Automatica 62 (2015) 26–31

Contents lists available at ScienceDirect

Automatica

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

Brief paper

A novel suboptimal method for solving polynomial filtering problems[☆]

Xue Luo[a], Yang Jiao[b], Wen-Lin Chiou[c], Stephen S.-T. Yau[d,1]

[a] School of Mathematics and Systems Science, Beihang University, Beijing, 100191, PR China

[b] Department of Mathematics, Statistics and Computer Science, University of Illinois at Chicago, Science and Engineering Offices (M/C 249), 851 S. Morgan Street, Chicago, IL 60607-7045, USA

[c] Department of Mathematics, Fu-Jen University, New Taipei City, Taiwan, ROC

[d] Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, PR China

ARTICLE INFO

Article history:
Received 22 April 2014
Received in revised form
31 May 2015
Accepted 19 August 2015

Keywords:
Polynomial filtering problems
Suboptimal filter
Recursive algorithm
Estimation theory

ABSTRACT

In this paper we derive the stochastic differentials of the conditional central moments of the nonlinear filtering problems, especially those of the polynomial filtering problem, and develop a novel suboptimal method by solving this evolution equation. The basic idea is to augment the state of the original nonlinear system by including the original states' conditional central moments such that the augmented states form a so-called bilinear system after truncating. During our derivation, it is clear to see that the stochastic differentials of the conditional central moments of the linear filtering problem (i.e., f, g and h are all at most degree one polynomials) form a closed system automatically without truncation. This gives one reason for the existence of optimal filtering for linear problems. On the contrary, the conditional central moments form an infinite dimensional system, in general. To reduce it to a closed-form, we let all the high enough central moments to be zero, as one did in the Carleman approach (Germani et al., 2007). Consequently, a novel suboptimal method is developed by dealing with the bilinear system. Numerical simulation is performed for the cubic sensor problem to illustrate the accuracy and numerical stability.

1. Introduction

The nonlinear filtering (NLF) problem has been extensively studied since the linear one has been satisfactorily solved by Kalman in 1960s. But until now there exists no universal optimal method for the general nonlinear settings. The main goal of NLF is to get "good" estimation of the conditional expectation, or perhaps even the conditional density function of the state, given the observation history. We refer the readers to the book by Jazwinski (1970) for excellent introduction to NLF.

One possible general method to NLF is the so-called global approaches, see the survey paper (Luo, 2014) for more detailed discussion. All these methods try to solve analytically or numerically for the conditional density function. One of the most recent global approach is Yau–Yau's on- and off-line algorithm, which is first derived in Yau and Yau (2008) and further generalized to nonlinear time-varying setting in Luo and Yau (2013). The Yau–Yau's method works for all NLF problems theoretically, however, there are still some technical works to be done for high-dimensional states' problem, say to overcome "the curse of dimensionality". Therefore, certain suboptimal methods still need to be developed.

Another possible way-out to solve the general NLF problems stems from the local approaches, especially the Kalman filter and its derivatives. The basic idea is to augment the states of the original NLF problem in certain way such that the augmented states satisfy a linear or so-called bilinear system (Carravetta, Germani, & Shuakayev, 2000). Generally speaking, one cannot obtain a closed system unless it is Benes' filter (Benes, 1981) or Yau filter (Yau, 1994). The suboptimal filtering therefore is derived by truncating the infinite-dimensional system in some way. In this direction, Basin (2003) is the first paper where the conditional higher order moments were employed for suboptimal polynomial filtering (PF). Later, a series of papers, say (Basin, 2008; Basin, Shi, & Calderon-Alvarez, 2010) and references therein follow this

[☆] The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Michael V. Basin under the direction of Editor Ian R. Petersen.

E-mail addresses: xluo@buaa.edu.cn (X. Luo), jiaoyang0715@gmail.com (Y. Jiao), chiou@math.fju.edu.tw, kubulinchiou@gmail.com (W.-L. Chiou), yau@uic.edu (S.S.-T. Yau).

[1] Tel.: +86 10 62787874; fax: +86 10 62798033.

http://dx.doi.org/10.1016/j.automatica.2015.09.001

line. Based on the suboptimal approach introduced for the bilinear system in Carravetta et al. (2000) and Germani, Manes, and Palumbo (2007) developed a Carleman approximation approach for the NLF problems. In their paper, the higher moments are omitted to form a finite-dimensional system. However, it is as early as in 1967 that Kushner (1967) considered the moment sequences. Even in the one-dimensional problem, the moment sequence has to satisfy the following inequalities:

$$m_2 > 0,\ m_4 > m_2^2,\ m_6 > m_4^2/m_2, \ldots,$$

where $m_s,\ s = 1, 2, \ldots$ represent the s-moment of some random variable. In particular, if the random variable is the standard Gaussian, then all the higher moments can be computed explicitly:

$$m_s = \begin{cases} 0, & \text{all odd } s \geq 1 \\ (s-1)!!\, m_2^s, & \text{all even } s \geq 2, \end{cases}$$

where $(s-1)!! = 1 \cdot 3 \cdot 5 \cdot \cdots \cdot (s-1)$, if s is even. It is easy to see that no matter how small m_2 is, the even moments grows without bound as $s \to \infty$. Therefore, it is inappropriate to let all the higher moments to be zero, even in the Gaussian case.

