Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

A Carleman discretization approach to filter nonlinear stochastic systems with sampled measurements

Filippo Cacace * **Valerio Cusimano** ** **Alfredo Germani** *** **Pasquale Palumbo** **

* *Università Campus Bio-Medico di Roma, Roma, Italy (e-mail: f.cacace@unicampus.it).*
** *IASI-CNR (e-mail: valerio.cusimano@iasi.cnr.it, pasquale.palumbo@iasi.cnr.it)*
*** *Dipartimento di Ingegneria Elettrica e dell'Informazione, Università degli Studi dell'Aquila, Poggio di Roio, 67040 L'Aquila, Italy (e-mail: germani@ing.univaq.it)*

Abstract:
The state estimation problem, here investigated, regards a class of nonlinear stochastic systems, characterized by having the state model described through stochastic differential equations meanwhile the measurements are sampled in discrete times. This kind of model (continuous-discrete system) is widely used in different frameworks (i.e. tracking, finance and systems biology). The proposed methodology is based on a proper discretization of the stochastic nonlinear system, achieved by means of a Carleman linearization approach. The result is a bilinear discrete-time system (i.e. linear drift and multiplicative noise), to which the Kalman Filter equations (or the Extended Kalman Filter equations in case of nonlinear measurements) can be applied. Because the approximation scheme is parameterized by a couple of indexes, related to the degree of approximation with respect to the deterministic and the stochastic terms, in the numerical simulations, different approximation orders have been used in comparison with standard methodologies. The obtained results encourage the use of the proposed approach.

Keywords: Nonlinear Filtering, Stochastic Systems, Nonlinear Systems, Kalman Filtering, Carleman Approximation

1. INTRODUCTION

Consider the following stochastic nonlinear differential system endowed with sampled measurements,

$$\begin{aligned} dx_t &= f(x_t)dt + \sum_{i=1}^{s} F_i dW_{i,t}, \qquad t \geqslant 0 \\ y_{i\Delta} &= h(x_{i\Delta}) + GN_i, \qquad i = 0, 1, \ldots \end{aligned} \tag{1}$$

where $x_t \in \mathbb{R}^n$ is the state vector, $y_{i\Delta} \in \mathbb{R}^q, i = 0, 1, \ldots$ are the sampled measurements at time $t = i\Delta$, $W_{i,t} \in \mathbb{R}$ are a set of pairwise independent standard Wiener processes with respect to a family of increasing σ-algebras $\{\mathcal{F}_t,\ t \geq 0\}$, $N_i \in \mathbb{R}^q, i = 0, 1, \ldots$ is a sequence of zero-mean independent random vectors, with identity covariance matrix, independent of the state noise. $f : \mathbb{R}^n \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}^q$ are analytical nonlinear maps. The initial state $x_0 = \bar{x}$ is an $\mathcal{F}_0$-measurable random vector, independent of both W_t and N_i. In order to avoid singular filtering problems, see Bucy and Jonckheere [1989], the standard assumption $\text{rank}(GG^T) = q$ is used.

A continuous-discrete modeling approach is necessary in many realistic frameworks where the model under investigation is required to satisfy the underlying physics of a dynamical system whilst measurements are constrained to be discrete. Filtering problems in such a continuous-discrete domain occur in a wide range of frameworks, including tracking [Teixeira et al., 2008], finance [Oksendal, 2003] and systems biology [Hartmann et al., 2012].

It is well known that the solution to the optimal filtering problem (in terms of the minimum variance estimate) is a difficult infinite-dimensional problem for nonlinear systems and, in general cases, it does not admit an implementable solution (see, e.g. Liptser and Shiryayev [1977]). Only in few cases does the optimal filter have a finite dimension (see Wong and Yau [1999] Basin and Calderon-Alvarez [2009]). One way to approach nonlinear filtering problems, in the general cases, is to approximate the original system by means of a simpler mathematical structure for which finite-dimensional, implementable state estimate algorithms are available. This is the case, for instance, of the Extended Kalman Filter (EKF), that makes use of the linearization of the original nonlinear system, or of the more recent polynomial extensions of the EKF, that suitably exploit higher order degree Carleman approximation (see, e.g. Germani et al. [2005, 2007] for both the discrete and continuous frameworks).

For the continuous-discrete scenario here considered, the EKF-based algorithms available in the literature (see, e.g. Jazwinski [1970], or the more recent Hartmann et al.

[2012], Jorgensen et al. [2007]) make use of the free evolution of the system (the mean value differential system):

$$\dot{\hat{x}}_t = f(\hat{x}_t), \qquad \hat{x}_{t=k\Delta} = \hat{x}_{k|k} \tag{2}$$

to compute the state prediction $\hat{x}_{k+1|k}$ at the endpoint of the time interval between a couple of consecutive measurements $[k\Delta, (k+1)\Delta)$, starting from the estimate $\hat{x}_{k|k}$, whilst the one-step prediction error covariance matrix $P_{k+1|k}$ is computed at the endpoint of the same interval as the evolution of the continuous-time Riccati equation

