Communications in Information and Systems
Volume 19, Number 2, 167–191, 2019

A survey of direct methods for Yau filtering systems

Xiuqiong Chen* and Wenhui Dong

Filtering problem has been studied since the early 1960s when the famous Kalman filter was proposed and there have been a variety of filtering algorithms springing up since then. One useful filtering method is the so called direct method for Yau filtering system which was proposed in 1990s. In the early works about the direct method, they all assume that the filtering system is time-invariant. However, this limitation excludes the situation where the filtering system also depends on the time, i.e., the filtering system is time-varying, which can be very common in real applications. Recently, Yau and his collaborators have extended the work to the time-varying case and make it more applicable for practical filtering problems. In this work, we shall briefly go through the direct method for both the time-invariant and the time-varying Yau filtering system.

1. Introduction

The aim of the filtering is to estimate the state of a stochastic dynamical system with observations corrupted by the noises and it has various applications in many applied fields such as communication, aerospace applications, and economics. About two centuries ago Gauss started the work about filtering problems, and later Wiener and Kalman made groundbreaking work in filtering theory. In 1960s, the most influential work in filtering theory are published including the classical Kalman filter (KF) [17] and its continuous counterpart Kalman-Bucy filter [18]. Both Kalman filter and Kalman-Bucy filter are only applicable to linear systems and most systems are nonlinear in real applications. Therefore there have been arising a lot of work applicable to the nonlinear filtering (NLF) problems, including the extended Kalman filter (EKF) [16], ensemble Kalman filter (EnKF) [13], unscented Kalman filter (UKF) [29, 33], particle filter (PF) [11, 14] and other methods [23, 25, 32]. When the dynamic system is significantly nonlinear, it is known that EKF always performs poorly and besides, it is very sensitive to initial value on account of the Taylor approximation. EnKF, which combines

*Corresponding author.

the data assimilation and ensemble generation problem, has lots of applications in numerical weather and ocean prediction applications [1, 2, 19]. PF is also one of the most popular methods nowadays and has been extended to different models, see [3, 4] and references therein. PF is applicable to nonlinear, non-Gaussian state update and observation equations, and can obtain asymptotically optimality as the number of particles goes to infinity. However, it is hard to be implemented in real time owing to its essence of Monte Carlo simulation.

All the filtering methods mentioned above only seek the approximation of conditional mean and variance, while another way to NLF problem is to derive the conditional probability of the state which is called the global approach [22]. Obviously all the statistical information, besides conditional mean and variance, can be acquainted by global approach. It is known that the unnormalized probability density function of the state satisfies the Duncan-Mortensen-Zakai (DMZ) equation [12, 27, 44]. However, it is not easy to solve the DMZ equation directly. The difficulty is that DMZ equation is a parabolic equation with coefficients containing observations. In 2008, Yau and Yau made a breakthrough in this problem by reducing the DMZ equation to forward Kolmogorov equation (FKE) which can be solved off-line [43]. Following this work, Luo and Yau proposed an algorithm [20, 21] to solve general NLF problems using DMZ equations in real-time manner. "Real-time" means that the estimation of the states is made on the spot instantaneously, while the observation data keep coming in. Thereafter they also proposed other numerical method to solve FKE with time-varying parameters equation [24]. We refer interested readers for the survey paper [22].

Though it is not possible to solve explicitly the DMZ equation in most situation, there are two methods have been found to the best of our knowledge for the past quarter of a century. One is to use Lie algebraic method proposed by Brockett [6] and Mitter [26], and Yau worked out the details of this method in [34]. The basic idea is that solving the DMZ equation is transformed into solving a series of ordinary differential equations (ODE), FKE, and some first-order linear partial differential equations (PDE). However, the basis of the estimation algebra must be known in this method. Yau and his co-workers [10, 31, 36, 41] have completely classified all finite dimensional estimation algebras of maximal rank. In particular, they have proved that for all finite dimensional filters, the observation terms $h_i(x), 1 \leq i \leq m$ in (1), must be polynomials of degree one.