In this paper, we propose a novel suboptimal method (NSM) for the PF problems by observing the evolution of the conditional *central* moments of the states. Instead of augmenting the states by their higher moments as in the Carleman approach, we derive the evolution of the higher *central* moments, and omit the high enough ones to form a finite-dimensional system. Another novelty of this paper is that we provide an explanation why the optimal method can be derived only for the linear/bilinear filtering problems from the viewpoint of the evolution of the conditional higher central moments. According to Theorem 2, the stochastic differentials of the conditional central moments form a closed system automatically without truncation, if the linear/bilinear filtering problem is considered.

2. Filtering model and notations

The model we consider here is:

$$\begin{cases} dx_t = f(x_t, t)dt + g(t)dv_t \\ dy_t = h(x_t, t)dt + dw_t, \end{cases} \tag{2.1}$$

where $x_t,\ v_t,\ y_t$, and w_t are $\mathbb{R}^n-$, $\mathbb{R}^p-$, $\mathbb{R}^m-$, and $\mathbb{R}^m-$valued processes, respectively, and $f : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$, $g : \mathbb{R}_+ \to \mathbb{R}^{n\times p}$, $h : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^m$ are polynomials with respect to x. Assume that $\{v_t, t \geqslant 0\}$ and $\{w_t, t \geqslant 0\}$ are Brownian motion processes with $\mathrm{Var}[dv_t] = Q(t)dt$ and $\mathrm{Var}[dw_t] = R(t)dt$, respectively. Moreover, $\{v_t, t \geqslant 0\}$, $\{w_t, t \geqslant 0\}$ and x_0 are independent, and $y_0 = 0$.

The conditional expectation for certain process x_t is denoted as $\widehat{x_t} := E[x_t \mid Y_t]$ for short, where $Y_t := \{y_s : 0 \leq s \leq t\}$ is the observation history. Also the a priori conditional expectation is denoted as $\widehat{(\circ)}_- := E[\circ \mid Y_{t_-}]$, where $Y_- := \{y_s : 0 \leq s \leq t_-\}$.

In this paper, we shall use the Kronecker algebra for conciseness. For the quick survey on the Kronecker product and its properties can be found in Carravetta, Germani, and Raimondi (1996). For the readers' convenience, we include some simple facts of Kronecker algebra here. The Kronecker product $\otimes$ is defined for any two matrices $M_{r\times s}$ and $N_{p\times q}$:

$$M \otimes N := \begin{bmatrix} m_{11}N & \cdots & m_{1s}N \\ \cdots & \cdots & \cdots \\ m_{r1}N & \cdots & m_{rs}N \end{bmatrix}.$$

Let $M^{[i]}$ denote the ith Kronecker power of the matrix M, which is defined as

$$M^{[0]} = 1; \qquad M^{[i]} = M \otimes M^{[i-1]} = M^{[i-1]} \otimes M.$$

The stack of the matrix $M_{r\times s} := [m_1, m_2, \ldots, m_s]$ is defined as

$$\mathrm{st}(M) = \left[m_1^T\ m_2^T\ \cdots\ m_s^T\right]^T,$$

where m_i is the i^{th} column of M. The inverse operation of the stack can reduce a vector into a matrix with proper size. That is, $M = \mathrm{st}^{-1}_{r\times s}\mathrm{st}(M)$. The following are properties of Kronecker product and the stack operation:

$$(A+B)\otimes(C+D) = A\otimes C + A\otimes D + B\otimes C + B\otimes D \tag{2.2}$$

$$(A\otimes B)\otimes C = A\otimes(B\otimes C) \tag{2.3}$$

$$(A\cdot C)\otimes(B\cdot D) = (A\otimes B)\cdot(C\otimes D), \tag{2.4}$$

$$u\otimes v = \mathrm{st}(v\cdot u^T), \tag{2.5}$$

where A, B, C, and D are matrices with suitable size, u and v are vectors. As the usual matrix multiplication, Kronecker product is not commutative. Given any two matrices $A \in \mathbb{R}^{r_a\times s_a}$ and $B \in \mathbb{R}^{r_b\times s_b}$, then

$$B\otimes A = C^T_{r_a,r_b}(A\otimes B)C_{s_a,s_b},$$

where $C_{a,b}$ is an orthonormal commutative matrix in $\{0, 1\}^{ab\times ab}$ with its entry (h, l) given by

$$\{C_{a,b}\}_{h,l} = \begin{cases} 1, & \text{if } l = (|h-1|_b)a + \left(\left[\dfrac{h-1}{b}\right]+1\right) \\ 0, & \text{otherwise,} \end{cases}$$

where $[\cdot]$ and $|\cdot|_s$ denote the integer part and s-modulo, respectively. The Kronecker power of a binomial, $(a+b)^{[i]}$ allows the following expansion:

$$(a+b)^{[i]} = \sum_{j=0}^{i} M_j^i\left(a^{[j]}\otimes b^{[i-j]}\right), \tag{2.6}$$

for any $a, b \in \mathbb{R}^n$, where $M_j^i \in \mathbb{R}^{n\times n}$ can be recursively computed

$$M_h^h = M_0^h = I_{n^h},$$

$$M_j^h = (M_j^{h-1}\otimes I_n) + (M_{j-1}^{h-1}\otimes I_n)(I_{n^{j-1}}\otimes G_{h-j}),$$

and $\{G_l\}$ is a sequence that satisfies the following equations:

$$G_1 = C^T_{n,n}, \qquad G_l = (I_n\otimes G_{l-1})\cdot(G_1\otimes I_{n^{l-1}}).$$

See detailed derivation in Carravetta et al. (1996).