$$\dot{P}_t = A(t)P_t + P_tA(t)^T + \sum_{i=1}^{s} F_iF_i^T, \tag{3}$$

starting from $P_{t=k\Delta}$ equal to the error covariance matrix, with matrix $A(t)$ given by the Jacobian matrix of $f(\cdot)$ computed in $\hat{x}_t$. Then the state estimate equations are updated (as well as the error covariance matrix and the Kalman gain) according to the optimal linear filter equations applied to the first-order approximation of the system equations. This note proposes the use of the polynomial Carleman approximation scheme (see, e.g., Kowalski and Steeb [1991]) to compute a discrete-time approximation of the continuous-time stochastic nonlinear system, with the aim to achieve a more accurate prediction step in the filtering algorithm. Indeed, it has been shown that, in a deterministic framework, the Carleman linearization technique provides the embedding of the original nonlinear system into an infinite-dimensional linear system, whose discretization can be achieved and easily implemented with a finer and finer degree of precision, Cacace et al. [2011]. The same idea is applied to the stochastic case here investigated, providing a discrete-time bilinear system, i.e. linear drift and multiplicative noise, whose system matrices will be properly exploited to build up the filter equations (the prediction step, actually).

The same continuous-discrete filtering problem has been recently addressed by the same authors, according to a different mixed observer-filter algorithm, Cacace et al. [2013a], in the spirit of the application of high-gain observers in the stochastic context, see e.g. Ahrens and Khalil [2009], Andrieu et al. [2009], Boizot et al. [2010], Sanfelice and Praly [2012], Cacace et al. [2013b].

2. THE CARLEMAN-BASED DISCRETIZATION

The filtering algorithm is based on the discretization of the nonlinear stochastic differential system (1), providing a discrete-time stochastic system whose state evolves according to the measurements sampling times. To this aim, consider the time interval $[k\Delta, (k+1)\Delta)$, where Δ is the measurement sampling interval. In the following $X(k)$ denotes the discrete state resembling $x_{k\Delta}$ and $\phi_t^k : [k\Delta, (k+1)\Delta) \to \mathbb{R}^n$ is defined as $\phi_t^k = x_t - X(k)$. Since the rest of the Section is devote to achieve the update equation for $X(k+1)$ within the sampling interval under investigation, in order to simplify the notation we drop the superscript k in ϕ_t^k.

According to the Taylor expansion of the nonlinear function $f(\cdot)$ around $X(k)$, it is:

$$d\phi_t = dx_t = \sum_{j=0}^{\infty} \Phi_j^1(k)\phi_t^{[j]}dt + \sum_{i=1}^{s} F_i dW_{i,t} \tag{4}$$

with

$$\Phi_j^1(k) = \left.\frac{\nabla_x^{[j]} \otimes f(x)}{j!}\right|_{x=X(k)} \in \mathbb{R}^{n\times n^j} \tag{5}$$

where the square brackets in $\phi^{[i]}$ are used for the Kronecker power of vector ϕ (that is $\phi^{[i]} = \phi \otimes \phi \otimes \cdots \otimes \phi$, repeated i times) and $\nabla_x \otimes \phi(x)$ provides the Jacobian of any vector function $\phi(x)$ (see Germani et al. [2007] and references therein for more details). Notice that $\Phi_0^1(k) = f\big(X(k)\big)$, $\Phi_1^1(k)$ is the Jacobian of f computed in $X(k)$.

According to the spirit of the Carleman approximation we aim to embed the nonlinear stochastic differential system of ϕ_t, evolving in the sampling interval $[k\Delta, (k+1)\Delta)$, into a stochastic infinite-dimensional linear system. To this aim we require to compute the differentials of $\phi_t^{[h]}$, $h > 1$, that is, according to the Ito formula and to (4):

$$\begin{aligned} d\left(\phi_t^{[h]}\right) &= \left(\nabla \otimes \phi_t^{[h]}\right) \sum_{j=0}^{\infty} \Phi_j^1(k)\phi_t^{[j]}dt \\ &\quad + \frac{1}{2}\left(\nabla^{[2]} \otimes \phi_t^{[h]}\right) F_0 dt + \left(\nabla \otimes \phi_t^{[h]}\right) \sum_{i=1}^{s} F_i dW_{i,t}, \end{aligned} \tag{6}$$

where

$$F_0 = \sum_{i=1}^{s} F_i^{[2]} \in \mathbb{R}^{n^2\times 1} \tag{7}$$

with F_i denoting the i-th column of matrix F, and

$$\begin{aligned} \nabla \otimes \phi_t^{[h]} &= U_n^h \left(I_n \otimes \phi_t^{[h-1]}\right), \quad U_n^h \in \mathbb{R}^{n^h\times n^h} \\ \nabla^{[2]} \otimes \phi_t^{[h]} &= O_n^h \left(I_{n^2} \otimes \phi_t^{[h-2]}\right), \quad O_n^h \in \mathbb{R}^{n^h\times n^h} \end{aligned} \tag{8}$$

see Germani et al. [2007] for more details on the explicit computation of matrices U_n^h and O_n^h.

