Direct method for time-varying nonlinear filtering problems

Xiuqiong Chen, Xue Luo, *Senior Member, IEEE,* and Stephen S.-T. Yau, *Fellow, IEEE*

***Abstract*—This paper discusses how to solve filtering problem for a class of continuous nonlinear time-varying systems via Duncan-Mortensen-Zakai (DMZ) equation. In this paper, the original DMZ equation is changed into Kolmogorov forward equation (KFE) by exponential transformations in each time interval, and then under some assumptions, the KFE can be transformed into time-varying Schrödinger equation which can be solved explicitly. The novelty of this paper lies in how to transform the KFE into Schrödinger equation. As a direct application, the results of [19] are extended for time-varying Yau systems.**

***Index Terms*—Nonlinear filtering, Duncan-Mortensen-Zakai (DMZ) equation, time-varying Yau system, Schrödinger Equation.**

I. INTRODUCTION

The problem of estimating the state of a stochastic dynamical system from noisy observations taken on the state is of central importance in engineering which known as filtering problem. The continuous time-varying filtering problem considered in this paper can be stated as follows:

$$\begin{cases} dx_t = f(x_t,t)dt + g(t)dv_t \\ dy_t = h(x_t,t)dt + dw_t \end{cases} \tag{1}$$

where $x_t, f \in \mathbb{R}^{n\times 1}$, g is an $n\times r$ matrix, v_t is an $r-$vector Brownian motion process with $E[dv_t dv_t^T] = \tilde{Q}(t)dt$ and $\tilde{Q}(t) > 0$, $y_t, h \in \mathbb{R}^{m\times 1}$ and w_t is an $m-$vector Brownian motion process with $E[dw_t dw_t^T] = S(t)dt$ and $S(t) > 0$. Here we refer x_t as the state of the system at time t, $f(x_t,t)$ as the drift term, $\tilde{Q}(t), S(t)$ as the variance of the noises and y_t as the observation at time t with $y_0 = 0$.

System (1) can model most practical physical situation. Taking a falling body in a constant field as an example, it can be modelled by a noise disturbed second-order system,

$$\ddot{z} = g_a + \tilde{v}_t,\ t \geq 0 \tag{2}$$

where the scalar z is the position, g_a is gravitational acceleration and $\tilde{v}_t$ is the white noise due to air friction. Let the position be $z = x_1$ and the velocity $\dot{z} = x_2$. Then, defining the state vector $x_t = [x_1, x_2]^T$, (2) can be written in system form

$$\dot{x}_t = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ g_a \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tilde{v}_t. \tag{3}$$

X. Q. Chen is with the Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China (cxq14@mails.tsinghua.edu.cn).

X. Luo is with the School of Mathematics and Systems Science, Beihang University, Beijing 100191, China (xluo@buaa.edu.cn).

S. S.-T. Yau is with the Department of Mathematical Sciences, Tsinghua University, Beijing, 100084, China (yau@uic.edu).

We get scalar observations of position

$$\tilde{y}_t = [1,0]x_t + \tilde{w}_t \tag{4}$$

where $\tilde{w}_t$ is the white noise due to measuring error. Since the white noise can be regarded as the derivative of Brownian motion in Itô sense [8], we can rewrite the mathematical model as follows:

$$\begin{cases} dx_t = \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_t + \begin{bmatrix} 0 \\ g_a \end{bmatrix} \right) dt + \begin{bmatrix} 0 \\ 1 \end{bmatrix} dv_t \\ dy_t = [1,0]x_t dt + dw_t \end{cases} \tag{5}$$

where $dv_t = \tilde{v}_t dt, dw_t = \tilde{w}_t dt, dy_t = \tilde{y}_t dt$.

It is widely acknowledged that the most influential work in filtering theory is the classical Kalman filter (KF) [6] which was published in 1960 and its continuous counterpart Kalman-Bucy filter [7]. Since then, there have been many derivatives of KF in filtering theory and control area. These methods approximate certain statistical quantities, such as mean and variance, among which the extended Kalman filter (EKF) is one of the most famous ones for the nonlinear filtering (NLF) problems. Till now, KF and its derivatives are still widely used in various industrial applications and scientific problems such as tracking, communications, economics and etc. Another possible way which is called global approach in [10] is to solve for the conditional density function of the states either directly or numerically. It is known that the unnormalized density function of the states satisfies the Duncan-Mortensen-Zakai (DMZ) equation [3], [11], [21]. Recently the second and the third author of this paper have developed a real-time algorithm for the general NLF problems without memory based on DMZ equation [9]. Its effectiveness has been numerically verified in very low dimensional problems. For more results on this direction, we refer interested readers for the survey paper [10].

In some sense, the NLF problems are said to be completely solved, if one can solve DMZ equation in real time and in a memoryless way since all the statistical information can be extracted from conditional density function of the states. For the past quarter of a century, as far as we know, there are two methods to solve DMZ equation explicitly. The first one is to use Lie algebraic method to solve the DMZ equation via the Wei-Norman approach. The basic idea is to reduce the DMZ equation to a finite system of ordinary differential equations (ODE), Kolmogorov equation, and several first-order linear partial differential equations (PDE). However, one must know the basis of the estimation algebra. The third author and his co-workers [4], [17] have completely classified all finite dimensional estimation algebras of maximal rank. In particular, they have proved that all the observation terms

$h_i(x), 1 \le i \le m$, must be polynomials of degree one. Another approach is the direct method introduced in [14], [16]. Compared with the Lie algebraic method, it does not need to know the basis. Furthermore, it is unnecessary to integrate n first-order linear PDEs, which is inevitable in the Lie algebraic method. We need to remark that, all the direct methods are for the Yau systems [19], i.e., $f(x,t)$ in (1) is of the form $f(x,t) = Lx + l + \nabla\phi(x)$ where L, l are matrices with proper dimensions and $\phi(x)$ is a C^∞ function. Though it seems restrictive, it includes Kalman-Bucy [7] and Beneš [1] filtering systems as its special cases. Under the assumption that the noises' covariances of state and observation are identity matrices and the system is time-invariant, the direct method has been extensively studied in [5], [15], [16], [18] and [19].

Time-invariant system can only be seen as an ideal model of practical applications. Thus, it is more meaningful to solve time-varying NLF problems. In this paper, we shall consider filtering problems for the time-varying Yau systems with time-varying covariances which extends the results of that in [19]. The novelty of this paper lies in how to transform the Kolmogorov forward equation (KFE) into a time-varying Schrödinger equation with respect to the time-varying nonlinear systems since the corresponding DMZ equation are much more complicated than that in [19].

This paper is organized as follows. The basic model and some preliminary results are stated in the next section. In section III, we shall construct a transformation to change the Kolmogorov equation into the one without drift term which is stated in Theorem 1. With further assumption on potential, we solve the KFE formally and directly by power series method in section IV. In section V, we use direct method to solve a numerical example and compare it with EKF. We arrive at the conclusion in the last section.

II. Preliminaries

Firstly we give some assumptions in terms of system (1). We assume that $G(t) \triangleq g(t)\tilde{Q}(t)g^T(t)$ is C^∞ smooth, $f(x,t)$ and $h(x,t)$ are C^∞ smooth in both state and time. For the sake of clarity we state some notations first: $*_{ij}$ denotes the $ij-$entry of the matrix $*$, $*_i$ denotes the $i-$th element of the vector $*$ and $*^T$ denotes the transposition of $*$.