In the derivation of our NSM for PF problems, the Itô formula for the computation of stochastic differentials is needed (see Liptser & Shiryayev, 1977). For any vector function $\psi(x_t) : \mathbb{R}^n \to \mathbb{R}^r$, we have

$$d\psi = (\nabla_x\otimes\psi)|_{x_t}\, dx_t + \frac{1}{2}\left(\nabla_x^{[2]}\otimes\psi\right)\Big|_{x_t}(dx_t\otimes dx_t). \tag{2.7}$$

The differential operator $\nabla_x^{[i]}\otimes$ applied to ψ is defined as

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

with $\nabla_x = \left[\frac{\partial}{\partial x_1}\ \frac{\partial}{\partial x_2}\ \cdots\ \frac{\partial}{\partial x_n}\right]$.

3. NSM for PF problems

3.1. Derivation of stochastic differentials for higher central moments

It is well-known from Jazwinski (1970) that the conditional mean of the process x_t satisfies

$$\begin{aligned} d\hat{x}_t &= \hat{f}dt + \widehat{P^{[1]}h^T}R^{-1}\left(dy - \hat{h}dt\right) \\ &= \left[\hat{f} + \widehat{P^{[1]}h^T}R^{-1}(h-\hat{h})\right]dt + \widehat{P^{[1]}h^T}R^{-1}dw_t, \end{aligned} \tag{3.8}$$

where $P^{[\alpha]} := (x - \hat{x}_-)^{[\alpha]}$, for any $\alpha \geq 1$. Thus, it is not hard to get that

$$
\begin{aligned}
dP^{[1]} :=& d\left(x - \hat{x}_-\right) \\
\overset{(2.1),(3.8)}{=}& \left[\left(f - \hat{f}_-\right) - \widehat{P^{[1]}h^T}_- R^{-1}\left(h - \hat{h}_-\right)\right] dt \\
&+ g dv_t - \widehat{P^{[1]}h^T}_- R^{-1} dw_t.
\end{aligned} \tag{3.9}
$$

Theorem 1. *For any $\alpha \geq 1$,*

$$
\begin{aligned}
dP^{[\alpha]} =& \left\{U_n^\alpha\left[\left(\left(f - \hat{f}_-\right) - \widehat{P^{[1]}h^T}_- R^{-1}\left(h - \hat{h}_-\right)\right) \otimes P^{[\alpha-1]}\right]\right. \\
&\left. + \frac{1}{2} O_n^\alpha\left[\mathrm{st}\left(gQg^T + \widehat{P^{[1]}h^T}_- R^{-1}\widehat{hP^{[1]T}}_-\right) \otimes P^{[\alpha-2]}\right]\right\} dt \\
&+ U_n^\alpha\left(I_n \otimes P^{[\alpha-1]}\right)\left[g dv_t - \widehat{P^{[1]}h^T}_- R^{-1} dw_t\right],
\end{aligned} \tag{3.10}
$$

with convention that $P^{[\alpha]} = 0$, if $\alpha < 0$, and $P^{[0]} = 1$, where I_n is the $n \times n$ identity matrix, U_n^α and O_n^α, for $\alpha > 1$, are recursively computed as

$$
U_n^1 = I_n, \qquad U_n^\alpha = I_{n^\alpha} + C_{n^{\alpha-1},n}^T\left(U_n^{\alpha-1} \otimes I_n\right) \tag{3.11}
$$

$$
O_n^\alpha = U_n^\alpha C_{n^{\alpha-1},n}^T\left(\left(U_n^{\alpha-1} C_{n^{\alpha-2},n}^T\right) \otimes I_n\right) C_{n^2,n^{\alpha-2}}^T.
$$

Proof. Taking ψ in (2.7) as the function of $P^{[1]}$ yields that

$$
\begin{aligned}
dP^{[\alpha]} =& \left(\nabla_{P^{[1]}} \otimes P^{[\alpha]}\right) dP^{[1]} \\
&+ \frac{1}{2}\left(\nabla_{P^{[1]}}^{[2]} \otimes P^{[\alpha]}\right)\left(dP^{[1]} \otimes dP^{[1]}\right).
\end{aligned} \tag{3.12}
$$

According to Lemma 1, Germani et al. (2007), we have

$$
\begin{aligned}
\nabla_x \otimes x^{[\alpha]} &= U_n^\alpha\left(I_n \otimes x^{[\alpha-1]}\right); \\
\nabla_x^{[2]} \otimes x^{[\alpha]} &= O_n^\alpha\left(I_{n^2} \otimes x^{[\alpha-2]}\right),
\end{aligned}
$$

where U_n^α and O_n^α are given in (3.11), see the proof in the appendix of Germani et al. (2007). Eq. (3.10) follows immediately from (3.9), (3.12) and the fact that

$$
\begin{aligned}
dP^{[1]} \otimes dP^{[1]} &= \mathrm{st}\left(dP^{[1]} \cdot \left(dP^{[1]}\right)^T\right) \\
&\overset{(3.9)}{=} \mathrm{st}\left[gQg^T + \widehat{P^{[1]}h^T}_- R^{-1}\widehat{hP^{[1]T}}_-\right] dt. \qquad \square
\end{aligned}
$$

