
[C_{u,v}]_{h,l} = \begin{cases} 1, & \text{if } l = (|h-1|_v)u + \left(\left[\dfrac{h-1}{v}\right] + 1\right); \\ 0, & \text{otherwise.} \end{cases} \tag{8}
$$

In the following the symbol I_n will denote the identity matrix of order n.

Proposition 1: [7, Lem. 5.1] *For any integer $h \geq 1$ and $x \in I\!R^n$, it results that:*

$$
\nabla_x \otimes x^{[h]} = U_n^h\big(I_n \otimes x^{[h-1]}\big), \tag{9}
$$

and, for any $h > 1$:

$$
\nabla_x^{[2]} \otimes x^{[h]} = O_n^h\big(I_{n^2} \otimes x^{[h-2]}\big), \tag{10}
$$

where:

$$
\begin{aligned}
U_n^h &= \sum_{\tau=0}^{h-1} C_{n,n^{h-1-\tau}}^T \otimes I_{n^\tau}, \\
O_n^h &= \sum_{\tau=0}^{h-1} \sum_{s=0}^{h-2} \big(C_{n,n^{h-1-\tau}}^T \otimes I_{n^\tau}\big)\big(I_n \otimes C_{n,n^{h-2-s}}^T \otimes I_n\big).
\end{aligned} \tag{11}
$$

Lemma 2: *Consider system (2) with the nonlinear maps ϕ, h analytical, so that the power expansion around a generic point $\tilde{x}$ in (4) may be used to describe the system. Then, for $k \geq 2$, the differential of the k-th order Kronecker power of the state x_t is given by:*

$$
d\big(x_t^{[k]}\big) = \sum_{r=0}^{\infty} A_r^{kn}(\tilde{x})\big(x_t - \tilde{x}\big)^{[r]} dt + \sum_{j=1}^{b} \mathbf{B}_{jk} x_t^{[k-1]} dW_j(t) \tag{12}
$$

with

$$A_r^{kn}(\tilde{x}) = \sum_{i=(r-k+1)\vee 0}^{r} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big) \cdot \big(I_{n^i} \otimes M_{r-i}^{k-1}\big)\big(I_{n^r} \otimes \tilde{x}^{[k-r+i-1]}\big), \tag{13}$$

and

$$\mathbf{B}_{jk} = U_n^k\big(F_j \otimes I_{n^{k-1}}\big). \tag{14}$$

Proof. Define $F_0 = \sum_{j=1}^b F_j^{[2]}$. According to the vector Ito formula [7, Thm. 5.2]:

$$d\big(x_t^{[k]}\big) = \Big(\nabla_x \otimes x^{[k]}\Big)\Big|_{x=x_t} dx_t + \frac{1}{2}\left(\nabla_x^{[2]} \otimes x^{[k]}\right)\Big|_{x=x_t} \big(d\Lambda(t)\big)^{[2]}, \tag{15}$$

with $\Lambda(t)$ the following martingale

$$\Lambda(t) = \int_0^t \sum_{j=1}^b F_j dW_j(\tau). \tag{16}$$

Then, thanks to Proposition 1:

$$\begin{aligned} d\big(x_t^{[k]}\big) &= U_n^k\big(I_n \otimes x_t^{[k-1]}\big)dx_t + \frac{1}{2}O_n^k\big(I_{n^2} \otimes x_t^{[k-2]}\big)F_0 dt \\ &= U_n^k\big(I_n \otimes x_t^{[k-1]}\big)\left(\phi(x_t)dt + \sum_{j=1}^b F_j dW_j(t)\right) \\ &\quad + \frac{1}{2}O_n^k\big(I_{n^2} \otimes x_t^{[k-2]}\big)F_0 dt \\ &= U_n^k\big(I_n \otimes x_t^{[k-1]}\big)\phi(x_t)dt + \frac{1}{2}O_n^k\big(I_{n^2} \otimes x_t^{[k-2]}\big)F_0 dt \\ &\quad + \sum_{j=1}^b U_n^k\big(I_n \otimes x_t^{[k-1]}\big)F_j dW_j(t), \end{aligned} \tag{17}$$