The direct method is the other approach to solve DMZ equation which works well especially for the Yau filtering system, i.e., $f(x,t)$ in (1) is of the form $f(x,t) = Lx + l + \nabla\phi(x)$ where L, l are constant matrices with proper dimensions and $\phi(x)$ is a C^∞ function. This method was introduced in [35] and generalized in [15, 37, 38]. Comparing with the Lie algebra method, direct method does not need to integrate several first-order linear PDEs. Nevertheless in [15, 35, 37] and [38], they need to assume that the observation terms $h_i(x), 1 \leq i \leq m$ in (1) are degree one polynomials. In [40], Yau and Yau transformed the DMZ equation to time varying Schrödinger equation in very general cases where observations terms are of linear growth. In [39], Yau and Lai solved DMZ equation by solving a series of ODEs when the initial distribution is Gaussian. Based on the work of Yau and Lai [39], Shi and Yau [30] proposed a useful Gaussian approximation algorithm such that each initial distribution can be decomposed into the sum of several Gaussian distributions. Therefore based on each Gaussian approximation of the initial condition, the Kolmogorov equation can be solved in terms of ODEs.

However, all these existing direct methods are for time-invariant systems and they need to assume that $g(t)\tilde{Q}(t)g^T(t)$ in (1) is an identity matrix. Recently, under some mild assumptions on the filtering system, we extend the related results to time-varying situations and make it more practicable in real applications [8, 9].

This survey paper is aim to present various direct methods studied in the literature with the emphases on the [8, 9, 30, 38]. Furthermore, we discuss and compare three different methods from the model and assumptions. The method in [38] reduces the DMZ equation into one Kolmogorov equation and several ODEs with respect to (w.r.t.) time-invariant Yau filtering system with assumption that the observation is linear. The second one [30] also considers the time-invariant system and only needs to assume that the observation is of linear growth. Besides, [30] reduces the DMZ equation to several ODEs by the use of Gaussian approximation. [8, 9] extends the work for time-invariant system to time-varying system and provides the direct computation of the solution to the DMZ equation.

This paper is organized as follows. In Section 2, we recall some basic concepts and existing results with respect to the filtering problem. We present the direct method for time-invariant Yau filtering system in Section 3 and conclude the corresponding result for time-varying system in Section 4. Finally, in the last section we draw the conclusion.

2. Basic concepts

2.1. Basic filtering problems

We consider the following continuous filtering problem:

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

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

In the continuous dynamic system (1), we now 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 first explain some notations here: A_{ij} denotes the ij-entry of an arbitrary matrix A, a_i denotes the i-th element of an arbitrary vector a, and A^T denotes the transposition of A.

Let $\rho(t,x)$ be density function of x_t conditioned on the observation history $\mathscr{F}_t \triangleq \{y_s : 0 \le s \le t\}$, then it must satisfy the normalization condition, i.e.,

$$\int \rho(t,x)dx = 1. \tag{2}$$

Actually, if there is any function $\sigma(t,x)$ satisfying

$$\rho(t,x) \propto \sigma(t,x) \text{ w.r.t. } x, \tag{3}$$

then $\rho(t,x)$ can be computed by normalization:

$$\rho(t,x) = \frac{\sigma(t,x)}{\int \sigma(t,x)dx}. \tag{4}$$

In [12], it is known that the unnormalized density function $\sigma(t,x)$ of x_t

conditioned on $\mathscr{F}_t$ satisfies the DMZ equation:

$$(5)\quad \begin{cases} d\sigma(t,x) = \left[\frac{1}{2}\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2\sigma}{\partial x_i \partial x_j}(t,x) - \sum_{i=1}^{n} f_i \frac{\partial \sigma}{\partial x_i}(t,x) \right. \\ \qquad\qquad \left. -\sigma(t,x)\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}(t,x)\right] dt + \sigma(t,x)h^T(x,t)S^{-1}(t)dy_t \\ \sigma(0,x) = \sigma_0(x), \end{cases}$$

where $\sigma_0(x)$ is the probability density of the initial sate x_0. For each arrived observation, making an invertible exponential transformation [28]:

