Systems & Control Letters 98 (2016) 37–43

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

A state predictor for continuous-time stochastic systems

Filippo Cacace [a], Valerio Cusimano [b,*], Alfredo Germani [c], Pasquale Palumbo [b]

[a] Università Campus Bio-Medico di Roma, Via Álvaro del Portillo, 21, 00128 Roma, Italy
[b] IASI-CNR Viale Manzoni 30, 00185 Roma, Italy
[c] Dipartimento di Ingegneria Elettrica e dell'Informazione, Università degli Studi dell'Aquila, Via Vetoio, Coppito 67100 L'Aquila, Italy

ARTICLE INFO

Article history:
Received 29 July 2015
Received in revised form 2 May 2016
Accepted 16 October 2016

Keywords:
Nonlinear filtering
Stochastic systems
Nonlinear systems
Kalman filtering
Carleman approximation

ABSTRACT

This work investigates the state prediction problem for nonlinear stochastic differential systems, affected by multiplicative state noise. This problem is relevant in many state-estimation frameworks such as filtering of continuous-discrete systems (i.e. stochastic differential systems with discrete measurements) and time-delay systems. A very common heuristic to achieve the state prediction exploits the numerical integration of the deterministic nonlinear equation associated to the noise-free system. Unfortunately these methods provide the exact solution only for linear systems. Instead here we provide the exact state prediction for nonlinear system in terms of the series expansion of the expected value of the state conditioned to the value in a previous time instant, obtained according to the Carleman embedding technique. The truncation of the infinite series allows to compute the prediction at future times with an arbitrary approximation. Simulations support the effectiveness of the proposed state-prediction algorithm in comparison to the aforementioned heuristic method.

1. Problem formulation and background

Consider the following nonlinear stochastic differential system in the Itô formulation

$$dx_t = f(x_t)dt + \sum_{j=1}^{p} g_j(x_t)dW_{j,t} \tag{1}$$

defined on a probability space $(\Omega, \mathcal{F}, P)$, where x_t is the state vector, f, $g_j : \mathbb{R}^n \to \mathbb{R}^n$ are nonlinear analytic maps and $\{W_j \in \mathbb{R}, j = 1, \ldots, p\}$ is a set of pairwise independent standard Wiener processes with respect to a family of increasing σ-algebras $\{\mathcal{F}_t,\ t \geq 0\}$. The initial state $x_0 = \bar{x}$ is an $\mathcal{F}_0$-measurable random vector, independent of the state noises $W_{j,t}$.

The problem here investigated is the prediction of x_t given the value x_s at a previous time instant $s < t$, that is, the aim is to compute the conditional expectation $\mathbb{E}(x_t|x_s)$. According to the explicit solution of (1), and to the properties of the Itô integral [1] the expected value is given by

$$\mathbb{E}(x_t|x_s) = x_s + \int_s^t \mathbb{E}(f(x_\xi)|x_s)d\xi. \tag{2}$$

Predictions are broadly exploited in the more general setting of nonlinear filters, according to the usual paradigm suggesting to write the filter equations as a "*prediction*" + a "*correction*" term (see e.g. [1,2]). The common denominator of such approaches, that include the well known Extended Kalman–Bucy filter for the continuous-time case, is that the prediction step is entrusted to the following coarse simplification:

$$\mathbb{E}(f(x_\xi)|x_s) \simeq f\big(\mathbb{E}(x_\xi|x_s)\big). \tag{3}$$

Clearly, such an approximation is exact only in special cases, like linear systems (see e.g. [3]). Such simplification provides an easy-to-handle heuristics since, by exploiting (3), the prediction $z_t = \mathbb{E}(x_t|x_s)$ of (2) can be computed as the solution of the deterministic differential system

$$\dot{z}_t = f(z_t), \qquad z_s = x_s. \tag{4}$$

Motivation for the present note stems from the need to build up a theory providing the solution to the state prediction problem that, in principle, could be applied to such a broad range of nonlinear filtering framework. This problem is especially important when designing continuous-discrete (CD) filters, that is, filters for stochastic differential systems with sampled measurements. CD filters are ubiquitous in problems such as tracking [4], finance [1] and systems biology [5], and they are receiving growing attention in recent years [6–8]. One reason for this interest is that new application areas such as systems biology often employ continuous-time models as in (1), coupled with sampled measurements with large sampling intervals. There can be found many solutions to CD filters in the literature, such as continuous-discrete extended