see also the proof of Theorem 6.1, reference [7] for more details. By using the following property of the Kronecker product:

$$(A \cdot B) \otimes (C \cdot D) = (A \otimes C) \cdot (B \otimes D), \tag{18}$$

which holds for any four matrices A, B, C, D suitably dimensioned according to the standard matricial product, the bilinear noise corresponding to the last term in (17) can be written as:

$$\begin{aligned} &\sum_{j=1}^b U_n^k\big(I_n \otimes x_t^{[k-1]}\big)(F_j \otimes 1)dW_j(t) \\ &\quad = \sum_{j=1}^b U_n^k\big(F_j \otimes x_t^{[k-1]}\big)dW_j(t) \\ &\quad = \sum_{j=1}^b U_n^k\big((F_j \cdot 1) \otimes (I_{n^{k-1}} \cdot x_t^{[k-1]})\big)dW_j(t) \\ &\quad = \sum_{j=1}^b \mathbf{B}_{jk} x_t^{[k-1]} dW_j(t) \end{aligned} \tag{19}$$

with $\mathbf{B}_{jk}$ as in (14). Then, taking into account the power expansion of the nonlinear map ϕ and applying (18), the first term in (17) becomes:

$$\begin{aligned} &\sum_{i=0}^{\infty} U_n^k\big(I_n \otimes x_t^{[k-1]}\big)\Phi_i(\tilde{x})\big(x_t - \tilde{x}\big)^{[i]} dt \\ &\quad = \sum_{i=0}^{\infty} U_n^k\Big(\Big(\Phi_i(\tilde{x})\big(x_t - \tilde{x}\big)^{[i]}\Big) \otimes x_t^{[k-1]}\Big)dt \\ &\quad = \sum_{i=0}^{\infty} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big)\Big(\big(x_t - \tilde{x}\big)^{[i]} \otimes x_t^{[k-1]}\Big)dt \\ &\quad = \sum_{i=0}^{\infty} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big) \\ &\qquad \cdot \Big(\big(x_t - \tilde{x}\big)^{[i]} \otimes \big((x_t - \tilde{x}) + \tilde{x}\big)^{[k-1]}\Big)dt. \end{aligned} \tag{20}$$

Recall that the Kronecker power of a binomial allows the following expansion:

$$(a+b)^{[k]} = \sum_{l=0}^{k} M_l^k\big(a^{[l]} \otimes b^{[k-l]}\big), \qquad \forall a, b \in I\!R^n, \tag{21}$$

with M_l^k suitably defined matricial coefficients in $I\!R^{n \times n}$, [6]. By applying (21) to the $(k-1)$-th Kronecker power in (20):

$$\begin{aligned} &\sum_{i=0}^{\infty} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big) \\ &\quad \cdot \Big(\big(x_t - \tilde{x}\big)^{[i]} \otimes \sum_{l=0}^{k-1} M_l^{k-1}\big((x_t - \tilde{x})^{[l]} \otimes \tilde{x}^{[k-1-l]}\big)\Big)dt \\ &= \sum_{i=0}^{\infty}\sum_{l=0}^{k-1} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big)\big(I_{n^i} \otimes M_l^{k-1}\big) \\ &\quad \cdot \Big(\big(x_t - \tilde{x}\big)^{[i+l]} \otimes \tilde{x}^{[k-1-l]}\Big)dt \\ &= \sum_{i=0}^{\infty}\sum_{l=0}^{k-1} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big)\big(I_{n^i} \otimes M_l^{k-1}\big) \\ &\quad \cdot \big(I_{n^{i+l}} \otimes \tilde{x}^{[k-1-l]}\big)\big(x_t - \tilde{x}\big)^{[i+l]} dt. \end{aligned} \tag{22}$$