The unnormalized density function $\sigma(t,x)$ of x_t conditioned on the observation history $\mathscr{F}_t \triangleq \{y_s : \ 0 \le s \le t\}$ satisfies the DMZ equation [3]:

$$\begin{cases} d\sigma(t,x) = L\sigma(t,x)dt + \sigma(t,x)h^T(x,t)S^{-1}(t)dy_t \\ \sigma(0,x) = \sigma_0(x), \end{cases} \tag{6}$$

where $\sigma_0(x)$ is the probability density of the initial sate x_0, and

$$L(*) \equiv \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2}{\partial x_i \partial x_j}\left[G_{ij}(t)*\right] - \sum_{i=1}^n \frac{\partial(f_i *)}{\partial x_i}. \tag{7}$$

For each arrived observation, making an invertible exponential transformation [12]

$$u(t,x) = exp\left[-h^T(x,t)S^{-1}(t)y_t\right]\sigma(t,x), \tag{8}$$

the DMZ equation is transformed into a deterministic PDE with stochastic coefficients

$$\begin{cases} \begin{aligned} \frac{\partial u}{\partial t}(t,x) =& \frac{1}{2}\sum_{i,j=1}^n G_{ij}(t)\frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) - \sum_{i=1}^n f_i \frac{\partial u}{\partial x_i}(t,x) \\ &+ \Bigg(-\frac{\partial}{\partial t}\left(h^T S^{-1}\right)^T y_t \\ &\quad + \frac{1}{2}\sum_{i,j=1}^n G_{ij}(t)\left[\frac{\partial^2 \tilde{K}}{\partial x_i \partial x_j} + \frac{\partial \tilde{K}}{\partial x_i}\frac{\partial \tilde{K}}{\partial x_j}\right] \\ &\quad - \sum_{i=1}^n f_i \frac{\partial \tilde{K}}{\partial x_i}(t,x) - \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(t,x) \\ &\quad - \frac{1}{2}\left(h^T S^{-1} h\right)\Bigg) u(t,x) \\ u(0,x) =& \sigma_0(x), \end{aligned} \end{cases} \tag{9}$$

in which

$$\tilde{K}(x,t) = h^T(x,t)S^{-1}(t)y_t. \tag{10}$$

We shall call (9) "pathwise-robust" DMZ equation in this paper. However, the exact solution to (9), generally speaking, does not have a closed form. Therefore, many mathematicians try to seek an efficient algorithm to construct a good approximation. Let us assume the observations arrive at discrete instants, therefore we construct the approximation as in [9] and get the robust DMZ equation (11) in each time interval.

Let us denote the observation time sequence as $\mathcal{P}_k = \{0 = \tau_0 < \tau_1 < \cdots < \tau_N = T\}$. Let u_k be the solution of the robust DMZ equation with $y_t = y_{\tau_{k-1}}$ on the time interval $\tau_{k-1} \le t \le \tau_k, \ k = 1, 2, \cdots, N$,

$$\begin{cases} \begin{aligned} \frac{\partial u_k}{\partial t}(t,x) =& \frac{1}{2}\sum_{i,j=1}^n G_{ij}(t)\frac{\partial^2 u_k}{\partial x_i \partial x_j}(t,x) - \sum_{i=1}^n f_i \frac{\partial u_k}{\partial x_i}(t,x) \\ &+ \Bigg(-\frac{\partial}{\partial t}\left(h^T S^{-1}\right)^T y_{\tau_{k-1}} \\ &\quad + \frac{1}{2}\sum_{i,j=1}^n G_{ij}(t)\left[\frac{\partial^2 K}{\partial x_i \partial x_j} + \frac{\partial \tilde{K}}{\partial x_i}\frac{\partial \tilde{K}}{\partial x_j}\right] \\ &\quad - \sum_{i=1}^n f_i \frac{\partial \tilde{K}}{\partial x_i}(t,x) - \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(t,x) \\ &\quad - \frac{1}{2}\left(h^T S^{-1} h\right)\Bigg) u_k(t,x) \\ u_1(0,x) =& \sigma_0(x), \\ u_k(\tau_{k-1},x) =& u_{k-1}(\tau_{k-1},x), \ k = 2, 3, \cdots N. \end{aligned} \end{cases} \tag{11}$$

with

$$\tilde{K}(x,t) = h^T(x,t)S^{-1}(t)y_{\tau_{k-1}}. \tag{12}$$

Define the norm of $\mathcal{P}_k$ by $|\mathcal{P}_k| = \sup_{1\le k\le N}(\tau_k - \tau_{k-1})$. By [20], we know that in both point-wise sense and L^2 sense,

$$u(\tau,x) = \lim_{|\mathcal{P}_k|\to 0} u_k(\tau,x). \tag{13}$$

Therefore, $u_k(t,x)$ is a good approximation of $u(t,x)$ in the interval $[\tau_{k-1}, \tau_k]$. We only need to seek the solution of DMZ equation (11).

In [9], the second and third author proposed an on - and off-line algorithm to solve the NLF problems in real time which has been verified numerically as an effective tool in very low dimension. The key observation is that the heavy computation of solving PDE can be moved to off-line by the following proposition.

Proposition 1. *(Proposition 2.1, [9]) For each $\tau_{k-1} \leq t \leq \tau_k,\ k = 1, 2, \cdots, N,\ u_k(t, x)$ satisfies (11) if and only if*

$$\tilde{u}_k(t, x) = \exp\left[h^T(x, t)S^{-1}(t)y_{\tau_{k-1}}\right] u_k(t, x), \tag{14}$$

satisfies the KFE

$$\frac{\partial \tilde{u}_k}{\partial t}(t, x) = \left(L - \frac{1}{2}h^T S^{-1} h\right) \tilde{u}_k(t, x), \tag{15}$$

where L is defined in (2). That is, $\tilde{u}_k(t, x)$ satisfies

$$\begin{cases} \dfrac{\partial \tilde{u}_k}{\partial t}(t, x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2 \tilde{u}_k}{\partial x_i \partial x_j}(t, x) - \sum_{i=1}^{n} f_i \frac{\partial \tilde{u}_k}{\partial x_i}(t, x) \\ \qquad - \left(\displaystyle\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(t, x) + \frac{1}{2}h^T S^{-1} h\right) \tilde{u}_k(t, x) \\ \tilde{u}_1(0, x) = \sigma_0(x), \\ \tilde{u}_k(\tau_{k-1}, x) = \exp\left[h^T(x, \tau_{k-1})S^{-1}(\tau_{k-1})(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] \\ \qquad \cdot \tilde{u}_{k-1}(\tau_{k-1}, x),\ k = 2, 3, \cdots N. \end{cases} \tag{16}$$

III. SCHRÖDINGER EQUATION

As mentioned in the introduction, all the direct methods are for time-invariant Yau system [19]. Even though they include a large class of systems such as time-invariant Kalman-Bucy [7] and Beneš [1] filtering systems, time-varying systems are more general in real applications. In this section, we aim to extend the results to the more general time-varying Yau systems

$$f(x, t) = L(t)x + l(t) + \nabla_x \phi(t, x), \tag{17}$$

where $L(t) = (l_{ij}(t)), 1 \leq i, j \leq n, l^T(t) = (l_1(t), \cdots l_n(t))$ and $\phi(t, x)$ is a C^∞ function on $\mathbb{R}^n$. For the conciseness of notation, we shall omit the t in l and L in the sequel if no confusion will arise.