Thus, by suitably exploiting (8) we have:

$$\begin{aligned} d\left(\phi_t^{[h]}\right) &= \sum_{j=0}^{\infty} U_n^h \left(I_n \otimes \phi_t^{[h-1]}\right) \left((\Phi_j^1(k)\phi_t^{[j]}) \otimes 1\right)dt \\ &\quad + \frac{1}{2} O_n^h \left(I_n \otimes \phi_t^{[h-2]}\right) (F_0 \otimes 1)dt \\ &\quad + \sum_{i=1}^{s} U_n^h \left(I_n \otimes \phi_t^{[h-1]}\right) (F_i \otimes 1)dW_{i,t} \end{aligned} \tag{9}$$

According to the following property of the Kronecker product

$$(A \otimes B) \cdot (C \otimes D) = (A \cdot C) \otimes (B \cdot D) \tag{10}$$

that holds true for suitably dimensioned matrices A, B, C, D, the following simplifications can apply

$$\begin{aligned} &U_n^h \left(I_n \otimes \phi_t^{[h-1]}\right) \left((\Phi_j^1(k)\phi_t^{[j]}) \otimes 1\right) \\ &\quad = U_n^h \left((\Phi_j^1(k)\phi_t^{[j]}) \otimes \phi_t^{[h-1]}\right) \\ &\quad = U_n^h \left((\Phi_j^1(k)\phi_t^{[j]}) \otimes \left(I_{n^{h-1}}\phi_t^{[h-1]}\right)\right) \\ &\quad = U_n^h \left(\Phi_j^1(k) \otimes I_{n^{h-1}}\right) \phi_t^{[j+h-1]} \end{aligned} \tag{11}$$

$$\begin{aligned} O_n^h \left(I_{n^2} \otimes \phi_t^{[h-2]}\right)(F_0 \otimes 1) &= O_n^h \left(F_0 \otimes \phi_t^{[h-2]}\right) \\ &= O_n^h \left((F_0 \cdot 1) \otimes \left(I_{n^{h-2}}\phi_t^{[h-2]}\right)\right) \\ &= O_n^h \left(F_0 \otimes I_{n^{h-2}}\right) \phi_t^{[h-2]} \end{aligned} \tag{12}$$

$$
\begin{aligned}
U_n^h\left(I_n \otimes \phi_t^{[h-1]}\right)(F_i \otimes 1) &= U_n^h\left(F_i \otimes \phi_t^{[h-1]}\right) \\
&= U_n^h\left((F_i \cdot 1) \otimes \left(I_{n^{h-1}}\phi_t^{[h-1]}\right)\right) \\
&= U_n^h\left(F_i \otimes I_{n^{h-1}}\right)\phi_t^{[h-1]}
\end{aligned} \tag{13}
$$

so that the differentials in (6)-(9) can be written, in a unified fashion for $h = 1, 2, \ldots$:

$$
\begin{aligned}
d\left(\phi_t^{[h]}\right) &= \sum_{j=0}^{\infty} U_n^h\left(\Phi_j^1(k) \otimes I_{n^{h-1}}\right)\phi_t^{[j+h-1]}dt \\
&\quad + \frac{1}{2}O_n^h\left(F_0 \otimes I_{n^{h-2}}\right)\phi_t^{[h-2]}dt \\
&\quad + \sum_{i=1}^{s} U_n^h\left(F_i \otimes I_{n^{h-1}}\right)\phi_t^{[h-1]}dW_{i,t} \\
&= \sum_{l=h-1}^{\infty} [A_k]_{hl}\phi_t^{[l]}dt + [L_k]_h dt \\
&\quad + \sum_{i=1}^{s}\left([B_i]_{h,h-1}\phi_t^{[h-1]} + [\mathcal{F}_i]_h\right)dW_{i,t}
\end{aligned} \tag{14}
$$

with

$$
[A_k]_{hl} = \begin{cases} \frac{1}{2}O_n^h\left(F_0 \otimes I_{n^{h-2}}\right), & l = h-2 > 1 \\ U_n^h\left(\Phi_{l-h+1}^1(k) \otimes I_{n^{h-1}}\right), & l \geq h-1 \\ 0, & \text{otherwise} \end{cases} \tag{15}
$$

$$
[L_k]_h = \begin{cases} \Phi_0^1(k), & \text{for } h = 1 \\ \frac{1}{2}O_n^2 F_0, & \text{for } h = 2 \\ 0, & \text{otherwise} \end{cases} \tag{16}
$$

$$
[B_i]_{h,h-1} = U_n^h(F_i \otimes I_{n^{h-1}}), \qquad h > 1 \tag{17}
$$

$$
[\mathcal{F}_i]_h = \begin{cases} F_i, & h = 1 \\ 0, & \text{otherwise} \end{cases} \tag{18}
$$