3.2. Stochastic differentials and NSM for PF problems

In the sequel, we shall apply Theorem 1 to the polynomial filtering problem (2.1) with f, g and h written in the Kronecker product notation:

$$
\begin{aligned}
f(x,t) &= \sum_{l=0}^{\deg(f)} F_l x^{[l]}; \qquad g_{:,i}(x,t) = \sum_{l=0}^{\deg(g)} G_{i,l} x^{[l]}; \\
h(x,t) &= \sum_{l=0}^{\deg(h)} H_l x^{[l]},
\end{aligned} \tag{3.13}
$$

where $\deg(f)$, $\deg(g)$ and $\deg(h)$ are the degrees of the polynomials f, g and h, respectively. $F_l : \mathbb{R}^+ \to \mathbb{R}^{n \times n^l}$, $G_{:,l} : \mathbb{R}^+ \to \mathbb{R}^{n \times n^l}$ and $H_l : \mathbb{R}^+ \to \mathbb{R}^{m \times n^l}$. For the conciseness of notations, we may omit the t-dependence in F_l, $G_{:,l}$ and H_l in the sequel.

Theorem 2. *Consider the system (2.1). Assume further that $Q = \mathrm{diag}(q_i)$, $i = 1, \ldots, p$, and $R = \mathrm{diag}(r_j)$, $j = 1, \ldots, m$. For $\alpha \geq 1$, the stochastic differentials of the central moments $P^{[\alpha]}$ are given*

$$
\begin{aligned}
& dP^{[\alpha]} \\
&= \left\{U_n^\alpha\left[\sum_{q=\alpha-1}^{\deg(f)+\alpha-1}\sum_{k=0}^{\deg(f)-q+\alpha-1}\right.\right. \\
&\qquad \left(F_{q+k-\alpha+1} M_k^{q+k-\alpha+1} \otimes I_{n^{\alpha-1}}\right)\left(\hat{x}_-^{[k]} \otimes I_{n^q}\right) P^{[q]} \\
&\qquad - \sum_{l=0}^{\deg(f)}\sum_{k=0}^{l} \\
&\qquad \left.\left(F_l M_k^l \otimes I_{n^{\alpha-1}}\right)\left(\hat{x}_-^{[k]} \otimes \widehat{P^{[l-k]}}_- \otimes I_{n^{\alpha-1}}\right) P^{[\alpha-1]}\right] \\
&- U_n^\alpha\left(\widehat{P^{[1]}h^T}_- R^{-1} \otimes I_{n^{\alpha-1}}\right) \\
&\quad \cdot \left[\sum_{q=\alpha-1}^{\deg(h)+\alpha-1}\sum_{k=0}^{\deg(f)-q+\alpha-1}\right. \\
&\qquad \left(H_{q+k-\alpha+1} M_k^{q+k-\alpha+1} \otimes I_{n^{\alpha-1}}\right)\left(\hat{x}_-^{[k]} \otimes I_{n^q}\right) P^{[q]} \\
&\qquad \left. - \sum_{l=0}^{\deg(h)}\sum_{k=0}^{l}\left(H_l M_k^l \otimes I_{n^{\alpha-1}}\right)\left(\hat{x}_-^{[k]} \otimes \widehat{P^{[l-k]}}_- \otimes I_{n^{\alpha-1}}\right) P^{[\alpha-1]}\right] \\
&+ \frac{1}{2} O_n^\alpha \sum_{i=1}^{p} q_i \left\{\left[\sum_{s=\alpha-2}^{\deg(g)+\alpha-2}\sum_{j=0}^{\deg(g)-s+\alpha-2}\sum_{k=0}^{j+s-\alpha+2}\right.\right. \\
&\qquad + \left(\sum_{s=\alpha-2}^{\deg(g)+\alpha-1}\sum_{j=\deg(g)-s+\alpha-1}^{\deg(g)-s+\alpha-2}\right. \\
&\qquad \left.\left. + \sum_{s=\deg(g)+\alpha}^{2\deg(g)+\alpha-2}\sum_{j=0}^{2\deg(g)-s+\alpha-2}\right)\sum_{k=-\deg(g)+j+s-\alpha+2}^{\deg(g)}\right] \\
&\qquad \left(\left(G_{i,j+s-\alpha+2-k} \otimes G_{i,k}\right) M_j^{j+s-\alpha+2} \otimes I_{n^{\alpha-2}}\right) \\
&\qquad \left. \cdot \left(\hat{x}_-^{[j]} \otimes I_{n^s}\right)\right\} P^{[s]} \\
&\left. + \frac{1}{2} O_n^\alpha \sum_{i=1}^{m} r_i^{-1}\left[\left(\widehat{P^{[1]}h^T}_-\right)_{:,i}^{[2]} \otimes I_{n^{\alpha-2}}\right] P^{[\alpha-2]}\right\} dt \\
&+ U_n^\alpha \sum_{i=1}^{p}\sum_{q=\alpha-1}^{\deg(g)+\alpha-1}\sum_{j=0}^{\deg(g)-q+\alpha-1} \\
&\quad \cdot \left(G_{i,q+j-\alpha+1} M_j^{q+j-\alpha+1} \otimes I_{n^{\alpha-1}}\right)\left(\hat{x}_-^{[j]} \otimes I_{n^q}\right) P^{[q]} (dv_t)_i \\
&- U_n^\alpha\left(\widehat{P^{[1]}h^T}_- R^{-1} dw_t \otimes I_{n^{\alpha-1}}\right) P^{[\alpha-1]},
\end{aligned} \tag{3.14}
$$