Proposition 2. *Suppose $\tilde{u}_k(t, x)$ is the solution to (16) in the interval $[\tau_{k-1}, \tau_k],\ k = 1, 2, \cdots N$ and $f(x, t)$ is of the form (17). Let*

$$\tilde{u}_k(t, x) = e^{\phi(t,x)}\tilde{v}_k(t, x), \tag{18}$$

then we have the following equation for $\tilde{v}_k(t, x)$.

$$\begin{cases} \dfrac{\partial \tilde{v}_k}{\partial t}(t, x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2 \tilde{v}_k}{\partial x_i \partial x_j}(t, x) \\ \qquad - (Lx + l)^T \nabla \tilde{v}_k(t, x) - \dfrac{1}{2}q(t, x)\tilde{v}_k(t, x), \\ \tilde{v}_1(0, x) = \sigma_0(x)e^{-\phi(0,x)}, \\ \tilde{v}_k(\tau_{k-1}, x) = \exp\left[h^T(x, \tau_{k-1})S^{-1}(\tau_{k-1})(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] \\ \qquad \cdot \tilde{v}_{k-1}(\tau_{k-1}, x),\ k = 2, 3, \cdots N, \end{cases} \tag{19}$$

where

$$\begin{aligned} q(t, x) = &\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2 \phi}{\partial x_i \partial x_j}(t, x) + \nabla_x \phi^T(t, x)G(t)\nabla_x \phi(t, x) \\ &+ 2(Lx + l)^T \nabla \phi_x(t, x) \\ &+ \sum_{p,l=1}^{n} S_{pl}^{-1}(t)h_p(x, t)h_l(x, t) + 2tr(L). \end{aligned} \tag{20}$$

Proof. Direct computations yield that

$$\frac{\partial \tilde{u}_k}{\partial t}(t, x) = e^{\phi(t,x)}\frac{\partial \tilde{v}_k}{\partial t}(t, x), \tag{21}$$

$$\frac{\partial \tilde{u}_k}{\partial x_i}(t, x) = e^{\phi(t,x)}\left[\frac{\partial \phi}{\partial x_i}(t, x)\tilde{v}_k(t, x) + \frac{\partial \tilde{v}_k}{\partial x_i}(t, x)\right], \tag{22}$$

and

$$\begin{aligned} \frac{\partial^2 \tilde{u}_k}{\partial x_i \partial x_j}(t, x) = &e^{\phi(t,x)}\frac{\partial \phi}{\partial x_j}(t, x)\left[\frac{\partial \phi}{\partial x_i}(t, x)\tilde{v}_k(t, x) + \frac{\partial \tilde{v}_k}{\partial x_i}(t, x)\right] \\ &+ e^{\phi(t,x)}\left[\frac{\partial^2 \phi}{\partial x_i \partial x_j}(t, x)\tilde{v}_k(t, x)\right. \\ &\left. + \frac{\partial \phi}{\partial x_i}(t, x)\frac{\partial \tilde{v}_k}{\partial x_j}(t, x) + \frac{\partial^2 \tilde{v}_k}{\partial x_i \partial x_j}(t, x)\right]. \end{aligned} \tag{23}$$

Put (18), (21)-(23) into (16), we arrive at (19). The initial conditions of $\tilde{v}_k$ follow from those in (16). □

By (13), (14) and (18), it can be easily concluded that the filtering problem for time-varying Yau system (17) becomes solving the Kolmogorov equation (19). In [19], the third author and his co-worker changed the KFE (19) into Schrödinger equation. However, the transformation is much more difficult here since the coefficients G_{ij} in front of the second derivative in (19) are time-varying rather than the identity matrix I. Some assumptions on the system are stated below and the transformation will be introduced in Theorem 1.

Assumption 1. *$G(t)$ is positive definite matrix.*

Since $G(t)$ is positive definite, then we can find a positive definite matrix $F(t) > 0$ such that

$$G(t) = F(t)F^T(t) \tag{24}$$

according to Cholesky decomposition.

Assumption 2. *$L(t)$ in (17) can be expressed as follows:*

$$L(t) = G(t)\Omega(t) + \frac{dF(t)}{dt}F^{-1}(t) \tag{25}$$

where $\Omega(t) \in \mathbb{R}^{n\times n}$ is an arbitrary symmetric matrix.

Remark 1. *If the state of system (1) is scalar or the state is a vector and $G(t), L(t)$ are diagonal, it is obvious that Assumption 2 is naturally satisfied.*

Under Assumption 1-2, we introduce a transformation to eliminate the drift term $\nabla\tilde{v}_k(t, x)$ in (19). So that the Schrödinger equation can be naturally connected to the NLF problems later, see details in section IV.

Theorem 1. *Under Assumption 1-2, suppose $\tilde{v}_k(t,x)$ is a solution of (19) and let*

$$\tilde{v}_k(t,x) = e^{x^T D(t) x} v_k(t,z), \tag{26}$$

where

$$\begin{aligned} z &= B(t)x + b(t), \\ B(t) &= F^{-1}(t), \\ b(t) &= \int_0^t B(s) l(s) ds, \end{aligned} \tag{27}$$

and

$$D(t) = \frac{1}{2}\Omega(t). \tag{28}$$

Then $v_k(t,z)$ is the solution of the following equation:

$$\begin{cases} \dfrac{\partial v_k}{\partial t}(t,z) = \dfrac{1}{2}\Delta v_k(t,z) \\ \qquad - \dfrac{1}{2}\tilde{q}\left(t, F(t)z - F(t)b(t)\right) v_k(t,z) \\ v_1(0,x) = \sigma_0(F(0)x) \\ \qquad \cdot \exp\left[-\phi\left(0, F(0)x\right) - \left(F(0)x\right)^T D(0)\left(F(0)x\right)\right] \\ v_k(\tau_{k-1}, x) = \exp\left[h^T(F(\tau_{k-1})x - F(\tau_{k-1})b(\tau_{k-1}), \tau_{k-1}) \right. \\ \qquad \left. S^{-1}(\tau_{k-1})(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] v_{k-1}(\tau_{k-1}, x), \end{cases} \tag{29}$$

$k = 2, 3, \cdots N$, where

$$\begin{aligned} \tilde{q}(t,x) =& q(t,x) + 2x^T \frac{dD(t)}{dt} x - tr\left(G(t)(D(t) + D^T(t))\right) \\ &- x^T\left(D(t) + D^T(t)\right) G(t) \left(D(t) + D^T(t)\right) x \\ &+ 2(L(t)x + l)^T \left(D(t) + D^T(t)\right) x. \end{aligned} \tag{30}$$