According to the change of index $r = i + l$, the sums in (22) become:

$$\sum_{i=0}^{\infty}\sum_{r=i}^{i+k-1} U_n^k\big(\Phi_i(\tilde{x}) \otimes I_{n^{k-1}}\big) \cdot \big(I_{n^i} \otimes M_{r-i}^{k-1}\big)\big(I_{n^r} \otimes \tilde{x}^{[k-r+i-1]}\big)\big(x_t - \tilde{x}\big)^{[r]} dt, \tag{23}$$

so that (12) with (13) is achieved by changing the order of the sums in (23). □

By neglecting in (4) and (12) the higher order terms, greater than a chosen degree ν, the above mentioned differentials $d\big(x_t^{[k]}\big)$, $k = 1, 2, \ldots$, and dy_t are approximated as follows:

$$d\big(x_t^{[k]}\big) \simeq \sum_{r=0}^{\nu} A_r^{kn}(\tilde{x})\big(x_t - \tilde{x}\big)^{[r]} dt + \sum_{j=1}^{b} \big(\mathbf{B}_{jk} x_t^{[k-1]} + \mathbf{F}_{jk}\big) dW_j(t),$$
$$dy_t \simeq \sum_{i=0}^{\nu} H_i(\tilde{x})\big(x_t - \tilde{x}\big)^{[i]} dt + \sum_{j=1}^{b} G_j dW_j(t), \tag{24}$$

with

$$\begin{aligned} &A_r^{1n}(\tilde{x}) = \Phi_r(\tilde{x}), && \forall r = 0, \ldots, \nu, \\ &\mathbf{B}_{j1} = 0, && \forall j = 1, \ldots, b, \\ &\mathbf{F}_{j1} = F_j, \qquad \mathbf{F}_{jk} = 0, && \forall j = 1, \ldots, b, \quad \forall k > 1, \end{aligned} \tag{25}$$

The ν-degree Carleman bilinearization of the stochastic differential system (2) consists in:

i) exploiting the powers of the binomial $(x_t - \tilde{x})$ in the sums of (24);

ii) substituting in (24) the powers of the state $x_t^{[k]}$ with a vector $X_k^{\nu}(t)$ of the same dimension (recall that $x_t^{[k]} \in I\!R^{n^k}$), and considering them up to $k = \nu$;

iii) substituting in (24) the output y_t with a vector $Y^{\nu}(t) \in I\!R^q$;

iv) consider the stochastic Ito equations in (24) after items i-iii) as the bilinear generation model for the extended vector:

$$X^{\nu}(t) = \begin{bmatrix} X_1^{\nu}(t) \\ X_2^{\nu}(t) \\ \vdots \\ X_{\nu}^{\nu}(t) \end{bmatrix} \in I\!R^{n_\nu}, \qquad n_\nu = \sum_{i=1}^{\nu} n^i. \tag{26}$$

The first step is achieved according to the Kronecker binomial formula (21):

$$\begin{aligned} \sum_{r=0}^{\nu} & A_r^{kn}(\tilde{x})\big(x_t - \tilde{x}\big)^{[r]} dt \\ &= \sum_{r=0}^{\nu}\sum_{j=0}^{r} A_r^{kn}(\tilde{x}) M_j^r \big(x_t^{[j]} \otimes (-\tilde{x})^{[r-j]}\big) dt \\ &= \sum_{j=0}^{\nu}\sum_{r=j}^{\nu} A_r^{kn}(\tilde{x}) M_j^r \big(I_{n^j} \otimes (-\tilde{x})^{[r-j]}\big) x_t^{[j]} dt \\ &= \sum_{j=0}^{\nu} \mathbf{A}_j^k(\tilde{x}) x_t^{[j]} dt \end{aligned} \tag{27}$$

with

$$\mathbf{A}_j^k(\tilde{x}) = \sum_{r=j}^{\nu} A_r^{kn}(\tilde{x}) M_j^r \big(I_{n^j} \otimes (-\tilde{x})^{[r-j]}\big). \tag{28}$$