* Corresponding author.
E-mail addresses: f.cacace@unicampus.it (F. Cacace), valerio.cusimano@iasi.cnr.it (V. Cusimano), germani@ing.univaq.it (A. Germani), pasquale.palumbo@iasi.cnr.it (P. Palumbo).

http://dx.doi.org/10.1016/j.sysconle.2016.10.004
0167-6911/

Kalman filter, CD-EKF [2,9], CD unscented Kalman filter, CD-UKF [10], and CD cubature Kalman filter, CD-CKF [11]. In the case of CD filters the crucial problem is the evolution of the state moments during the sampling interval, since the optimal state estimate over the inter-sampling period (e.g. for $t \in [k\Delta, (k+1)\Delta)$, where $t = k\Delta, k = 0, 1, \ldots$ are the time instants when measurements are acquired) is provided by the prediction $\mathbb{E}(x_t|\hat{x}_{k\Delta})$, formally defined by

$$\mathbb{E}(x_t|\hat{x}_{k\Delta}) = \hat{x}_{k\Delta} + \int_{k\Delta}^{t} \mathbb{E}(f(x_\xi)|\hat{x}_{k\Delta})d\xi. \tag{5}$$

This is, clearly, the same problem introduced in (2) with $s = k\Delta$ and $x_s = \hat{x}(k\Delta)$. In the literature, CD filters share the same approach to solve the optimal state prediction problem based on (3), and Eq. (4) (written with $s = k\Delta$) endowed with the equation of the covariance of the prediction error are also called *differential moment equations*: in this framework several recent works have been devoted to propose precise and efficient methods to evaluate the solution of (4) on small discretization intervals, see for example [6–8]. However, these methods provide precise solutions to the approximate equation (4), whereas the exact prediction (2) cannot be obtained by solving an ordinary differential equation (see for example [2], p. 168). The right hand-side of (2) involves an expectation that requires the whole conditional density for its evaluation. Stated differently, the evaluation of the first two moments of the prediction depends on all the other moments. It should be mentioned that when the discretization interval is not negligible the solutions of (2) and (4) can be quite different even in the simple case of scalar systems.

Furthermore, the application of state predictors for stochastic systems is not limited to CD filters, for example they are useful in the area of stochastic delay equations [12] as well as predictors and filters for stochastic systems with delays in the input and/or the output [13,14].

From a theoretical viewpoint, the solution to the prediction problem can be pursued by searching for the conditional density $p(x_t, t|x_s, s)$ provided by the solution of the Kolmogorov forward equation, and then use p to compute the conditional expectation (2). Since the solution of the Kolmogorov forward equation can be obtained by analytic means only in few cases, a number of numerical methods have been proposed to this aim, including finite-difference method [15], finite elements [16], adaptive finite-elements ([17], pp. 115–123), quadrature-based methods [18], the adjoint method [19], Galerkin's method [20,21], particle methods [22–24] and Markov chain Monte Carlo methods [25]. The computational complexity of solving the Kolmogorov equation increases exponentially with the dimension of the state vector. For this reason, this approach is not well suited for the implementation of real-time predictors (or filters) even for systems of moderate size.

In this work we introduce an approximation scheme for the state prediction equation (2). The proposed solution has several positive features. In the first place we provide the exact solution to the correct problem statement: its analytical form is expressed in terms of a Taylor series expansion, thus the state prediction (2) can be computed with any arbitrary precision. In the second place, and for the same reason, the prediction can be made precise on arbitrary sampling intervals. A final advantage is that, because of the analytic expression provided by the method, the state prediction $\mathbb{E}(x_t|x_s)$ can be used for analysis purposes, and not only as a numerical value. As a drawback, it can be mentioned that our method applies only to systems of the form (1), that is, to time-invariant nonlinear systems, in contrast with other methods that include also the time-varying case.