Proof. Through direct computations, we have

$$\begin{aligned} \frac{\partial \tilde{v}_k}{\partial t}(t,x) =& e^{x^T D(t) x}\left[x^T \frac{dD(t)}{dt} x v_k(t,z) + \frac{\partial v_k}{\partial t}(t,z) \right. \\ &\left. + \sum_{i,j=1}^n \frac{\partial v_k}{\partial z_i}(t,z)\left(\frac{dB_{ij}(t)}{dt}x_j + \frac{db_i(t)}{dt}\right)\right] \\ =& e^{x^T D(t) x}\left[x^T \frac{dD(t)}{dt} x v_k(t,z) + \frac{\partial v_k}{\partial t}(t,z) \right. \\ &\left. + \left(\frac{dB(t)}{dt}x + \frac{db(t)}{dt}\right)^T \nabla v_k(t,z)\right], \end{aligned} \tag{31}$$

$$\begin{aligned} \frac{\partial \tilde{v}_k}{\partial x_i}(t,x) =& e^{x^T D(t) x}\left[\sum_{l=1}^n (D_{il} + D_{li}) x_l v_k(t,z) \right. \\ &\left. + \sum_{l=1}^n \frac{\partial v_k}{\partial z_l}(t,z) b_{li}\right] \\ =& e^{x^T D(t) x}\left[\sum_{l=1}^n (D_{il} + D_{li}) x_l v_k(t,z) \right. \\ &\left. + \left(B^T(t)\nabla v_k(t,z)\right)_i\right], \end{aligned} \tag{32}$$

and

$$\begin{aligned} \frac{\partial^2 \tilde{v}_k}{\partial x_i \partial x_j}(t,x) =& e^{x^T D(t) x}\sum_{l=1}^n (D_{jl} + D_{lj}) x_l \\ &\cdot \left[\sum_{p=1}^n (D_{ip} + D_{pi}) x_p v_k(t,z) \right. \\ &\left. + \left(B^T(t)\nabla v_k(t,z)\right)_i\right] \\ &+ e^{x^T D(t) x}\left[(D_{ij} + D_{ij}) v_k(t,z) \right. \\ &+ \sum_{l=1}^n (D_{il} + D_{li}) x_l \left(B^T(t)\nabla v_k(t,z)\right)_j \\ &\left. + \sum_{p,l=1}^n \frac{\partial^2 v_k}{\partial z_p \partial z_l}(t,z) b_{pi} b_{lj}\right], \end{aligned} \tag{33}$$

where $*_{ij}$ denotes the $ij-$entry of the matrix $*$, $*_i$ denotes the $i-$th element of the vector $*$.

Let us write (32) and (33) compactly:

$$\begin{aligned} \nabla \tilde{v}_k(t,x) =& e^{x^T D(t) x}\left[\left(D(t) + D^T(t)\right) x v_k(t,z) \right. \\ &\left. + B^T(t)\nabla v_k(t,z)\right], \end{aligned} \tag{34}$$

$$\begin{aligned} &\sum_{i,j=1}^n G_{ij}(t) \frac{\partial^2 \tilde{v}_k}{\partial x_i \partial x_j}(t,x) \\ =& e^{x^T D(t) x}\left[\left(D(t)x + D^T(t)x\right)^T G(t) \left(D(t)x + D^T(t)x\right) v_k(t,z) \right. \\ &+ 2\left(D(t)x + D^T(t)x\right)^T B^T(t) \nabla v_k(t,z) \\ &+ tr\left(G(t)(D(t) + D^T(t))\right) v_k(t,z) \\ &\left. + \sum_{i,j=1}^n \sum_{p,l=1}^n B_{il} G_{lp} B_{jp} \frac{\partial^2 v_k}{\partial z_i \partial z_j}(t,z)\right]. \end{aligned} \tag{35}$$

Put (26), (31), (34) and (35) into (19), we can get

$$\begin{aligned} \frac{\partial v_k}{\partial t}(t,z) =& \frac{1}{2} tr\left(G(t)B^T(t)H(v_k(t,z))B(t)\right) \\ &+ \left[\left(B(t)G(t)\left(D(t) + D^T(t)\right) \right.\right. \\ &\left. - B(t)L(t) - \frac{dB(t)}{dt}\right) x \\ &\left. - \left(\frac{db(t)}{dt} + B(t)l(t)\right)\right]^T \nabla v_k(t,z) \\ &- \frac{1}{2}\left[q(t,x) + 2x^T \frac{dD(t)}{dt} x \right. \\ &- x^T (D(t) + D^T(t)) G(t) (D(t) + D^T(t)) x \\ &- tr\left(G(t)(D(t) + D^T(t))\right) \\ &\left. + 2(L(t)x + l(t))^T (D(t) + D^T(t)) x\right] v_k(t,z), \end{aligned} \tag{36}$$

where $H(v_k(t,z))$ is the Hessian matrix of $v_k(t,z)$.

The form

$$\frac{\partial v_k}{\partial t}(t,z) = \frac{1}{2}\Delta v_k(t,z) - \frac{1}{2}\tilde{q}(t,x) v_k(t,z) \tag{37}$$

can be obtained by choosing $B(t), b(t), D(t)$ such that

$$\begin{aligned}&B(t)G(t)B^T(t) = I_{n\times n},\\&\frac{db(t)}{dt} + B(t)l(t) = 0,\\&B(t)G(t)\left(D(t)+D^T(t)\right) - B(t)L(t) - \frac{dB(t)}{dt} = 0.\end{aligned} \tag{38}$$

Then we can easily get $B(t)$ and $b(t)$ from (24) and the first two equations of (38). The last equation of (38) is equivalent to

$$\begin{aligned}D(t)+D^T(t) &= G^{-1}B^{-1}(BL + \frac{dB}{dt})\\&= G^{-1}L + G^{-1}B^{-1}\frac{dB}{dt}\\&= G^{-1}L - G^{-1}\frac{dF}{dt}F^{-1}.\end{aligned} \tag{39}$$

Under Assumption 2 we can get

$$D(t)+D^T(t) = G^{-1}\left(G\Omega + \frac{dF}{dt}F^{-1}\right) - G^{-1}\frac{dF}{dt}F^{-1} = \Omega(t). \tag{40}$$

Without loss of generality, we choose $D(t)$ to be symmetric, i.e. $D(t) = D^T(t) = \frac{1}{2}\Omega(t)$. Eqn. (29) follows from (37) immediately, by noting that $x = B^{-1}(t)z - B^{-1}b(t)$. □

IV. Filtering problem

If $q(t,x)$ in (29) is quadratic in x, then it is called Schrödinger equation. Though it feels very restrictive, it includes Kalman-Bucy [7] and Beneš [1] filtering.

Assumption 3. *$\tilde{q}(t,x)$ defined in (20) is quadratic with respect to (w.r.t.) x.*

Notice that observation term $h_i(x,t)$ can be nonlinear which extends the Kalman-Bucy filtering system. Since $q(t,x)$ is quadratic, $h_i(x,t)$, $1 \le i \le m$, are of linear growth w.r.t. the state x, i.e., $h_i^2(x,t) \le M(t)(1+|x|^2)$ for some $M(t)$ from (20).