$$(6)\quad u(t,x) = \exp\left[-h^T(x,t)S^{-1}(t)y_t\right]\sigma(t,x),$$

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

$$(7)\quad \begin{cases} \frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2 u}{\partial x_i \partial x_j}(t,x) + \sum_{i=1}^{n}\left(\sum_{j=1}^{n} G_{ij}(t)\frac{\partial \bar{K}}{\partial x_j} - f_i\right)\frac{\partial u}{\partial x_i}(t,x) \\ \qquad + \left(-\frac{\partial}{\partial t}\left(h^T S^{-1}\right)^T y_t + \frac{1}{2}\sum_{i,j=1}^{n} G_{ij}(t)\left[\frac{\partial^2 \bar{K}}{\partial x_i \partial x_j} + \frac{\partial \bar{K}}{\partial x_i}\frac{\partial \bar{K}}{\partial x_j}\right]\right. \\ \qquad \left. - \sum_{i=1}^{n} f_i \frac{\partial \bar{K}}{\partial x_i}(t,x) - \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}(t,x) - \frac{1}{2}\left(h^T S^{-1} h\right)\right) u(t,x), \\ u(0,x) = \sigma_0(x), \end{cases}$$

in which

$$(8)\quad \bar{K}(x,t) = h^T(x,t)S^{-1}(t)y_t.$$

We shall call (7) "pathwise-robust" DMZ equation in this paper. In general, the exact solution to (7) does not have a closed form. Assuming the observations arrive at discrete instants, then we can construct the approximation as in [21, 43] and get the robust DMZ equation (9) 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 $\tau_{k-1} \le t \le \tau_k,\ k = 1, 2, \cdots, N$,

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

with

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

Define the norm of $\mathcal{P}_k$ by $|\mathcal{P}_k| = \sup_{1\le k\le N}(\tau_k - \tau_{k-1})$. By [42], we know that in both point-wise sense and L^2 sense,

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

Therefore, $u_k(t,x)$ is a good approximation of $u(t,x)$ in the interval $[\tau_{k-1}, \tau_k]$. Then we only need to seek the solution of DMZ equation (9).

In [21], Luo and Yau proposed an on- and off-line algorithm solving the NLF problems in real time and it has been verified numerically as an efficient tool in very low dimension. The key step is that the heavy computation of solving PDE can be moved to off-line using the following proposition.

Proposition 1 ([43]). *For each $\tau_{k-1} \le t \le \tau_k,\ k = 1, 2, \cdots, N,\ u_k(t,x)$ satisfies (9) if and only if*

$$\tilde{u}_k(t,x) = \exp\left[h^T(x,t)S^{-1}(t)y_{\tau_{k-1}}\right] u_k(t,x), \tag{12}$$

satisfies the Kolmogorov forward equation

$$\begin{cases} \dfrac{\partial \tilde{u}_k}{\partial t}(t,x) = & \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2 \tilde{u}_k}{\partial x_i \partial x_j}(t,x) - \sum_{i=1}^{n} f_i \frac{\partial \tilde{u}_k}{\partial x_i}(t,x) \\ & - \left(\displaystyle\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(t,x) + \frac{1}{2}h^T S^{-1} h\right) \tilde{u}_k(t,x), \\ \tilde{u}_1(0,x) = & \sigma_0(x), \\ \tilde{u}_k(\tau_{k-1},x) = & \exp\left[h^T(x,\tau_{k-1})S^{-1}(\tau_{k-1})(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] \\ & \cdot \tilde{u}_{k-1}(\tau_{k-1},x), \;\; k = 2,3,\cdots N. \end{cases} \tag{13}$$

Usually, the filtering system can be divided into two categories including time-invariant system and time-varying system, and (1) is the general form of the time-varying system. When $f(x_t,t), h(x_t,t)$ are only functions w.r.t. x, and $g(t), \tilde{Q}(t), S$ are constants in (1), we call (1) the time-invariant filtering system. More importantly, $G(t) = I$ in the time-invariant case, while $G(t)$ is a time-varying matrix in the time-varying case. Therefore, the Kolmogorov forward equation (13) can be much more complicated in the time-varying case. In the following two sections, we shall show how to get the explicit solution of the DMZ equation w.r.t. different systems.

3. Time-invariant filtering system

The time-invariant form of (1) is as follows:

$$\begin{cases} dx_t = f(x_t)dt + gdv_t \\ dy_t = h(x_t)dt + dw_t \end{cases} \tag{14}$$

where g is assumed to be an orthogonal matrix, covariance $\tilde{Q}, S$ of the noises are identity matrices.