In order to write the infinite dimensional Carleman embedding (4)-(9) in the interval $t \in [k\Delta, (k+1)\Delta)$ in a more compact form, define the extended state $\Psi_t = \left[\phi_t^T, \phi_t^{[2]T}, \ldots\right]^T$, evolving according to the following stochastic differential system:

$$
\begin{aligned}
d\Psi_t &= A_k\Psi_t dt + L_k dt + \sum_{i=1}^{s}\left(B_i\Psi_t + G_i\right)dW_{i,t} \\
\Psi_{t=k\Delta} &= 0
\end{aligned} \tag{19}
$$

where the infinite dimensional matrices L_k, A_k, B_i and $\mathcal{F}_i$ have the block-structure

$$
L_k = \begin{bmatrix} [L_k]_1 \\ [L_k]_2 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \quad A_k = \begin{bmatrix} [A_k]_{11} & [A_k]_{12} & [A_k]_{13} & \cdots \\ [A_k]_{21} & [A_k]_{22} & [A_k]_{23} & \cdots \\ [A_k]_{31} & [A_k]_{32} & [A_k]_{33} & \cdots \\ 0 & [A_k]_{42} & [A_k]_{42} & \cdots \\ 0 & 0 & [A_k]_{53} & \cdots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}
$$
$$
\mathcal{F}_i = \begin{bmatrix} [\mathcal{F}_i]_1 \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix}, \quad B_i = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ [B_i]_{21} & 0 & 0 & \cdots \\ 0 & [B_i]_{32} & 0 & \cdots \\ 0 & 0 & [B_i]_{43} & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \tag{20}
$$

with 0 denoting null matrices of suitable size .

The integral equation associated to (19) is given by:

$$
\begin{aligned}
\Psi_t &= \int_{k\Delta}^{t} e^{A_k(t-\tau)}L_k d\tau \\
&\quad + \sum_{i=1}^{s}\int_{k\Delta}^{t} e^{A_k(t-\tau)}\left(B_i\Psi_\tau + G_i\right)dW_{i,\tau}.
\end{aligned} \tag{21}
$$

Since $\phi_t = x_t - X(k)$ we can use only the first n components of vector Ψ_t to compute the update $X(k+1) = x_{(k+1)\Delta}$. Indeed, it is $\phi_t = \mathcal{L}\Psi_t$, with $\mathcal{L} = [I_n \, 0_{n\times n^2} \, \cdots]$, therefore, from (21):

$$
X(k+1) = X(k) + \mathcal{U}_k + \mathcal{V}_k. \tag{22}
$$

with

$$
\begin{aligned}
\mathcal{U}_k &= \int_{k\Delta}^{(k+1)\Delta}\mathcal{L}e^{A_k((k+1)\Delta-\tau)}L_k d\tau = \int_0^{\Delta}\mathcal{L}e^{A_k\theta}L_k d\theta \\
&= \int_0^{\Delta}\sum_{i=0}^{\infty}\mathcal{L}\frac{A_k^i\theta^i}{i!}L_k d\theta = \sum_{i=1}^{\infty}\mathcal{L}A_k^{i-1}L_k\frac{\Delta^i}{i!}
\end{aligned} \tag{23}
$$

and

$$
\mathcal{V}_k = \sum_{i=1}^{s}\int_{k\Delta}^{(k+1)\Delta}\mathcal{L}e^{A_k((k+1)\Delta-\tau)}\left(B_i\Psi_\tau + G_i\right)dW_{i,\tau} \tag{24}
$$

Remark 1. Notice that the computation for the deterministic drift $\mathcal{U}_k$ can be done in a finer and finer way, according to the degree of tolerance required. Indeed, each term of the infinite sum in (23) is a finite dimensional matrix coming by suitably combining infinite-dimensional matrices (see Cacace et al. [2011] for more details). On the other hand, the stochastic term $\mathcal{V}_k$ is a multiplicative noise, since the components of Ψ_t are involved in the stochastic integral. In presence of a complete statistical characterization of the stochastic driving term $\mathcal{V}_k$, by taking into account all the infinite terms for the computation of $\mathcal{U}_k$ we would have an exact discretization if starting from $X(k) = x_{k\Delta}$.

As far as the sequence $\mathcal{V}_k$, $k = 0, 1, \ldots$, it is easy to verify that it is a sequence of zero-mean uncorrelated random vectors, that is $\mathbb{E}\left[\mathcal{V}_k\mathcal{V}_j^T\right] = 0$, $k \neq j$.

Lemma 2. Define $\Xi_k = \mathbb{E}\left[\mathcal{V}_k\mathcal{V}_k^T\right]$ the covariance matrix of the random sequence $\mathcal{V}_k$ in (22). Then it is:

$$
\begin{aligned}
\Xi_k &= \sum_{i=1}^{s}\int_{k\Delta}^{(k+1)\Delta}\mathcal{L}e^{A_k((k+1)\Delta-\tau)} \\
&\quad \cdot Q_i(\eta_\tau, m_\tau)e^{A_k^T((k+1)\Delta-\tau)}\mathcal{L}^T d\tau
\end{aligned} \tag{25}
$$

where

$$
\begin{aligned}
Q_i(\eta_\tau, m_\tau) &= \mathbb{E}\left[\left(B_i\Psi_\tau + \mathcal{F}_i\right)\left(B_i\Psi_\tau + \mathcal{F}_i\right)^T\right] \\
&= B_i m_\tau B_i^T + \left(B_i\eta_\tau + \mathcal{F}_i\right)\left(B_i\eta_\tau + \mathcal{F}_i\right)^T
\end{aligned} \tag{26}
$$

with

$$
\begin{aligned}
\eta_t &= \mathbb{E}\left[\Psi_t\right] \\
m_t &= \mathrm{Cov}(\Psi_t) = \mathbb{E}\left[\left(\Psi_t - \eta_t\right)\left(\Psi_t - \eta_t\right)^T\right]
\end{aligned} \tag{27}
$$

mean value and covariance matrix of Ψ_t.