Since $\tilde{q}(t,x)$ is quadratic in x by (30) under Assumption 3. Thus we can assume that

$$\tilde{q}(t,x) = x^TQ(t)x + p^T(t)x + r(t). \tag{41}$$

Theorem 2. *Let $K(t,x,y)$ be the fundamental solution of*

$$\frac{\partial v_k}{\partial t}(t,x) = \frac{1}{2}\Delta v_k(t,x) - \frac{1}{2}\tilde{q}\left(t, F(t)x - F(t)b(t)\right)v_k(t,x), \tag{42}$$

where

$$\begin{aligned}&\tilde{q}\left(t, F(t)x - F(t)b(t)\right)\\=&x^TF^T(t)Q(t)F(t)x\\&-\left[2b^T(t)F^T(t)Q(t)F(t) - p^T(t)F(t)\right]x\\&+b^T(t)F^T(t)Q(t)F(t)b(t) - p^T(t)F(t)b(t) + r(t).\end{aligned} \tag{43}$$

Assume the fundamental solution $K(t,x,y)$ can be written as

$$K(t,x,y) = (2\pi t)^{-n/2}exp\left\{x^T\tilde{A}(t)x + x^T\tilde{B}(t)y + y^T\tilde{C}(t)y + \tilde{D}^T(t)x + \tilde{E}^T(t)y + s(t)\right\} \tag{44}$$

where $\tilde{A}(t), \tilde{C}(t)$ are $n\times n$ symmetric matrices, $\tilde{B}(t)$ is a $n\times n$ matrix, and $\tilde{D}(t)$ and $\tilde{E}(t)$ are column $n-$vector. Then the coefficients $\tilde{A}(t) - \tilde{E}(t)$ satisfy the following ODEs:

$$\frac{d\tilde{A}}{dt}(t) = 2\tilde{A}^2(t) - \frac{1}{2}F^T(t)Q(t)F(t) \tag{45}$$

$$\frac{d\tilde{B}}{dt}(t) = 2\tilde{A}(t)\tilde{B}(t) \tag{46}$$

$$\frac{d\tilde{C}}{dt}(t) = \frac{1}{2}\tilde{B}^T(t)\tilde{B}(t) \tag{47}$$

$$\frac{d\tilde{D}}{dt}(t) = 2\tilde{A}(t)\tilde{D}(t) + F^T(t)Q(t)F(t)b(t) - \frac{1}{2}F^T(t)p(t) \tag{48}$$

$$\frac{d\tilde{E}}{dt}(t) = \tilde{B}^T(t)\tilde{D}(t) \tag{49}$$

$$\begin{aligned}\frac{ds}{dt}(t) =& \frac{1}{2}\tilde{D}^T(t)\tilde{D}(t) + tr(\tilde{A}(t))\\&-\frac{1}{2}\left[b^T(t)F^T(t)Q(t)F(t)b(t)\right.\\&\left.-p^T(t)F(t)b(t) + r(t)\right] + \frac{n}{2t}\end{aligned} \tag{50}$$

Proof. The proof is the same as that of Theorem 4 in [19]. □

Since $G(t)$ is C^∞ smooth, $f(x,t)$ and $h(x,t)$ are C^∞ smooth in both state and time, it can be easily concluded that $F(t), b(t), Q(t), p(t)$ and $r(t)$ are analytic. Therefore we shall solve the ODEs (45)-(50) formally by power series method.

Theorem 3. *Under Assumption 1-3, the solution $v_k(t,z)$ in $\tau_{k-1} \le t \le \tau_k$ of (29) with $\tilde{q}\left(t, F(t)x - F(t)b(t)\right)$ in (43) is given by*

$$v_k(t,x) = \int_{\mathbb{R}^n} K(t,x,y)v_k(\tau_{k-1},y)dy, \tag{51}$$

where

$$\begin{aligned}K(t,x,y) =&(2\pi(t-\tau_{k-1}))^{-n/2}\\&\cdot\exp\left\{-\frac{|x-y|^2}{2(t-\tau_{k-1})} + x^T\tilde{A}(t-\tau_{k-1})x\right.\\&+x^T\tilde{B}(t-\tau_{k-1})y + y^T\tilde{C}(t-\tau_{k-1})y\\&+\tilde{D}^T(t-\tau_{k-1})x + \tilde{E}^T(t-\tau_{k-1})y\\&\left.+s(t-\tau_{k-1})\right\},\end{aligned} \tag{52}$$

with $\tilde{A}(t-\tau_{k-1}) = \sum_{\nu=1}^{\infty}\tilde{A}_\nu(t-\tau_{k-1})^\nu, \tilde{B}(t-\tau_{k-1}) = \sum_{\nu=1}^{\infty}\tilde{B}_\nu(t-\tau_{k-1})^\nu, \tilde{C}(t-\tau_{k-1}) = \sum_{\nu=1}^{\infty}\tilde{C}_\nu(t-\tau_{k-1})^\nu, \tilde{D}(t-\tau_{k-1}) = \sum_{\nu=1}^{\infty}\tilde{D}_\nu(t-\tau_{k-1})^\nu, \tilde{E}(t-\tau_{k-1}) = \sum_{\nu=1}^{\infty}\tilde{E}_\nu(t-\tau_{k-1})^\nu, s(t-\tau_{k-1}) = \sum_{\nu=1}^{\infty}s_\nu(t-\tau_{k-1})^\nu, b(t-\tau_{k-1}) = \sum_{\nu=0}^{\infty}b_\nu(t-\tau_{k-1})^\nu, F(t-\tau_{k-1}) = \sum_{\nu=0}^{\infty}F_\nu(t-\tau_{k-1})^\nu, Q(t-\tau_{k-1}) = \sum_{\nu=0}^{\infty}Q_\nu(t-\tau_{k-1})^\nu, p(t-\tau_{k-1}) = \sum_{\nu=0}^{\infty}p_\nu(t-$

$\tau_{k-1})^\nu, r(t-\tau_{k-1}) = \sum_{\nu=0}^{\infty} r_\nu (t-\tau_{k-1})^\nu$, *where*

$$\tilde{A}_{\nu+1} = \frac{2}{\nu+3}\sum_{i=0}^{\nu}\tilde{A}_i\tilde{A}_{\nu-i} - \frac{1}{2(\nu+3)}\sum_{j=0}^{\nu}\sum_{i=0}^{j}F_i^TQ_{j-i}F_{\nu-j}, \tag{53}$$

$$\tilde{B}_{\nu+1} = \frac{2}{\nu+2}\sum_{i=0}^{\nu+1}\tilde{A}_i\tilde{B}_{\nu-i}, \tag{54}$$

$$\tilde{C}_{\nu+1} = \frac{1}{2(\nu+1)}\sum_{i=-1}^{\nu+1}\tilde{B}_i^T\tilde{B}_{\nu-i}, \tag{55}$$

$$\tilde{D}_{\nu+1} = \frac{2}{\nu+2}\sum_{i=0}^{\nu+1}\tilde{A}_i\tilde{D}_{\nu-i} - \frac{1}{2(\nu+2)}\sum_{i=0}^{\nu}F_i^Tp_{\nu-i} - \frac{1}{2(\nu+2)}\sum_{j=0}^{\nu}\sum_{i=0}^{j}\sum_{l=0}^{i}F_l^TQ_{i-l}F_{j-i}b_{\nu-j}, \tag{56}$$