Analogously, it comes that:

$$\sum_{i=0}^{\nu} H_i(\tilde{x})\big(x_t - \tilde{x}\big)^{[i]} dt = \sum_{j=0}^{\nu} \mathbf{C}_j(\tilde{x}) x_t^{[j]} dt, \tag{29}$$

with

$$\mathbf{C}_j(\tilde{x}) = \sum_{i=j}^{\nu} H_i(\tilde{x}) M_j^i \big(I_{n^j} \otimes (-\tilde{x})^{[i-j]}\big). \tag{30}$$

Then, according to the items previously mentioned, the generation model for the pair (X^{ν}, Y^{ν}) is given by:

$$\begin{aligned} dX^{\nu}(t) &= \mathbf{A}^{\nu}(\tilde{x}) X^{\nu}(t) dt + \mathbf{N}^{\nu}(\tilde{x}) dt \\ &\quad + \sum_{j=1}^{b} \big(\mathbf{B}_j^{\nu} X^{\nu}(t) + \mathbf{F}_j^{\nu}\big) dW_j(t), \quad X^{\nu}(0) = X_0^{\nu}, \\ dY^{\nu}(t) &= \mathbf{C}^{\nu}(\tilde{x}) X^{\nu}(t) dt + \mathbf{D}^{\nu}(\tilde{x}) dt + \sum_{j=1}^{b} \mathbf{G}_j^{\nu} dW_j(t), \end{aligned} \tag{31}$$

with $X_0^{\nu} = \big(x_0^T \ \cdots \ x_0^{\nu T}\big)^T$, $x_0^k = x_0^{[k]}$, $k = 1, \ldots, \nu$, and:

$$\mathbf{A}^{\nu} = \begin{bmatrix} \mathbf{A}_1^1 & \cdots & \mathbf{A}_{\nu}^1 \\ \vdots & \ddots & \vdots \\ \mathbf{A}_1^{\nu} & \cdots & \mathbf{A}_{\nu}^{\nu} \end{bmatrix}, \quad \mathbf{N}^{\nu} = \begin{bmatrix} \mathbf{A}_0^1 \\ \vdots \\ \mathbf{A}_0^{\nu} \end{bmatrix}, \tag{32}$$

$$\mathbf{C}^{\nu} = [\, \mathbf{C}_1 \quad \cdots \quad \mathbf{C}_{\nu} \,], \qquad \mathbf{D}^{\nu} = \mathbf{C}_0, \tag{33}$$

$$\mathbf{B}_j^{\nu} = \begin{bmatrix} \mathbf{B}_{j1} \\ \vdots \\ \mathbf{B}_{j\nu} \end{bmatrix}, \qquad \mathbf{F}_j^{\nu} = \begin{bmatrix} \mathbf{F}_{j1} \\ \vdots \\ \mathbf{F}_{j\nu} \end{bmatrix}, \qquad \mathbf{G}_j^{\nu} = G_j. \tag{34}$$

III. THE FILTERING ALGORITHM

In the following it will be assumed that $\mathbf{D}^{\nu} = 0$. It is not a loss of generality, in that it can always be defined an auxiliary output $\widetilde{Y}^{\nu}$ such that:

$$d\widetilde{Y}^{\nu}(t) = dY^{\nu}(t) - \mathbf{D}^{\nu}(\tilde{x}) dt. \tag{35}$$

The state x_t is estimated as a linear transformation of the extended state estimate $\widehat{X}^{\nu}(t)$ described by (31), that is:

$$\hat{x}_t^{\nu} = \big[I_n \quad O_{n \times (n_\nu - n)}\big] \widehat{X}^{\nu}(t). \tag{36}$$