with the convention that $P^{[\alpha]} = 0$, if $\alpha < 0$ and $P^{[0]} = 1$, where

$$
\widehat{P^{[1]}h^T}_- = \left[\sum_{l=0}^{\deg(h)}\sum_{k=0}^{l} H_l M_k^l\left(\hat{x}_-^{[k]} \otimes \mathrm{st}_{n^{l-k} \times n}^{-1}\widehat{P^{[l-k+1]}}_-\right)\right]^T,
$$

and $(\circ)_{:,i}$ denotes the ith column of the matrix $(\circ)$.

Proof. Direct computation of the terms on the right-hand side of (3.8) with f, g and h as polynomials (3.13) yields that

$$
\hat{f}_- = \sum_{l=0}^{\deg(f)} F_l \widehat{x^{[l]}}_- \overset{(2.6)}{=} \sum_{l=0}^{\deg(f)}\sum_{k=0}^{l} F_l M_k^l\left(\hat{x}_-^{[k]} \otimes \widehat{P^{[l-k]}}_-\right), \tag{3.15}
$$

and similarly,

$$
\hat{h}_- = \sum_{l=0}^{\deg(h)}\sum_{k=0}^{l} H_l M_k^l\left(\hat{x}_-^{[k]} \otimes \widehat{P^{[l-k]}}_-\right). \tag{3.16}
$$

The first term on the right-hand side of (3.10) can be obtained by Eqs. (3.17)–(3.19):

$$
\begin{aligned}
&\left(f-\hat f_-\right)\otimes P^{[\alpha-1]}\\
&\overset{(3.15)}{=}\sum_{l=0}^{\deg(f)}\sum_{k=0}^{l}\left\{F_lM_k^l\left[\hat x_-^{[k]}\otimes\left(P^{[l-k]}-\widehat{P^{[l-k]}}_-\right)\right]\right\}\otimes P^{[\alpha-1]}\\
&\overset{(2.4)}{=}\sum_{l=0}^{\deg(f)}\sum_{k=0}^{l}\left(F_lM_k^l\otimes I_{n^{\alpha-1}}\right)\\
&\quad\cdot\left[\hat x_-^{[k]}\otimes\left(P^{[l-k+\alpha-1]}-\widehat{P^{[l-k]}}_-\otimes P^{[\alpha-1]}\right)\right]\\
&=\sum_{q=\alpha-1}^{\deg(f)+\alpha-1}\;\sum_{k=0}^{\deg(f)-q+\alpha-1}\left(F_{q+k-\alpha+1}M_k^{q+k-\alpha+1}\otimes I_{n^{\alpha-1}}\right)\\
&\quad\cdot\left(\hat x_-^{[k]}\otimes I_{n^q}\right)P^{[q]}\\
&\quad-\sum_{l=0}^{\deg(f)}\sum_{k=0}^{l}\left(F_lM_k^l\otimes I_{n^{\alpha-1}}\right)\\
&\quad\cdot\left(\hat x_-^{[k]}\otimes\widehat{P^{[l-k]}}_-\otimes I_{n^{\alpha-1}}\right)P^{[\alpha-1]},
\end{aligned}
\tag{3.17}
$$

where the last equality holds due to reordering the summation. Similarly, we have

$$
\begin{aligned}
&\left(h-\hat h_-\right)\otimes P^{[\alpha-1]}\\
&=\sum_{q=\alpha-1}^{\deg(h)+\alpha-1}\;\sum_{k=0}^{\deg(h)-q+\alpha-1}\left(H_{q+k-\alpha+1}M_k^{q+k-\alpha+1}\otimes I_{n^{\alpha-1}}\right)\\
&\quad\cdot\left(\hat x_-^{[k]}\otimes I_{n^q}\right)P^{[q]}\\
&\quad-\sum_{l=0}^{\deg(h)}\sum_{k=0}^{l}\left(H_lM_k^l\otimes I_{n^{\alpha-1}}\right)\\
&\quad\cdot\left(\hat x_-^{[k]}\otimes\widehat{P^{[l-k]}}_-\otimes I_{n^{\alpha-1}}\right)P^{[\alpha-1]},
\end{aligned}
\tag{3.18}
$$

and

$$
\begin{aligned}
&\widehat{P^{[\alpha]}h^T}_-=\left(\widehat{hP^{[\alpha]T}}_-\right)^T\\
&\overset{(2.4),(2.5)}{=}\left[\sum_{l=0}^{\deg(h)}\sum_{k=0}^{l}H_lM_k^l\left(\hat x_-^{[k]}\otimes \mathrm{st}^{-1}_{n^{l-k}\times n^{\alpha}}\widehat{P^{[l-k+\alpha]}}_-\right)\right]^T.
\end{aligned}
\tag{3.19}
$$

The second term on the right-hand side of (3.10) follows immediately by the following fact that

$$
\mathrm{st}\left(gQg^T\right)=\sum_{i=1}^{p}q_ig_{:,i}^{[2]}\triangleq\tilde g^{[2]},
\tag{3.20}
$$

where $(\circ)_{:,i}$ denotes the ith column of the matrix $\circ$.