$$\tilde{E}_{\nu+1} = \frac{2}{\nu+1}\sum_{i=-1}^{\nu+1}\tilde{B}_i\tilde{D}_{\nu-i}, \tag{57}$$

$$s_{\nu+1} = \frac{1}{2(\nu+1)}\sum_{i=-1}^{\nu+1}\tilde{D}_i^T\tilde{D}_{\nu-i} + \frac{1}{\nu+1}tr(\tilde{A}_\nu) - \frac{1}{2(\nu+1)}\left[\sum_{i=0}^{\nu}\sum_{j=0}^{i}\sum_{m=0}^{j}\sum_{l=0}^{m}b_l^TF_{m-l}^TQ_{j-m}F_{i-j}b_{\nu-i} - \sum_{j=0}^{\nu}\sum_{i=0}^{j}p_i^TF_{j-i}b_{\nu-j} + r_\nu\right], \tag{58}$$

with

$$\begin{aligned}\tilde{A}_{-1} = \tilde{C}_{-1} = -\frac{1}{2}I,\ \tilde{B}_{-1} = I,\\ \tilde{D}_{-1} = \tilde{E}_{-1} = s_{-1} = 0,\\ \tilde{A}_0 = \tilde{B}_0 = \tilde{C}_0 = \tilde{D}_0 = \tilde{E}_0 = s_0 = 0.\end{aligned} \tag{59}$$

Proof. Suppose that all matrices in (45)-(50) can be expanded as follows:

$$\begin{aligned}&\tilde{A}(t) = \sum_{i=-1}^{\infty}\tilde{A}_it^i, \quad \tilde{B}(t) = \sum_{i=-1}^{\infty}\tilde{B}_it^i, \quad \tilde{C}(t) = \sum_{i=-1}^{\infty}\tilde{C}_it^i,\\ &\tilde{D}(t) = \sum_{i=-1}^{\infty}\tilde{D}_it^i, \quad \tilde{E}(t) = \sum_{i=-1}^{\infty}\tilde{E}_it^i, \quad s(t) = \sum_{i=-1}^{\infty}s_it^i,\\ &b(t) = \sum_{i=0}^{\infty}b_it^i, \quad F(t) = \sum_{i=0}^{\infty}F_it^i, \quad Q(t) = \sum_{i=0}^{\infty}Q_it^i,\\ &p(t) = \sum_{i=0}^{\infty}p_it^i, \quad r(t) = \sum_{i=0}^{\infty}r_it^i.\end{aligned} \tag{60}$$

Put (60) into (45)-(50), we can easily know that:

1) Equation (45) is equivalent to

$$-\tilde{A}_{-1} = 2\tilde{A}_{-1}^2 \tag{61}$$

$$0 = 2(\tilde{A}_{-1}\tilde{A}_0 + \tilde{A}_0\tilde{A}_{-1}) \tag{62}$$

$$(\nu+1)\tilde{A}_{\nu+1} = 2\sum_{i=-1}^{\nu+1}\tilde{A}_i\tilde{A}_{\nu-i} - \frac{1}{2}\sum_{j=0}^{\nu}\sum_{i=0}^{j}F_i^TQ_{j-i}F_{\nu-j},\ \nu \ge 0 \tag{63}$$

2) Equation (46) is equivalent to

$$-\tilde{B}_{-1} = 2\tilde{A}_{-1}\tilde{B}_{-1} \tag{64}$$

$$0 = 2(\tilde{A}_{-1}\tilde{B}_0 + \tilde{A}_0\tilde{B}_{-1}) \tag{65}$$

$$(\nu+1)\tilde{B}_{\nu+1} = 2\sum_{i=-1}^{\nu+1}\tilde{A}_i\tilde{B}_{\nu-i},\ \nu \ge 0 \tag{66}$$

3) Equation (47) is equivalent to

$$-\tilde{C}_{-1} = \frac{1}{2}\tilde{B}_{-1}^T\tilde{B}_{-1} \tag{67}$$

$$0 = \frac{1}{2}(\tilde{B}_{-1}^T\tilde{B}_0 + \tilde{B}_0^T\tilde{B}_{-1}) \tag{68}$$

$$(\nu+1)\tilde{C}_{\nu+1} = \frac{1}{2}\sum_{i=-1}^{\nu+1}\tilde{B}_i^T\tilde{B}_{\nu-i},\ \nu \ge 0 \tag{69}$$

4) Equation (48) is equivalent to

$$-\tilde{D}_{-1} = 2\tilde{A}_{-1}\tilde{D}_{-1} \tag{70}$$

$$0 = 2(\tilde{A}_{-1}\tilde{D}_0 + \tilde{A}_0\tilde{D}_{-1}) \tag{71}$$

$$(\nu+1)\tilde{D}_{\nu+1} = 2\sum_{i=-1}^{\nu+1}\tilde{A}_i\tilde{D}_{\nu-i} - \frac{1}{2}\sum_{i=0}^{\nu}F_i^Tp_{\nu-i} - \frac{1}{2}\sum_{j=0}^{\nu}\sum_{i=0}^{j}\sum_{l=0}^{i}F_l^TQ_{i-l}F_{j-i}b_{\nu-j},\ \nu \ge 0 \tag{72}$$

5) Equation (49) is equivalent to

$$-\tilde{E}_{-1} = \tilde{B}_{-1}^T\tilde{D}_{-1} \tag{73}$$

$$0 = (\tilde{B}_{-1}^T\tilde{D}_0 + \tilde{B}_0^T\tilde{D}_{-1}) \tag{74}$$

$$(\nu+1)\tilde{E}_{\nu+1} = 2\sum_{i=-1}^{\nu+1}\tilde{B}_i\tilde{D}_{\nu-i},\ \nu \ge 0 \tag{75}$$

6) Equation (50) is equivalent to

$$-s_{-1} = \frac{1}{2}\tilde{D}_{-1}^T\tilde{D}_{-1} \tag{76}$$

$$0 = \frac{1}{2}(\tilde{D}_{-1}^T\tilde{D}_0 + \tilde{D}_0^T\tilde{D}_{-1}) + tr(\tilde{A}_{-1}) + \frac{n}{2} \tag{77}$$

$$(\nu+1)s_{\nu+1} = \frac{1}{2}\sum_{i=-1}^{\nu+1}\tilde{D}_i^T\tilde{D}_{\nu-i} + tr(\tilde{A}_\nu) - \frac{1}{2}\left[\sum_{i=0}^{\nu}\sum_{j=0}^{i}\sum_{m=0}^{j}\sum_{l=0}^{m}b_l^TF_{m-l}^TQ_{j-m}F_{i-j}b_{\nu-i} - \sum_{j=0}^{\nu}\sum_{i=0}^{j}p_i^TF_{j-i}b_{\nu-j} + r_\nu\right],\ \nu \ge 0 \tag{78}$$

TABLE I
NUMERICAL IMPLEMENTATION OF DIRECT METHOD

Algorithm numerical implementation of direct method for (1)

Initialization: give $T_0, T, \Delta, \sigma_0(x), M \geq 0$
Calculate $N = (T - T_0)/\Delta$
Calculate $F(t), B(t), b(t), D(t)$ by (24),(25),(27),(28)
Calculate $Q(t), p(t), r(t)$ by (20),(30),(41)
Calculate $\tilde{A}_\nu, \tilde{B}_\nu, \tilde{C}_\nu, \tilde{D}_\nu, \tilde{E}_\nu, s_\nu, 0 \leq \nu < M$ by (53)-(58)
Calculate $\hat{x}_{t_0}$
for $k = 1$ to N **do**
  Calculate $v_k(t_{k-1}, x), v_k(t_k, x)$ by (29),(51)
  Calculate $\tilde{v}_k(t_k, x), \tilde{u}_k(t_k, x)$ by (26),(18)
  Calculate $u_k(t_k, x), \sigma(t_k, x)$ by (14),(8)
  Calculate $\hat{x}_{t_k}$
  Assign $k := k + 1$
end for

$[T_0, T]$ is the appointed time period, Δ is discrete time step size, M is the assumed truncate order, and $\sigma_0(x)$ is the assumed probability density of the initial sate x_0.