As is well known, the optimal choice for $\widehat{X}^{\nu}(t)$ would be the conditional expectation w.r.t. all the Borel transformations of the measurements:

$$\widehat{X}^{\nu}(t) = I\!E\big[X^{\nu}(t) | \mathcal{F}_t^{Y^{\nu}}\big], \tag{37}$$

whose computation in general can not be obtained through algorithms of finite dimension. Nevertheless, from an applicative point of view, it is useful to look for finite-dimensional approximations of the optimal filter, that is estimates described by stochastic differential equations of the form:

$$\begin{aligned} d\xi_t &= \alpha(\xi_t)dt + \beta(\xi_t)dY^\nu(t), \\ \widehat{X}^\nu(t) &= \gamma(\xi_t), \end{aligned} \tag{38}$$

where $\{\xi_t,\ t \geq t_0\}$ is a process taking its values on a finite-dimensional space. The finite-dimensional system (38) is an optimal linear filter for the random process $X^\nu(t)$ if:

$$\widehat{X}^\nu(t) = \Pi\big[X^\nu(t)|L(Y_t^\nu)\big], \tag{39}$$

where $\Pi\big[\cdot|L(Y_t^\nu)\big]$ denotes the projection onto the space $L(Y_t^\nu)$ linearly spanned by the family of random variables $\{Y^\nu(\tau),\ t_0 \leq \tau \leq t\}$.

Theorem 3: *Let $m_X^\nu(t) = I\!E\big[X^\nu(t)\big]$ and $\Psi_X^\nu(t) = \mathrm{Cov}(X^\nu(t))$ be the mean value and the covariance matrix of $X^\nu(t)$ respectively, whose evolutions are given by the following equations:*

$$\dot{m}_X^\nu = \mathbf{A}^\nu m_X^\nu(t) + \mathbf{N}^\nu(\tilde{x}), \qquad m_X^\nu(0) = \big(\zeta_1^T \ \cdots \ \zeta_\nu^T\big)^T, \tag{40}$$

$$\begin{aligned} \dot{\Psi}_X^\nu(t) = \mathbf{A}^\nu(\tilde{x})\Psi_X^\nu(t) + \Psi_X^\nu \mathbf{A}^{\nu T}(\tilde{x}) + \sum_{i=1}^b \mathbf{B}_i^\nu \Psi_X^\nu(t) \mathbf{B}_i^{\nu T} \\ + \sum_{i=1}^b (\mathbf{B}_i^\nu m_X^\nu(t) + \mathbf{F}_i)(\mathbf{B}_i^\nu m_X^\nu(t) + \mathbf{F}_i)^T, \end{aligned} \tag{41}$$

with $\Psi_X^\nu(0) = \mathrm{Cov}(X_0^\nu)$. Assume that $\mathrm{rank}(G) = q$. Then, the optimal linear estimate of the state process $X^\nu(t)$, namely $\widehat{X}^\nu(t)$, is given by:

$$\begin{aligned} d\widehat{X}^\nu(t) = \mathbf{A}^\nu(\tilde{x})\widehat{X}^\nu(t)dt + \Bigg(\sum_{i=1}^b (\mathbf{B}_i^\nu m_X^\nu(t) + \mathbf{F}_i)\mathbf{G}_i^T \\ + P(t)\mathbf{C}^{\nu T}(\tilde{x})\Bigg) R^{-1}\big(dY^\nu(t) - \mathbf{C}^\nu(\tilde{x})\widehat{X}^\nu(t)\big)dt \end{aligned} \tag{42}$$

with $R = \sum_{i=1}^b \mathbf{G}_i \mathbf{G}_i^T$ and $P(t)$ the error covariance matrix:

$$P(t) = I\!E\Big[\big(X^\nu(t) - \widehat{X}^\nu(t)\big)\big(X^\nu(t) - \widehat{X}^\nu(t)\big)^T\Big] \tag{43}$$

evolving according to the following equation

$$\begin{aligned} \dot{P}(t) &= \mathbf{A}^\nu(\tilde{x})P(t) + P(t)\mathbf{A}^{\nu T}(\tilde{x}) + R \\ &- \left(\sum_{i=1}^p (\mathbf{B}_i^\nu m_X^\nu(t) + \mathbf{F}_i)\mathbf{G}_i^T + P(t)\mathbf{C}^{\nu T}(\tilde{x})\right) \\ &\cdot R^{-1}\left(\sum_{i=1}^p (\mathbf{B}_i^\nu m_X^\nu(t) + \mathbf{F}_i)\mathbf{G}_i^T + P(t)\mathbf{C}^{\nu T}(\tilde{x})\right)^T \end{aligned} \tag{44}$$

with $P(0) = \Psi_X^\nu(0)$.