Proof Eq.(25) is a straightforward consequence of the Ito isometry, since:

$$\begin{aligned}\Xi_k = &\sum_{i=1}^{s}\sum_{j=1}^{s}\mathbb{E}\Bigg[\int_{k\Delta}^{(k+1)\Delta}\int_{k\Delta}^{(k+1)\Delta} \\ &\mathcal{L}e^{A_k((k+1)\Delta-\tau_1)}\big(B_i\Psi_{\tau_1}+\mathcal{F}_i\big)\big(B_j\Psi_{\tau_2}+\mathcal{F}_j\big)^T \\ &\cdot e^{A_k^T((k+1)\Delta-\tau_2)}\mathcal{L}^T dW_{i,\tau_1}dW_{j,\tau_2}\Bigg] \\ = &\sum_{i=1}^{s}\int_{k\Delta}^{(k+1)\Delta}\mathcal{L}e^{A_k((k+1)\Delta-\tau)} \\ &\cdot\mathbb{E}\left[\big(B_i\Psi_\tau+\mathcal{F}_i\big)\big(B_i\Psi_\tau+\mathcal{F}_i\big)^T\right]e^{A_k^T((k+1)\Delta-\tau)}\mathcal{L}^T d\tau\end{aligned} \tag{28}$$

and

$$\begin{aligned}&\mathbb{E}\left[\big(B_i\Psi_\tau+\mathcal{F}_i\big)\big(B_i\Psi_\tau+\mathcal{F}_i\big)^T\right] \\ &\quad=\mathbb{E}\big[\big(B_i(\Psi_\tau-\eta_\tau)+B_i\eta_\tau+\mathcal{F}_i\big) \\ &\quad\cdot\big(B_i(\Psi_\tau-\eta_\tau)+B_i\eta_\tau+\mathcal{F}_i\big)^T\big]=Q_i\big(\eta_\tau,m_\tau\big)\end{aligned} \tag{29}$$

□

Unfortunately, Lemma 2 is of poor practical use, since the exact computation of Ξ_k (an $n\times n$ finite-dimensional matrix) passes through the computation of a pair of infinite-dimensional objects like η_t and m_t. Therefore, in order to provide a finite-dimensional approximation of both η_t and m_t, the μ-degree Carleman approximation of Ψ_t will be considered, by taking into account the differentials $d(\phi_t^{[h]})$ for $h\le\mu$ and trivially neglecting in the right-hand-side of in (14) the higher-than-μ Kronecker powers of ϕ_t. This way, by denoting $\Psi_t^{h,\mu}$ the μ-degree Carleman approximation of $\phi_t^{[h]}$, the infinite-dimensional equation (19) may be replaced by the finite-dimensional equation for $\Psi_t^\mu=\left[(\Psi_t^{1,\mu})^T,(\Psi_t^{2,\mu})^T,\dots,(\Psi_t^{\mu,\mu})^T\right]^T$:

$$d\Psi_t^\mu = A_k^\mu\Psi_t^\mu dt + L_k^\mu dt + \sum_{i=1}^{s}\left(B_i^\mu\Psi_t^\mu+\mathcal{F}_i^\mu\right)dW_{i,t} \tag{30}$$

where L_k^μ, $\mathcal{F}_i^\mu$ (and A_k^μ, B_i^μ) are provided by the first μ blocks (the first μ row-blocks and column-blocks) in (20).

The integral equation associated to (30) is:

$$\begin{aligned}\Psi_t^\mu = &\int_{k\Delta}^{t}e^{A_k^\mu(t-\tau)}L_k^\mu d\tau \\ &+\sum_{i=1}^{s}\int_{k\Delta}^{t}e^{A_k^\mu(t-\tau)}\left(B_i^\mu\Psi_\tau^\mu+\mathcal{F}_i^\mu\right)dW_{i,\tau}.\end{aligned} \tag{31}$$

according to which:

$$\eta_t^\mu=\mathbb{E}[\Psi_t^\mu]=\int_{k\Delta}^{t}e^{A_k^\mu(t-\tau)}L_k^\mu d\tau \tag{32}$$

and

$$\begin{aligned}m_t^\mu = &\mathrm{Cov}(\Psi_t^\mu)=\mathbb{E}\Bigg[\sum_{i=1}^{s}\sum_{j=1}^{s}\int_{k\Delta}^{t}\int_{k\Delta}^{t}e^{A_k^\mu(t-\tau_1)} \\ &\cdot\big(B_i^\mu\Psi_{\tau_1}^\mu+\mathcal{F}_i^\mu\big)\big(B_i^\mu\Psi_{\tau_2}^\mu+\mathcal{F}_i^\mu\big)^T \\ &\cdot e^{A_k^{\mu T}(t-\tau_2)}dW_{i,\tau_1}dW_{j,\tau_2}\Bigg] \\ = &\sum_{i=1}^{s}\int_{k\Delta}^{t}e^{A_k^\mu(t-\tau)}Q_i\big(\eta_\tau^\mu,m_\tau^\mu\big)e^{A_k^{\mu T}(t-\tau)}d\tau\end{aligned} \tag{33}$$