$$
\begin{aligned}
&\mathrm{st}\left(gQg^T\right)\otimes P^{[\alpha-2]}\\
&\overset{(3.20),(3.13)}{=}\sum_{l,k=0}^{\deg(g)}\sum_{i=1}^{p}q_i\left[\left(G_{i,l}\otimes G_{i,k}\right)x^{[l+k]}\right]\otimes P^{[\alpha-2]}\\
&\overset{(2.4)}{=}\sum_{i=1}^{p}q_i\sum_{l,k=0}^{\deg(g)}\sum_{j=0}^{l+k}\left(\left(G_{i,l}\otimes G_{i,k}\right)M_j^{l+k}\otimes I_{n^{\alpha-2}}\right)\\
&\quad\cdot\left(\hat x_-^{[j]}\otimes I_{n^{l+k-j+\alpha-2}}\right)P^{[l+k-j+\alpha-2]}.
\end{aligned}
\tag{3.21}
$$

Notice that the summation can be reordered as

$$
\begin{aligned}
&\sum_{l,k=0}^{\deg(g)}\sum_{j=0}^{l+k}\\
&\overset{q=l+k}{=}\left(\sum_{q=0}^{\deg(g)}\sum_{k=0}^{q}+\sum_{q=\deg(g)+1}^{2\deg(g)}\;\sum_{k=-\deg(g)+q}^{\deg(g)}\right)\sum_{j=0}^{q}\\
&\overset{s=q-j+\alpha-2}{=}\sum_{s=\alpha-2}^{\deg(g)+\alpha-2}\;\sum_{j=0}^{\deg(g)-s+\alpha-2}\;\sum_{k=0}^{j+s-\alpha+2}\\
&\quad+\left(\sum_{s=\alpha-2}^{\deg(g)+\alpha-1}\;\sum_{j=\deg(g)-s+\alpha-1}^{\deg(g)-s+\alpha-2}\right.\\
&\quad\left.+\sum_{s=\deg(g)+\alpha}^{2\deg(g)+\alpha-2}\;\sum_{j=0}^{2\deg(g)-s+\alpha-2}\right)\\
&\quad\cdot\sum_{k=-\deg(g)+j+s-\alpha+2}^{\deg(g)}.
\end{aligned}
$$

Thus, (3.21) can be written as

$$
\begin{aligned}
&\mathrm{st}\left(gQg^T\right)\otimes P^{[\alpha-2]}\\
&=\sum_{i=1}^{p}q_i\left\{\left[\sum_{s=\alpha-2}^{\deg(g)+\alpha-2}\;\sum_{j=0}^{\deg(g)-s+\alpha-2}\;\sum_{k=0}^{j+s-\alpha+2}\right.\right.\\
&\quad+\left(\sum_{s=\alpha-2}^{\deg(g)+\alpha-1}\;\sum_{j=\deg(g)-s+\alpha-1}^{\deg(g)-s+\alpha-2}\right.\\
&\quad\left.\left.+\sum_{s=\deg(g)+\alpha}^{2\deg(g)+\alpha-2}\;\sum_{j=0}^{2\deg(g)-s+\alpha-2}\right)\sum_{k=-\deg(g)+j+s-\alpha+2}^{\deg(g)}\right]\\
&\quad\times\left(\left(G_{i,j+s-\alpha+2-k}\otimes G_{i,k}\right)M_j^{j+s-\alpha+2}\otimes I_{n^{\alpha-2}}\right)\\
&\quad\left.\cdot\left(\hat x_-^{[j]}\otimes I_{n^s}\right)\right\}P^{[s]}.
\end{aligned}
$$

Similarly as in (3.20), we have

$$
\begin{aligned}
&\mathrm{st}\left(\widehat{P^{[1]}h^T}_-R^{-1}\widehat{hP^{[1]T}}_-\right)\otimes P^{[\alpha-2]}\\
&\overset{(3.20)}{=}\sum_{i=1}^{m}r_i^{-1}\left[\left(\widehat{P^{[1]}h^T}_-\right)_{:,i}^{[2]}\otimes I_{n^{\alpha-2}}\right]P^{[\alpha-2]},
\end{aligned}
\tag{3.22}
$$

where $\widehat{P^{[1]}h^T}_-$ is in (3.19) with $\alpha=1$. The last term is given by (3.19) and

$$
\begin{aligned}
&gdv_t\otimes P^{[\alpha-1]}\\
&=\sum_{i=1}^{p}\;\sum_{q=\alpha-1}^{\deg(g)+\alpha-1}\;\sum_{j=0}^{\deg(g)-q+\alpha-1}\\
&\quad\cdot\left(G_{i,q+j-\alpha+1}M_j^{q+j-\alpha+1}\otimes I_{n^{\alpha-1}}\right)\left(\hat x_-^{[j]}\otimes I_{n^q}\right)P^{[q]}(dv_t)_i.
\end{aligned}
\tag{3.23}
$$

Eq. (3.14) follows immediately from Theorem 1 and (3.17)–(3.23). □

Remark 3. (1) If $\deg(f)$, $\deg(h)$ and $\deg(g)\leq 1$, the system is a linear (or so-called bilinear Carravetta et al., 2000) system, then from Theorem 2, (3.14) with arbitrary truncation $\alpha\geq 2$, (3.14) forms a closed system, and consequently the optimal estimation can be obtained. However, (3.14) cannot be closed with one of the degrees of $\deg(f)$, $\deg(h)$ and $\deg(g)$ greater than one.