Proof. The proof is a straightforward consequence of [7, Thm. 4.4]. □

IV. SIMULATION RESULTS

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

$$\begin{aligned} dx_1(t) &= \big(-x_1(t) + x_1(t)x_2(t)\big)dt + adW_t, \\ dx_2(t) &= \big(-2x_2(t) - 2x_1(t)x_2(t)\big)dt + bdW_t, \\ dy(t) &= \big(x_1(t) - x_1(t)x_2(k)\big)dt + \gamma dW_t, \end{aligned} \tag{45}$$

with $a = 4$, $b = 1$, $\gamma = 2$. The initial state x_0 is a Gaussian random variable, with zero mean and unitary variance. x_0 has been chosen equal to $[0.1 \ \ 0.2]^T$. Numerical simulations are obtained through the construction of an exact stochastic realization of system (45) at discrete times $t_k = k\Delta$, with $\Delta = 0.1$ on the interval $[0, T]$ with $T = 100$, according to the Euler-Maruyama method [14].

In the following plots, the estimates obtained with the proposed filtering algorithm with $\nu = 2$ are compared to those obtained by applying the classical Kalman-Bucy filter to the linearized differential system ($\nu = 1$).

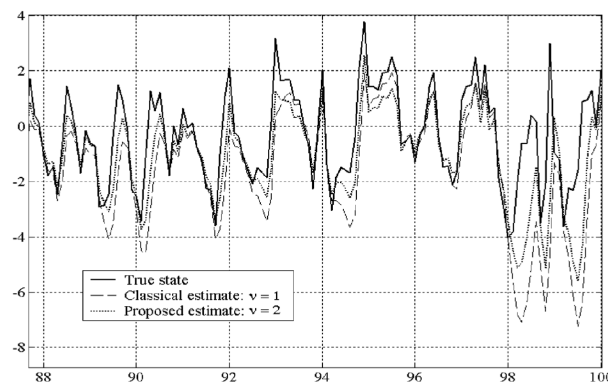

Fig. 4.1 – True and estimated state: x_1.

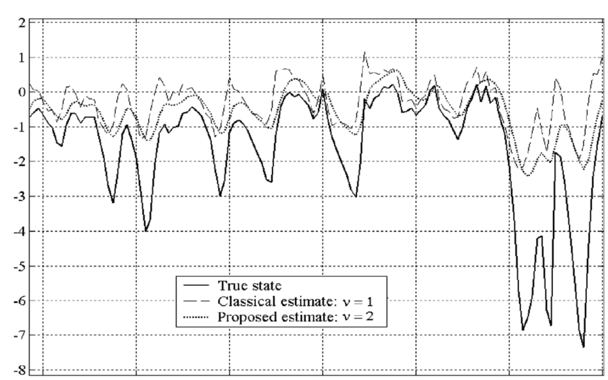

Fig. 4.2 – True and estimated state: x_2.

The improvements obtained by increasing the index ν can be recognized by comparing the sampling variances of the estimation errors:

$$\begin{aligned} \sigma_1^2(\nu=1) &= 4.5592, \qquad \sigma_1^2(\nu=2) = 2.5684, \\ \sigma_2^2(\nu=1) &= 3.2081, \qquad \sigma_2^2(\nu=2) = 2.4942. \end{aligned} \tag{46}$$

V. CONCLUSIONS

The problem of state estimation for a nonlinear differential system driven by a standard Wiener process has been investigated in this paper. The filtering algorithm here proposed is based on two steps: first the nonlinear system is approximated by using the Carleman bilinearization approach, taking into account all the powers of the series expansion up to a fixed degree ν; next, the optimal linear filter of the approximating system is achieved. This step is based on a well known literature concerning suboptimal estimates for bilinear state space representations [5, 6]. When the index $\nu = 1$, the proposed algorithm gives back the classical Kalman-Bucy filter applied to the linearized differential system.