The verification of $K(t, x, y)$ in (52) is a fundamental solution of (29) is the same as that of Theorem 5 in [19]. □

V. Simulations

In this section, we use an example to show the efficiency of the proposed direct method and compare it with EKF.

A. Algorithm

To implement the proposed direct method numerically, we need to truncate the higher order in (60), which means that we only need to compute $\tilde{A}_\nu, \tilde{B}_\nu, \tilde{C}_\nu, \tilde{D}_\nu, \tilde{E}_\nu, s_\nu, 0 \leq \nu < M$ by (53)-(58) where M is the assumed order. The numerical procedure of direct method for nonlinear filtering problem (1) is listed in TABLE I.

B. Numerical example

The numerical example considered here is as follows:

$$\begin{cases} \quad dx_t = \dfrac{t}{40} \cdot x_t dt + dv_t \\ dy_1(t) = x_t \sin x_t dt + dw_1(t) \\ dy_2(t) = x_t \cos x_t dt + dw_2(t), \end{cases} \tag{79}$$

where $x_t \in \mathbb{R}$ is state with initial state x_0. Here the true initial state x_0 has been chosen to be 0.1, $v_t \in R$ is standard Brownian motion, and $y(t) = [y_1(t), y_2(t)]^T \in \mathbb{R}^2$ is the two dimensional measurement, $w_t = [w_1(t), w_2(t)]^T$ is the two dimensional standard Brownian motion.

Numerical simulations are obtained through the construction of an exact stochastic realization of system (79) at discrete times $t_k = k\Delta$ with $\Delta = 0.01$ on the interval $[0, T]$ with $T = 10$ according to Euler-Maruyama method. All ODEs and integrals in the simulations by two methods are solved by Euler method. The initial values for EKF are $\hat{x}_0$ and P_0 and $\sigma_0(x)$ is Gaussian distribution $\mathcal{N}(x_0, P_0)$ and the integral step size for direct method is $h = 0.05$.

We shall take different initial to examine the effectiveness of our direct method. The initial value $\hat{x}_0$ is generated by posing various P_0. We observe that when $P_0 = 3$, EKF blows up 8 times in 100 simulations and this number becomes 80 when $P_0 = 10$, while our method seems not so sensitive as EKF.

TABLE II
AVERAGE OF MSE BY DIRECT METHOD AND EKF

Algorithm	Direct method	EKF
$P_0 = 0.1$	1.9671	7.1694
$P_0 = 3$	2.1196	——
$P_0 = 10$	2.2753	——

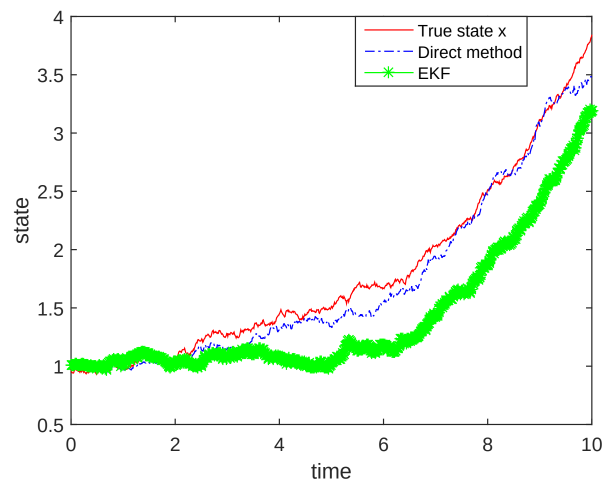

Fig. 1. Estimation results of state x with $\hat{x}_0 \sim \mathcal{N}(x_0, 0.1)$

We define the mean of the squared estimation error (MSE) for one realization

$$\mu_x = \frac{1}{N+1} \sum_{k=0}^{N} (x_{t_k} - \hat{x}_{t_k})^2, \tag{80}$$

and the average of MSE over 100 simulations for different methods are listed in Table II. The average estimation results of EKF with $P_0 = 3, 10$ cannot be obtained due to the blow up.

The average results over 100 simulations of different methods are displayed in Fig. 1-3. It can be clearly seen that the proposed direct method always performs better than the EKF and our method can trace the real state well even with large P_0 which means that the estimation initial value $\hat{x}_0$ is far from the real initial value x_0. Furthermore, comparing estimation results of direct method in Fig. 1-3, we can observe that $\hat{x}_0$ only effects the initial estimation result and direct method has the same performance with different $\hat{x}_0$ after some time. This property is very useful, since in real applications, we can hardly know the real initial state and P_0 can be very large, so the direct method is superior from this point of view.

VI. Conclusion

In this paper, we extend the algorithm developed in [19] to solve the NLF problems for the time-varying Yau filtering system with arbitrary initial conditions. Under some mild assumptions, we get the corresponding time-varying Schrödinger equation by the transformation of the DMZ equation, and then we write down its fundamental solution in terms of the solution

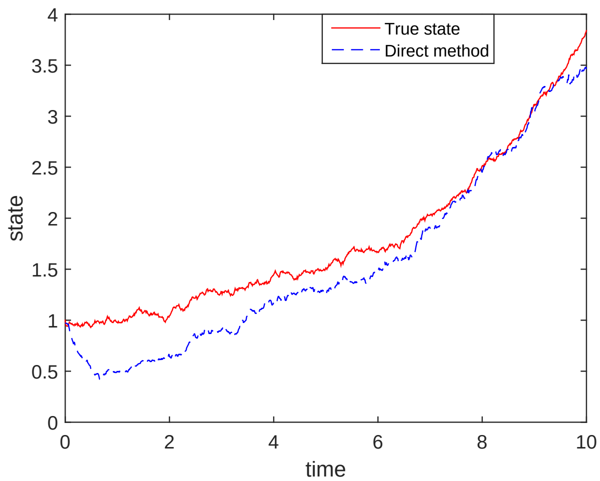

Fig. 2. Estimation results of state x with $\hat{x}_0 \sim \mathcal{N}(x_0, 3)$

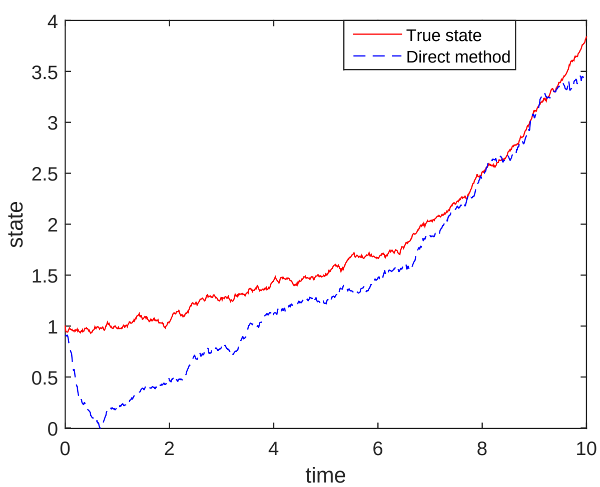

Fig. 3. Estimation results of state x with $\hat{x}_0 \sim \mathcal{N}(x_0, 10)$

of a system of ODEs, which are solved by power series method when the potential is quadratic in state. Besides, the numerical simulation shows the efficiency of the proposed method.