The approximation scheme is in essence based on the Carleman embedding technique, already exploited in [26] for a slightly different stochastic differential system (nonlinear drift + additive Gaussian noise, instead of the more general nonlinear diffusion term here considered) with the aim of continuous-time filters. The Carleman technique results in the embedding of the original finite-dimensional nonlinear system into an infinite-dimensional bilinear one. Differently from [26], where the state-estimation problem required a further finite-dimensional approximation of the Carleman embedding, here we propose the exact stochastic discretization of the Carleman embedding without any approximation, thus obtaining the optimal prediction as the sum of a series, that we name the *Carleman prediction* in analogy to the deterministic case [27,28].

Section 2 describes the approach, while Section 3 is devoted to its evaluation. In particular we consider a class of systems widely used in financial mathematics and for which the solution of the prediction problem here described has important applications. In a few cases the exact solution can be found by analytic tools. We consider one such cases, with the aim of comparing the exact solution of (2) with the standard approximation (4) and the one provided by our method.

Notation. I_n denotes the identity matrix in $\mathbb{R}^n$. $0_{n,m}$ denotes a matrix of zeros in $\mathbb{R}^{n\times m}$. The symbol $\otimes$ denotes the Kronecker matrix product, the notation $A^{[i]}$ is used for the Kronecker power of matrix A, that is $A \otimes A \otimes \cdots \otimes A$, repeated i times. The standard Jacobian of $f : \mathbb{R}^n \to \mathbb{R}^n$ can be formally written as $\nabla_x \otimes f$, where ∇_x denotes the operator $[\partial/\partial x_1 \ldots \partial/\partial x_n]$. Higher-order derivatives of f are represented as $\nabla_x^{[i]} \otimes f = \nabla_x \otimes \left(\nabla_x^{[i-1]} \otimes f\right)$, where $\nabla_x^{[i]} \otimes f : \mathbb{R}^n \to \mathbb{R}^{n\times n^i}$.

2. Carleman predictor for stochastic systems

Consider the problem of computing the state prediction $\mathbb{E}(x_t|\bar{x})$, with $x_s = \bar{x}$, $t \geq s$ and x_t that evolves according to (1). Whenever useful, the non negative displacement $t - s$ will be referred to as $\Delta \geq 0$. The proposed prediction algorithm is based on the following steps.

1. Define the displacement
$$\varphi_t = x_t - \bar{x}. \tag{6}$$
2. By using the Carleman embedding technique [27] we transform the nonlinear stochastic differential system for φ_t into an infinite-dimensional bilinear system (linear drift and multiplicative noise).
3. Since the system is bilinear, the exact prediction of the state in the embedded space is obtained by integrating the corresponding linear drift without the noise terms.
4. Finally, we project the solution onto the original finite-dimensional space to obtain $\mathbb{E}(x_t|\bar{x})$.

To exploit the Carleman embedding, we shall make use of the representation of the analytic maps $f(x_t)$ and $g_j(x_t)$ as Taylor expansions around $\bar{x}$, written according to the Kronecker formalism (see [26] for details):

$$f(x_t) = \sum_{i=0}^{\infty} A_i(\bar{x})\varphi_t^{[i]}, \quad g_j(\varphi_t) = \sum_{i=0}^{\infty} G_i^j(\bar{x})\varphi_t^{[i]}, \tag{7}$$

$$A_i(\bar{x}) = \left.\frac{\nabla_x^{[i]} \otimes f(x)}{i!}\right|_{x=\bar{x}} \in \mathbb{R}^{n\times n^i}, \tag{8}$$

$$G_i^j(\bar{x}) = \left.\frac{\nabla_x^{[i]} \otimes g_j(x)}{i!}\right|_{x=\bar{x}} \in \mathbb{R}^{n\times n^i}. \tag{9}$$