From an implementing point of view it is useful to achieve the differential equations associated to (32)-(33) for $t\in[k\Delta,(k+1)\Delta)$:

$$\begin{aligned}\dot\eta_t^\mu = &L_k^\mu+\int_{k\Delta}^{t}A_k^\mu e^{A_k^\mu(t-\tau)}L_k^\mu d\tau = A_k^\mu\eta_t^\mu+L_k^\mu, \\ \eta_{k\Delta}^\mu = &0.\end{aligned} \tag{34}$$

and

$$\begin{aligned}\dot m_t^\mu = &\sum_{i=1}^{s}Q_i\big(\eta_t^\mu,m_t^\mu\big) \\ &+\sum_{i=1}^{s}\int_{k\Delta}^{t}A_k^\mu e^{A_k^\mu(t-\tau)}Q_i\big(\eta_\tau^\mu,m_\tau^\mu\big)e^{A_k^{\mu T}(t-\tau)}d\tau \\ &+\sum_{i=1}^{s}\int_{k\Delta}^{t}e^{A_k^\mu(t-\tau)}Q_i\big(\eta_\tau^\mu,m_\tau^\mu\big)e^{A_k^T(t-\tau)}A_k^{\mu T}d\tau \\ = &\sum_{i=1}^{s}Q_i\big(\eta_t^\mu,m_t^\mu\big)+A_k^\mu m_t^\mu+m_t^\mu A_k^{\mu T},\quad m_{k\Delta}^\mu=0.\end{aligned} \tag{35}$$

In summary, the Carleman-based discretization of the nonlinear continuous time system (1) is formally provided by:

$$X^{\mu,\xi}(k+1)=X^{\mu,\xi}(k)+\mathcal{U}_k^\xi+\mathcal{V}_k^\mu. \tag{36}$$

where the integer ξ defines the approximation of the infinite sum in (23):

$$\mathcal{U}_k^\xi=\sum_{i=1}^{\xi}\mathcal{L}A_k^{i-1}L_k\frac{\Delta^i}{i!} \tag{37}$$

and $\{\mathcal{V}_k^\mu\}$ is a sequence of zero-mean, uncorrelated random vectors, with covariance matrix

$$\begin{aligned}\Xi_k^\mu = &\sum_{i=1}^{s}\int_{k\Delta}^{(k+1)\Delta}\mathcal{L}^\mu e^{A_k^\mu((k+1)\Delta-\tau)} \\ &\cdot Q_i\big(\eta_\tau^\mu,m_\tau^\mu\big)e^{A_k^{\mu T}((k+1)\Delta-\tau)}\mathcal{L}^{\mu T}d\tau\end{aligned} \tag{38}$$

and $\mathcal{L}^\mu=[I_n\,0_{n\times n^2}\,\cdots\,0_{n\times n^\mu}]$.

3. THE FILTERING ALGORITHM

The state estimates at the sampling times are achieved by suitably exploiting the Carleman-based discretization of the nonlinear stochastic differential system (36). Notice that the Carleman approach provides a bilinear discrete-time system (i.e. linear drift and multiplicative noise), the only nonlinearity relying in the output equation. Below follow the steps of the filtering algorithm.

Initialization at $k=0$. We set the initial conditions for the (*a priori*) prediction $\hat x_{1|0}$ and for the (*a priori*) one-step prediction covariance matrix $P_{1|0}$.

Steps of the filtering algorithm.

- *Update:* In the update phase, the current *a priori* prediction is combined with current observation information to refine the state estimate. This improved estimate is termed the *a posteriori* state estimate.
  (1) Innovation or measurement residual:
  $$\tilde y_{k\Delta}=y_{k\Delta}-h(\hat x_{k|k-1})$$
  (2) Innovation covariance:
  $$S_k=H_kP_{k|k-1}H_k^T+GG^T$$

with the observation matrix defined by the following Jacobian

$$H_k = \left. \frac{\partial h}{\partial x} \right|_{\hat{x}_{k|k-1}} \tag{39}$$

(3) Kalman gain:

$$K_k = P_{k|k-1} H_k^T S_k^{-1}$$

(4) Updated state estimate:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_{k\Delta}$$

(5) Updated covariance estimate:

$$P_{k|k} = (I_n - K_k H_k) P_{k|k-1}$$

- *Prediction:* The prediction phase suitably exploits the stochastic discretization (36) achieved by means the Carleman approximation scheme around $X^{\mu,\xi}(k) = \hat{x}_{k|k}$.
  (1) Compute $\mathcal{U}_k^{\xi}$ according to a chosen degree of precision ξ, by means of (37)
  (2) Compute Ξ_k^{μ} according to the chosen degree μ of the Carleman approximation, by means of (38) and (34)-(35)
  (3) Predicted state estimate:
  $$\hat{x}_{k+1|k} = \hat{x}_{k|k} + \mathcal{U}_k^{\xi}$$
  (4) Predicted covariance estimate:
  $$P_{k+1|k} = P_{k|k} + \Xi_k^{\mu}$$

Remark 3. Notice that the computation of Ξ_k^{μ} requires the computation of the s integrals in (38), that is achieved by means of the numerical solutions of η_t^{μ} and m_t^{μ} provided by (32)-(33).