In terms of this time-invariant filtering system (14), the robust DMZ equation (7) is reduced to

$$
(15)\quad \begin{cases}
\dfrac{\partial u}{\partial t}(t,x) = \dfrac{1}{2}\displaystyle\sum_{i=1}^{n} \dfrac{\partial^2 u}{\partial x_i^2}(t,x) + \sum_{i=1}^{n}\left(\dfrac{\partial \tilde{K}}{\partial x_i} - f_i\right)\dfrac{\partial u}{\partial x_i}(t,x) \\
\qquad + \left(\dfrac{1}{2}\displaystyle\sum_{i=1}^{n}\left[\dfrac{\partial^2 \tilde{K}}{\partial x_i^2} + \left(\dfrac{\partial \tilde{K}}{\partial x_i}\right)^2\right] - \sum_{i=1}^{n} f_i \dfrac{\partial \tilde{K}}{\partial x_i}(t,x)\right. \\
\qquad \left. - \displaystyle\sum_{i=1}^{n}\dfrac{\partial f_i}{\partial x_i}(t,x) - \dfrac{1}{2}\left(h^T h\right)\right) u(t,x), \\
u(0,x) = \sigma_0(x).
\end{cases}
$$

Now we need the following conditions before we present the work in [30, 38].

C_1) $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$, where c_{ij} are constants for $1 \le i, j \le n$. This is the so-called Yau filtering system in [7]. This condition has been proved to be equivalent to [38]

$$
(16)\qquad f_i(x) = l_i(x) + \frac{\partial \phi}{\partial x_i}(x),
$$

for $1 \le i \le n$, where $l_i(x) = \sum_{j=1}^{n} d_{ij} x_j + d_i$ and ϕ is a C^∞ function;

C_2) $\sum_{i=1}^{m} h_i(x) = \sum_{i,j=1}^{n} q_{ij} x_j + q_i$;

C_2') $\sum_{i=1}^{m} h_i^2(x) = \sum_{i,j=1}^{n} q_{ij} x_i x_j + \sum_{i=1}^{n} q_i x_i + q_0$;

C_3) $\eta(x) = \sum_{i,j=1}^{n} \eta_{ij} x_i x_j + \sum_{i=1}^{n} \eta_i x_i + \eta_0$, where $\eta(x)$ is defined as follows:

$$
(17)\qquad \eta(x) = \sum_{i=1}^{n} f_i^2(x) + \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(x) + \sum_{i=1}^{m} h_i^2(x);
$$

where $d_{ij}, d_i, q_{ij} = q_{ji}, q_i, q_0, \eta_{ij}, \eta_i, \eta_0, 1 \le i, j \le n$ are constants.

3.1. Reducing the DMZ equation to Kolmogorov equation and ODEs

In the work of [38], Yau and his collaborator proposed the direct method to solve the DMZ equation (15) by reducing it to the Kolmogorov equation and ODEs which is stated in the following theorem.

Theorem 1 (DMZ equation → Kolmogorov equation + ODEs, Theorem 3.3 in [38]). *Consider the nonlinear system (14) with conditions C_1), C_2)*

and C_3). Then the solution $u(t,x)$ for the DMZ equation (15) is reduced to the solution $\tilde{u}(t,x)$ for the following Kolmogorov equation:

$$(18)\qquad \begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t,x) = & \dfrac{1}{2}\Delta \tilde{u}(t,x) - \displaystyle\sum_{i=1}^{n} f_i(x)\frac{\partial \tilde{u}}{\partial x_i}(t,x) \\ & + \dfrac{1}{2}\left(\displaystyle\sum_{i=1}^{n} f_i^2(x) - \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}(x) - \eta(x)\right)\tilde{u}(t,x) \\ \tilde{u}(0,x) = & \sigma_0(x) \end{cases}$$

where

$$(19)\quad \tilde{u}(t,x) = \exp\left[c(t) + \sum_{i=1}^{n} a_i(t)x_i + \phi(x) - \phi(x+b(t))\right]\cdot u(t, x+b(t))$$

and $a_i(t), b_i(t)$ and $c(t)$ satisfy ODEs:

$$(20)\qquad \begin{cases} b_i'(t) - a_i(t) - \displaystyle\sum_{j=1}^{n} d_{ij}b_j(t) + \sum_{j=1}^{m} c_{ji}y_j(t) = 0 \\ b_i(0) = 0, \end{cases}$$