REFERENCES

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

[2] Z. Bai, "Krylov Subspace Techniques for Reduced-Order Modeling of Large-Scale Dynamical Systems," *Applied Numerical Mathematics*, Vol. 43, pp. 9-44, 2002.

[3] C. Bruni, G. Di Pillo and G. Koch, "Bilinear systems: an appealing class of "nearly linear" systems in theory and applications," *IEEE Trans. Autom. Contr.*, Vol. 19, No.4, pp. 334–348, 1974.

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

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

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

[7] F. Carravetta, A. Germani, and M. K. Shuakayev, "A new suboptimal approach to the filtering problem for bilinear stochastic differential systems," *SIAM J. Control and Optim.*, Vol. 38, No. 4, pp. 1171–1203, 2000.

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

[9] G. B .Di Masi, M. Pratelli, W. J. Runggaldier, "An Approximation for the Nonlinear Filtering Problem, with Error Bound," *Stochastics*, Vol. 14, , pp.247–271, 1985.

[10] R. J. Elliot and R. Glowinski, "Approximations to solutions of the Zakai filtering equation," *Stoch. Anal. Appl.*, Vol. 7, No. 2, pp. 145–168, 1989.

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

[12] A. Germani and M. Piccioni, "Finite dimensional approximation for the equations of nonlinear filtering derived in mild form," *Applied Math. and Optim.*, Vol. 16, pp. 51–72, 1987.

[13] A. Germani, C. Manes and P. Palumbo, "Polynomial extended Kalman filtering for discrete-time nonlinear stochastic systems," *42nd IEEE Conf. on Decision and Control*, pp. 886–891, Maui, Hawaii, 2003.

[14] D. J. Higham, "An algorithmic introduction to numerical simulation of stochastic differential equations," *SIAM Review*, Vol. 43, No. 3, pp. 525–546, 2001.

[15] K. Ito, "Approximation of the Zakai equation for nonlinear filtering theory," *SIAM J. Control Optim.*, Vol. 34, No. 2, pp. 620–634, 1996.

[16] K. Ito and K. Xiong, "Gaussian filters for nonlinear filtering problems," *IEEE Trans. Autom. Contr.*, Vol. 45, pp. 910–927, 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. of IEEE*, Vol. 92, pp. 401–422, 2004.

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

[20] H. J. Kushner, "Nonlinear filtering: the exact dynamical equations satisfied by the conditional models," *IEEE Trans. Autom. Contr.*, Vol. 12, No. 3, pp. 262–267, 1967.

[21] R. S. Liptser and A. N. Shiryayev, *Statistics of Random Processes I and II*, Berlin, Springer, 1977.

[22] E. Pardoux, "Stochastic partial differential equations and filtering of diffusion processes," *Stochastics.*, Vol. 3, pp. 127–167, 1979.

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

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

[25] S. Svoronos, G. Stephanopoulos, and R. Aris, "Bilinear approximation of general nonlinear dynamic systems with linear inputs," *International Journal of Control*, Vol. 31(1), pp. 109-126, 1980.

[26] J. Ting-Ho Lo, "Global bilinearization of systems with control appearing linearly," *SIAM J. Control*, Vol. 13, No. 4, pp. 879–885, 1975.

[27] W. S. Wong and S. T. Yau, "The estimation algebra of nonlinear filtering systems," *Mathematical Control Theory*, J. Baillieul and J. C. Willems eds., Springer Verlag, pp. 33–65, 1999.

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

[29] M. Zakai, *On the optimal filtering of diffusion processes*, Z. Wahrschein. Verw. Geb., 11, pp. 230–243, 1969.