(2) (3.14) is linear with respect to $P^{[\alpha]}$s, with $\alpha\geq 1$.

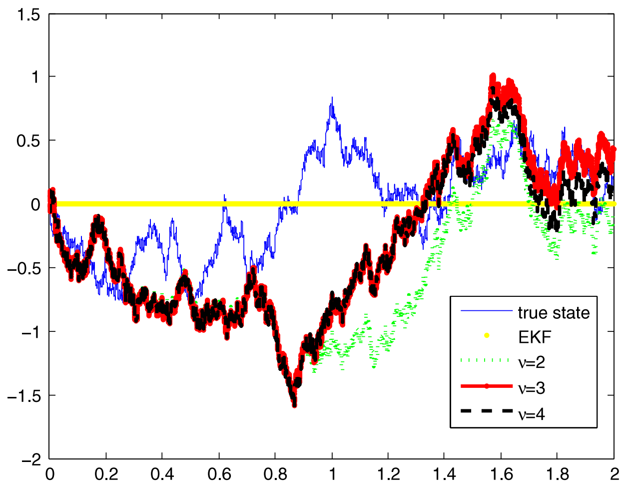

Fig. 1. Blue solid line: the true state generated by *randn*(‘*state*’,*1*); green dotted line: the estimation by using NSM with $\nu = 2$; red dotted and dashed line: with $\nu = 3$; black dashed line: with $\nu = 4$; yellow dotted line: EKF. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

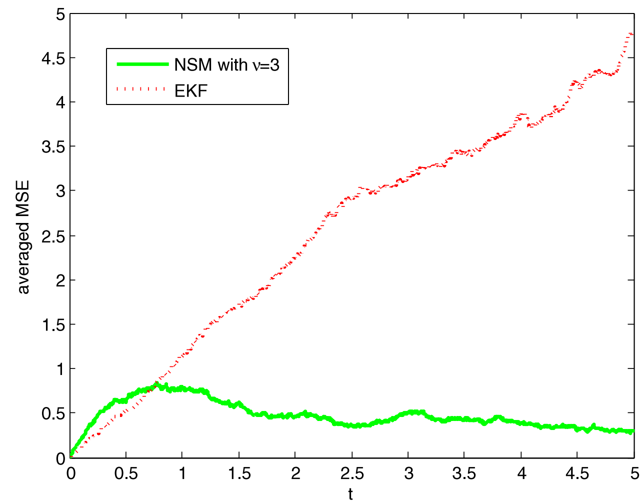

Fig. 2. Green solid line: the averaged timewise MSE of the true state and our NSM with $\nu = 3$ averaged over 300 simulation runs; Red dotted line: the averaged timewise MSE by using EKF. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Algorithm for NSM:

Step 1 Form a closed system for (3.14) by omitting the higher degree terms. For example, if the truncation is chosen to be some $N_\alpha \geq 2$, then all the $P^{[\alpha]}$ with $\alpha \geq N_a + 1$ is set to be zero.

Step 2 Apply the suboptimal approach in Carravetta et al. (2000) for bilinear system.

4. Numerical experiment

We experiment our NSM on the following benchmark example, cubic sensor problem:

$$\begin{cases} dx_t = dv_t \\ dy_t = x_t^3 dt + dw_t, \end{cases} \tag{4.24}$$

where $Q = R = 1$. The initial state x_0 obeys the standard Gaussian. The results are obtained by using the Euler–Maruyama method (Higham, 2001), with time step $\triangle t = 0.01$. To form a closed system, we need to truncate the system (3.14) with some truncation mode ν. That is, we ignore all the terms $P^{[\alpha]}$ with $\alpha > \nu$. In Fig. 1, we illustrate the performances corresponding to different truncation modes $\nu = 2, 3$ and 4. The true state is generated by *randn*(‘*state*’,*1*). The mean of the squared estimation error (MSE) are 0.5881, 0.3965 and 0.3934 for different $\nu = 2, 3$ and 4, respectively. For comparison, the result from EKF has also been plotted. The MSE for one realization is defined as:

$$\text{MSE}\left(\hat{x}\right) = \frac{1}{N+1} \sum_{k=0}^{N} (x(t_k) - \hat{x}(t_k))^2,$$

where N is the number of the time steps.

Fig. 2 compares the timewise MSE of our NSM with truncation mode $\nu = 3$ with that of extended Kalman filter (EKF) averaged over 300 simulation runs, which is generated by *randn*(‘*state*’,*s*), $s = 1, 2, \ldots, 300$. The timewise MSE is defined as

$$\text{Timewise MSE}\left(\hat{x}\right) = \left[(x(t_k) - \hat{x}(t_k))^2\right]_{k=0}^{N},$$

which is a $(N + 1)$-vector of MSE at each time step. It is clear to see that our method has the averaged timewise MSE under control, while that of EKF grows without bound.