4. SIMULATION

Numerical simulation results are here reported in order to show the effectiveness of the proposed algorithm. The example is taken from the systems biology framework, where the problem to infer information from discrete sampled measurements is particularly recurring.

The continuous-time model considered is the one that describes the HIV dynamics (see Phillips [1996]), already adopted to evaluate state estimation algorithms in the recent past (see Cacace et al. [2013a,b]), whose equations are presented below:

$$\begin{aligned} dx_{1t} &= (s - d_1 x_{1t} - \beta x_{1t} x_{3t})dt + f_1 dW_{1,t} \\ dx_{2t} &= (\beta x_{1t} x_{3t} - d_2 x_{2t})dt + f_2 dW_{2,t} \\ dx_{3t} &= (p x_{2t} - c x_{3t})dt + f_3 dW_{3,t} \\ y_{i\Delta} &= x_{1,i\Delta} + x_{2,i\Delta} + G N_i \end{aligned} \tag{40}$$

where x_{1t}, x_{2t}, x_{3t} are the target cells, infected cells and serum viral concentrations, respectively and the output represents the sum of the total cells presented in the blood ($y = x_1 + x_2$).

Notice that in the physical system all state variables are bounded and positive, thus the noise amplitude $F = \text{diag}(f_1, f_2, f_3)$ must be small enough to satisfy this constraint, as it actually happens in the simulations, which have been performed in the time interval of $[0, 100]$ minutes.

Regarding the parameters, s is the rate of the constant influx of target cells, d_1 the target cell loss, β the target cells infection, d_2 the target cells infection loss, p the viral production and c the viral clearance and were set at the values presented in Table 1 (see Phillips [1996]).

Table 1. Parameters

s	d_1	β	d_2	p	c
$1 \cdot 10^2$	$1 \cdot 10^{-3}$	$1.3 \cdot 10^{-6}$	1	$1 \cdot 10^3$	3

To effectively evaluate the performance of the proposed algorithm, we have chosen different values for the measurement noise G in the range of $10 - 100$ and the state noise $F = \eta * \text{diag}(50, 1, 1)$ with $\eta = 1, 2$. Two classes of simulations have been considered, according to a couple of different measurement sampling time: $\Delta = 0.5$ and $\Delta = 1$.

For each set of noise parameters, one hundred random realizations were run and the performance was evaluated by using the Mean Square Error (MSE) for each of the state component x_j, $j = 1, 2, 3$:

$$\text{MSE}_j = \frac{1}{N} \sum_{\nu=1}^{N} \frac{1}{M} \sum_{k=1}^{M} \sqrt{(x_{j,k} - \hat{x}_{j,k|k})^2} \tag{41}$$

where $x_{j,k}$ is the real value of component x_j at time $t = k\Delta$, $x_{j,k|k}$ is the estimated value of component X_j, $N = 100$ number of realizations and M the number of measurements samples.

As far as the indexes of the approximation scheme, ξ has been fixed equal to 10 for all the simulations. On the other hand, the filtering algorithm has been tested for $\mu = 2$ and $\mu = 3$ (2-EKF and 3-EKF for short). The comparison with the standard Extended Kalman Filter (EKF) algorithm (see the Introduction Section or Jazwinski [1970] for more details) shows the very good results of the proposed filtering scheme.

In Table 2 and 3, some numerical results for different types of noises are presented. In particular in Table 2 it can be noted that the EKF algorithm works better than the 2-EKF, regards to x_2 and x_3, reasonably due to the very small noises; however, if we increase the order of index μ, the 3-EKF definitely improves the EKF performances. By increasing the noise statistics, Table 3, it clearly appears that even also the 2-EKF improves the EKF estimates.

Finally, in Figure 1 it is presented the evolution of the concentration of healthy cells (x_1) and the corresponding estimates of 2-EKF and EKF algorithms. Notice that in the figure the simulation time has been reduced to allow to distinguish the different curves.

Table 2. Numerical results: MSE

	$G = 10$			$F = 1 * \text{diag}(50, 1, 1)$		
	$\Delta = 0.5$			$\Delta = 1$		
	x_1	x_2	x_3	x_1	x_2	x_3
2-EKF	11.1	6.4	213	13.4	7.4	246
3-EKF	10.5	5.6	185	12.2	6.1	201
EKF	22.0	6.2	197	23.2	7.2	234

5. CONCLUDING REMARKS

In summary, the algorithm here proposed has involved two different and sequential tasks: the discretization of stochastic nonlinear systems and the state estimate.

Regarding the first argument, in a deterministic framwork, the Carleman linearization technique provides the embedding of the original nonlinear system into an infinite-dimensional linear system. In the stocastic case, the result is a bilinear discrete-time system composed by a linear drift and a multiplicative noise.

The proposed algorithm is able to give a statistical characterization of the stochastic term. Increasing the order μ of the Carleman approximation, the estimate of the discrete-time state covariance matrix becomes more and more accurate. In presence of a complete statistical characterization of $\mathcal{V}_k$, by taking into account all the infinite terms for the computation of $\mathcal{U}_k$ we would have an exact discretization of the original system, if starting from $X(k) = x_{k\Delta}$.