$1 \le i \le n$.

$$(21)\qquad \begin{cases} a_i'(t) - \dfrac{1}{2}\displaystyle\sum_{j=1}^{n}(\eta_{ij}+\eta_{ji})b_j(t) + \sum_{j=1}^{n} d_{ji}b_j'(t) = 0 \\ a_i(0) = 0, \end{cases}$$

$1 \le i \le n$.

$$(22)\qquad \begin{cases} c'(t) = & -\dfrac{1}{2}\displaystyle\sum_{i=1}^{n}(b_i'(t))^2 + \sum_{i=1}^{n} a_i(t)b_i'(t) - \sum_{i=1}^{n} d_i(t)b_i'(t) \\ & + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n}\eta_{ij}b_i(t)b_j(t) + \frac{1}{2}\sum_{i=1}^{n}\eta_i b_i(t) \\ c(0) = & 0. \end{cases}$$

3.2. Reducing the DMZ equation to ODEs

However, [38] only considers the system with condition C_2), i.e., the observation must be linear. Recently, [30] considers a less constrained condition

C_2') and can reduce the computation of the DMZ equation to the solutions of the ODEs. Similarly, the FKE (13) for time-invariant system (14) is reduced to:

$$\begin{aligned}\frac{\partial \tilde{u}_k}{\partial t}(t,x) &= \frac{1}{2}\Delta \tilde{u}_k(t,x) - \sum_{i=1}^{n} f_i(x)\frac{\partial \tilde{u}_k}{\partial x_i}(t,x) \\ &\quad - \left(\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^{m} h_i^2\right)\tilde{u}_k(t,x)\end{aligned} \tag{23}$$

with

$$u_k(t,x) = \exp\,(-\sum_{i=1}^{m} y_i(\tau_{k-1})h_i(x))\tilde{u}_k(t,x). \tag{24}$$

Based on (23), [30] first transforms the Kolmogorov equation (23) of $\tilde{u}_k(t,x)$ into another Kolmogorov equation of $\hat{u}_k(t,x)$ in Theorem 2, and then reduces the computation of the Kolmogorov equation into ODEs with arbitrary initial conditions using Gaussian approximation in Algorithm 1.

Theorem 2 (Corollary 1 in [30])**.** *Consider the nonlinear system* (14) *with conditions C_1), C_2') and C_3). Then for each k, $\tau_{k-1} \le t \le \tau_k$, the solution $\tilde{u}_k(t,x)$ for* (23) *is reduced to the solution $\hat{u}_k(t,x)$ for the following Kolmogorov equation*

$$\begin{aligned}\frac{\partial \hat{u}_k}{\partial t}(t,x) &= \frac{1}{2}\Delta \hat{u}_k(t,x) - \sum_{i=1}^{n} l_i(x)\frac{\partial \hat{u}_k}{\partial x_i}(t,x) \\ &\quad + \frac{1}{2}\left(\sum_{i=1}^{n} l_i^2(x) - \sum_{i=1}^{n}\frac{\partial l_i}{\partial x_i}(x) - \eta(x)\right)\hat{u}_k(t,x)\end{aligned} \tag{25}$$

with

$$\hat{u}_k(\tau_{k-1},x) = \begin{cases} e^{-\phi(x)}\sigma_0(x), & k=1, \\ \exp\left(\sum_{j=1}^{m}(y_j(\tau_{k-1}) - y_j(\tau_{k-2}))h_j(x)\right)\cdot \hat{u}_{k-1}(\tau_{k-1},x), & k \ge 2 \end{cases} \tag{26}$$

where

$$\tilde{u}_k(t,x) = e^{\phi(x)}\hat{u}_k(t,x). \tag{27}$$

Suppose $\hat{u}_k(\tau_{k-1}, x)$ is well approximated by a sum of finite number of Gaussian distributions, it follows that a well approximated solution of (25) is obtained by linear combination of solutions of (25) with Gaussian initial condition since (25) is a linear PDE. The following theorem give the solution of (25) with Gaussian initial distribution in terms of ODEs.