VII. Acknowledgements

This paper is sponsored by National Natural Science Foundation of China (11471184). X. Luo acknowledges the support from the Fundamental Research Funds for the Central Universities (YWF-16-SXXY-001) and National Natural Science Foundation of China (11501023). S. S.-T. Yau thanks the financial support of start-up fund from Tsinghua University.

References

[1] Beneš, V. Exact finite dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, 5 (1981), 65-92.

[2] Chen, J. On ubiquity of Yau filters. *In Proceedings of the American Control Conference*, Baltimore, MD, (June 1994), 252-254.

[3] Duncan, T. E. *Probability density for diffusion processes with applications to nonlinear filtering theory and diffusion theory*. Ph.D. dissertation, Stanford University, Stanford, CA, 1967.

[4] Hu, G-Q., Yau, S. S-T., and Chiou, W-L. Finite dimensional filters with nonlinear drift XIII: Classification of finite-dimensional estimation algebras of maximal rank with state space dimension less than or equal to five. *Asian Journal of Mathematics*, 4.4 (2000), 905-932.

[5] Hu, G-Q., and Yau, S. S-T. Finite dimensional filters with nonlinear drift XV: New direct method for construction of universal finite dimensional filter. *IEEE Transactions on Aerospace and Electronic Systems*, 38.1 (2002), 50-57.

[6] Kalman, R. E. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1960), 35-45.

[7] Kalman, R. E. and Bucy, R. S. New results in linear filtering and prediction theory. *Transactions of the Asme-Journal of Basic Engineering*, 83(1961), 95-108.

[8] Karatzas, I., and Shreve, S. E. *Brownian motion and stochastic calculus*. World Scientific, (2006).

[9] Luo, X. and Yau, S. S.-T. Complete real time solution of the general nonlinear filtering problem with out memory. *IEEE Transactions on Automatic Control*, 58.10 (2013), 2563-2578.

[10] Luo, X. On recent advance of nonlinear filtering theory: emphases on global approaches. *Pure & Applied Mathematics Quarterly*, 10.4 (2014), 685-721.

[11] Mortensen, N. E. *Optimal control of continuous-time stochastic systems*. Ph.D. dissertation, University of California, Berkeley, CA, (1966).

[12] Rozovsky, B. L. Stochastic partial differential equations arising in nonlinear filtering problems. *Uspekhi Matematicheskikh Nauk*, 27 (1972), 213-214.

[13] Yau, S. S-T., and Yau, S-T. Existence and decay estimates for time dependent parabolic equation with application to Duncan-Mortensen-Zakai equation. *IEEE Transactions on Information Theory,* 24.5 1978, 649-650.

[14] Yau, S. S-T., and Yau, S-T. New direct method for Kalman-Bucy filtering system with arbitrary initial condition. *In Proceedings of the 33rd Conference on Decision and Control*, Lake Buena Vista, FL, Dec. 14-16, 1994, 1221-1225.

[15] Yau, S. S-T., and Hu, G-Q. Direct method without Riccati equation for Kalman-Bucy filtering system with arbitrary initial conditions. *In Proceedings of the 13th World Congress IFAC*, vol. H, San Francisco, CA, June 30-July 5, 1996, 469-474.

[16] Yau, S. S-T., and Yau, S-T. Finite dimensional filters with nonlinear drift III: Duncan-Mortensen-Zakai equation with arbitrary initial condition for linear filtering system and the Beneš filtering system. *IEEE Transactions on Aerospace and Electronic Systems*, 33 (Oct. 1997), 1277-1294.

[17] Yau, S. S-T. Brocketts problem on nonlinear filtering theory. *In Lectures on Systems, Control and Information, AMS/IP, Studies in Advanced Mathematics*, 17 (2000), 177-212.

[18] Yau, S. S-T. and Hu, G-Q. Finite dimensional filters with nonlinear drift X: Explicit solution of DMZ equation. *IEEE Transactions on Automatic Control*, 46.1 (Jan. 2001), 142-148.

[19] Yau, S. T. and Yau, S.S.-T. Nonlinear filtering and time varying Schrödinger equation I. *IEEE Transactions on Aerospace and Electronic Systems*, 40.1 (2004), 284-292.

[20] Yau, S. S.-T. Solution of filtering problem with nonlinear observations. *Mathematische Nachrichten*, 10.3-4(2006), 187-196.

[21] Zakai M. On the optimal filtering of diffusion processes, *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete,* 11 (1969), 230-243.

Xiuqiong Chen received the B.S. degree in School of Mathematics and Systems Science, Beihang University, Beijing, China, in 2014. She is currently pursuing the Ph.D. degree in Applied Mathematics from the Department of Mathematical Sciences, Tsinghua University, Beijing, China.

Her research interests include nonlinear filtering, active disturbance rejection control.

Xue Luo (M' 15, SM' 16) 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. 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.

Stephen S.-T. Yau (F' 03) received the Ph.D. degree in mathematics from the State University of New York at Stony Brook, NY, USA in 1976.

He was a Member of the Institute of Advanced Study at Princeton from 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, During 2005-2011, he became a joint Professor with the Department of Electrical and Computer Engineering at the MSCS, UIC. After his retirement in 2012, he joined Tsinghua University, Beijing, China, where he is a full-time professor in the Department of Mathematical Sciences. His research interests include nonlinear filtering, bioinformatics, complex algebraic geometry, CR geometry and sigularities theory.

Dr. Yau is the Managing Editor and founder of the *Journal of Algebraic Geometry* since 1991, and the Editor-in-Chief and founder of *Communications in Information and Systems* from 2000 to the present. He was the General Chairman of the IEEE International Conference on Control and Information, which was held in the Chinese University of Hong Kong in 1995. He was awarded the Sloan Fellowship in 1980, the Guggenheim Fellowship in 2000, and the AMS Fellow Award in 2013. In 2005, he was entitled the UIC Distinguished Professor.

Figure captions:

Fig. 1 Estimation results of state x with $\hat{x}_0 \sim \mathcal{N}(x_0, 0.1)$

Fig. 2 Estimation results of state x with $\hat{x}_0 \sim \mathcal{N}(x_0, 3)$

Fig. 3 Estimation results of state x with $\hat{x}_0 \sim \mathcal{N}(x_0, 10)$