The bilinear fashion of the obtained discrete-time system motivates the use of the optimal linear filter for bilinear systems, that requires the knowledge of the second order state noise statistics. Further improvements will be based on the application a higher order polynomial filter (see Carravetta et al. [1996]).

REFERENCES

J.H. Ahrens and H.K. Khalil. High gain observers in the presence of measurement noise: a switched-gain approach. *Automatica*, 45(4):936–943, 2009.

V. Andrieu, L. Praly, and A. Astolfi. High gain observers with updated gain and homogeneous correction term. *Automatica*, 45(2):422–428, 2009.

Table 3. Numerical results: MSE

	$G = 100$			$F = 2 * \text{diag}(50, 1, 1)$		
	$\Delta = 0.5$			$\Delta = 1$		
	x_1	x_2	x_3	x_1	x_2	x_3
2-EKF	82.2	19.5	636	77.4	20.2	667
3-EKF	81.7	18.2	592	75.1	16.8	556
EKF	129	23	743	102	22.7	742

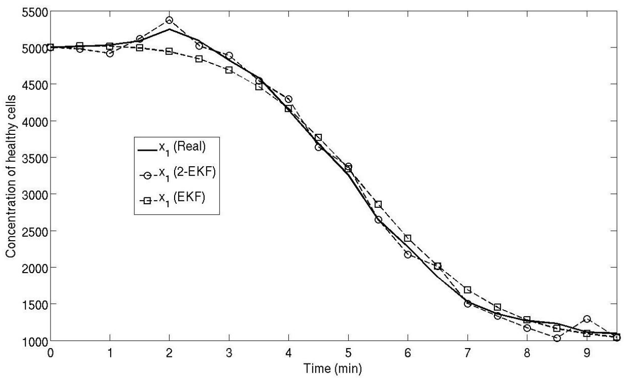

Fig. 1. Evolution of x_1 and comparison between 2-EKF and EKF algorithms

M. Basin and D. Calderon-Alvarez. Optimal filtering for incompletely measured polynomial systems with multiplicative noise. *Circuits Systems and Signal Processing*, 28:223–239, 2009.

N. Boizot, E. Busvelle, and J.P. Gauthier. An adaptive high-gain observer for nonlinear system. *Automatica*, 46(9):1483–1488, 2010.

R. Bucy and E. Jonckheere. Singular filtering problems. *Syst. Contr. Lett.*, 13:339–344, 1989.

F. Cacace, V. Cusimano, and A. Germani. An efficient approach to the design of observers for continuous-time systems with discrete-time measurements. In *Proceedings of the 52nd IEEE Conference on Decision and Control*, pages 7549–7554, 2011.

F. Cacace, V. Cusimano, A. Germani, and P. Palumbo. The observer follower filter for stochastic differential systems with sampled measurements. In *Proceedings of the 52nd IEEE Conference on Decision and Control*, 2013a.

F. Cacace, A. Germani, and P. Palumbo. The observer follower filter: A new approach to nonlinear suboptimal filtering. *Automatica*, 49:548–553, 2013b.

F. Carravetta, A. Germani, and M. Raimondi. Polynomial filtering for linear discrete time non-gaussian systems. *SIAM Control and Optimization*, 34(5):1666–1690, September 1996.

A. Germani, C. Manes, and P. Palumbo. Polynomial extended kalman filter. *IEEE Trans. on Automatic Control*, 50:2059–2064, 2005.

A. Germani, C. Manes, and P. Palumbo. Filtering of stochastic differential systems via a carleman approximation approach. *IEEE Trans. on Automatic Control*, 52:2166–2172, 2007.

Andras Hartmann, S. Vinga, and J.M. Lemos. Online bayesian time-varying parameter estimation of hiv-1 time-series. In *Preprints of the 16th IFAC Symposium on System Identification*, 2012.

A. H. Jazwinski. *Stochastic Processes and Filtering Theory.* Academic Press, 1970.

J.B. Jorgensen, P.G. Thomsen, H. Madsen, and M.R. Kristensen. A computationally efficient and robust implementation of the continuous-discrete extended kalman filter. In *Proceedings of the 2007 American Control Conference*, 2007.

K. Kowalski and W.H. Steeb. *Dynamical systems and Carleman linearization.* World Scientific, 1991.

R.S. Liptser and A.N. Shiryayev. *Statistics of random processes I and II.* Springer, Berlin, 1977.

B. Oksendal. *Stochastic Differential Equations: An Introduction with Applications.* Springer, New York, 2003.

A. N. Phillips. Reduction of hiv concentration during acute infection: independence from a specific immune response. *Science*, 271(5248):497–499, 1996.

R. G. Sanfelice and L. Praly. On the performance of high-gain observers with gain adaptation under measurement noise. *Automatica*, 47:2165–2176, 2012.

B.O.S. Teixeira, M.A. Santillo, R.S. Erwin, and D. S. Bernstein. Spacecraft tracking using sampled-data kalman filters. *IEEE Contr. Syst. Mag.*, 28(4):78–94, 2008.

W.S. Wong and S.T. Yau. *The estimation algebra of nonlinear filtering systems.* J. Baillieul & J.C. Willems, 1999.