Acknowledgments

X. Luo thanks the support from the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, the Fundamental Research Funds for the Central Universities (YWF-15-SXXY-006, YWF-15-SXXY-002), Beijing Natural Science Foundation (1154011) and National Natural Science Foundation of China (11501023). The research of S.S-T. Yau is supported by the start-up fund from Tsinghua University. Both X. Luo and S.S-T. Yau gratefully acknowledge the support of the National Natural Science Foundation of China (11471184).

References

Basin, M. (2003). On optimal filtering for polynomial system states. *ASME Transactions on Journal of Dynamic Systems, Measurement and Control*, *125*(1), 123–125.
Basin, M. (2008). *Lecture notes in control and information sciences*, *New trends in optimal filtering and control for polynomial and time-delay systems*. Berlin: Springer-Verlag.
Basin, M., Shi, P., & Calderon-Alvarez, D. (2010). Approximate finite-dimensional filtering for polynomial states over polynomial observations. *International Journal of Control*, *83*(4), 724–730.
Benes, V. E. (1981). Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, *5*(1–2), 65–92.
Carravetta, F., Germani, A., & Raimondi, M. (1996). Polynomial filtering for linear discrete time non-Gaussian systems. *SIAM Journal on Control and Optimization*, *34*(5), 1666–1690.
Carravetta, F., Germani, A., & Shuakayev, M. (2000). A new suboptimal approach to the filtering problem for bilinear stochastic differential systems. *SIAM Journal on Control and Optimization*, *38*(4), 1171–1203.
Germani, A., Manes, C., & Palumbo, P. (2007). Filtering of stochastic nonlinear differential systems via a Carleman approximation approach. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, *52*(11), 2166–2172.
Higham, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review*, *43*, 525–546.
Jazwinski, A. H. (1970). *Stochastic processes and filtering theory*. New York: Academic Press.
Kushner, H. (1967). Approximations to optimal nonlinear filters. *Institute of Electrical and Electronics Engineers Transactions on Automatic Control*, *12*(5), 546–556.
Liptser, R. S., & Shiryayev, A. N. (1977). *Statistics of random processes I and II*. New York, Heidelberg: Springer-Verlag.
Luo, X. (2014). On recent advance of nonlinear filtering theory: emphases on global approaches (survey paper). *Pure and Applied Mathematics Quarterly*, *10*(4), 685–721.
Luo, X., & Yau, S. S.-T. (2013). Complete real time solution of the general nonlinear filtering problem without memory. *Institute of Electrical and Electronics Engineers. Transactions on Automatic Control*, *58*(10), 2563–2578.

Yau, S. S.-T. (1994). Finite dimensional filters with nonlinear drift I: A class of filters including both Kalman-Bucy filters and Benes filters. *Journal of Mathematical Systems, Estimation, and Control*, *4*(2), 181–203.
Yau, S.-T., & Yau, S. S.-T. (2008). Real time solutions of the nonlinear filtering problem without memory II. *SIAM Journal on Control and Optimization*, *47*(1), 163–195.

Xue Luo received her first Ph.D. degree in Mathematics from East China Normal University, Shanghai, PR China in 2010 and her second Ph.D. degree in Applied Mathematics from the department of mathematics, statistics and computer science, University of Illinois at Chicago (UIC) in 2013. Around 2009 and 2010 respectively, she visited University of Connecticut and UIC as visiting scholar. Till now, she is currently an assistant professor in school of mathematics and systems science, Beihang University, Beijing, PR China.

Her research interests include analysis of partial differential equations, nonlinear filtering theory, numerical analysis of spectral methods, sparse grid algorithm and fluid mechanics.

Yang Jiao received his Bachelor's degree in Applied Mathematics from the University of Science and Technology of China in 2006, and M.S. and Ph.D. degrees in Applied Mathematics from University of Illinois at Chicago in 2008 and 2012 respectively.

Wen-Lin Chiou received her Bachelor's and M.S. degrees in Mathematics from the Fu-Jen University in 1982 and 1984, respectively, and a Ph.D. degree in Mathematics from the University of Illinois at Chicago in 1991. She joined the Department of Mathematics, Fu-Jen Catholic University, Taiwan, in 1991. Currently, she is an associated Professor at Fu-Jen University, Taiwan. Her main research interests include nonlinear filtering theory and cooperative game.

Stephen S-T. Yau received the Ph.D. degree in Mathematics from the State University of New York at Stony Brook, NY, US, in 1976. He was a member of Institute of Advanced Study at Princeton 1976–1977 and 1981–1982, and a Benjamin Pierce Assistant Professor at Harvard University during 1977–1980. After that, he joined the department of mathematics, statistics and computer science (MSCS), University of Illinois at Chicago (UIC), and served for over 30 years. He was awarded Sloan Fellowship in 1980, Guggenheim Fellowship in 2000, IEEE Fellow Award in 2003 and AMS Fellow Award in 2013. In 2005, he was entitled the UIC distinguished professor. During 2005–2011, he became a joint-professor of department of electrical and computer engineering and MSCS, UIC. In 2011, he joined Tsinghua University, Beijing, PR China as full-time professor in department of mathematical sciences.

His research interests include nonlinear filtering, bioinformatics, complex algebraic geometry, CR geometry and singularities theory.

He is the Managing Editor and founder of Journal of Algebraic Geometry from 1991, and the Editors-in-Chief and founder of Communications in Information and Systems from 2000 till now. He was the General Chairman of IEEE International Conference on Control and Information, which was held in the Chinese University of Hong Kong in 1995.