Theorem 3 (Kolmogorov equation → ODEs, Theorem 3.2 in [39]). *Consider the following Kolmogorov equation with Gaussian initial condition*

$$\begin{cases} \dfrac{\partial \hat{u}_k}{\partial t}(t,x) = \dfrac{1}{2}\Delta \hat{u}_k - \displaystyle\sum_{i=1}^{n} l_i(x)\dfrac{\partial \hat{u}_k}{\partial x_i}(t,x) \\ \qquad\qquad\qquad + \dfrac{1}{2}\left(\displaystyle\sum_{i=1}^{n} l_i^2(x) - \sum_{i=1}^{n}\dfrac{\partial l_i}{\partial x_i}(x) - \eta(x)\right)\hat{u}_k(t,x), \\ \hat{u}_k(\tau_{k-1}, x) = e^{x^T A(\tau_{k-1})x + B^T(\tau_{k-1})x + C(\tau_{k-1})}, \end{cases} \tag{28}$$

where $A(\tau_{k-1}) = (A_{ij}(\tau_{k-1}))$ is a $n \times n$ symmetric matrix, $B^T(\tau_{k-1}) = (B_1(\tau_{k-1}), \cdots, B_n(\tau_{k-1}))$, $x^T = (x_1, \cdots x_n)$ are row vectors and $C(\tau_{k-1})$ is a scalar.

Let

$$\begin{aligned} q(x) &= \frac{1}{2}\left(\sum_{i=1}^{n} l_i^2(x) - \sum_{i=1}^{n}\frac{\partial l_i}{\partial x_i}(x) - \eta(x)\right) \\ &= x^T Q x + p^T x + r \end{aligned} \tag{29}$$

where $l_i(x) = \sum_{j=1}^{n} d_{ij}x_j + d_i, Q = (q_{ij})$ is a $n \times n$ symmetric matrix, $p^T = (p_1, \cdots, p_n)$ is a row vector and r is a scalar.

Then the solution of (28) is of the following form

$$\hat{u}_k(t,x) = e^{x^T A(t)x + B(t)^T x + C(t)} \tag{30}$$

where $A(t) = (A_{ij}(t))$ is a $n \times n$ symmetric matrix valued function of t, $B^T(t) = (B_1(t), \cdots, B_n(t)$ is a row vector valued function of t, and $C(t)$ is a scalar function of t. Moreover, $A(t), B(t)$ and $C(t)$ satisfy the following *system of nonlinear ODEs:*

$$
(31)\qquad
\begin{aligned}
\frac{dA(t)}{dt} &= 2A^2(t) - [A(t)D(t) + D^T A(t)] + Q(t),\\
\frac{dB^T(t)}{dt} &= 2B^T(t)A(t) - B^T(t)D(t) - 2d^T(t)A(t) + p^T(t),\\
\frac{dC(t)}{dt} &= trA(t) + \frac{1}{2}B^T B(t) - d^T(t)B(t) + r(t),
\end{aligned}
$$

where $D = (d_{ij})$ is a $n \times n$ matrix and $d^T = (d_1, \cdots, d_n)$ is a $1 \times n$ vector.

Given a probability density $\varphi(x)$ and the threshold E, [30] proposed a numerical algorithm to get a Gaussian approximation $\tilde{\varphi}(x) = \sum_{i=1}^{\widetilde{N}} \alpha_i \mathcal{N}(\mu_i, \sigma_i)$ which satisfies $\max_x |\varphi(x) - \tilde{\varphi}(x)| \le E$, and $\widetilde{N}, \alpha_i, \mu_i, \sigma_i$ are determined by probability density $\varphi(x)$ and the threshold E. This Gaussian approximation method is summarized in Algorithm 1.

Algorithm 1 Gaussian approximation

1: Let $a(x) = \varphi(x)$ and the threshold $E = \alpha * \max \varphi(x)$, where α is a given small parameter.
2: Fitting the peaks of $a(x)$ which are larger than E with gaussian distributions. Suppose the sum of gaussian distributions in this step is $g(x)$.
3: Let $a_1(x) = a(x) - g(x)$. If $a_1(x)$ has no peaks whose values larger than E, then go to step 4. Otherwise, let $a(x) = a_1(x)$ and go to step 2.
4: Let $a_2(x) = -a_1(x)$. If $a_2(x)$ has no peaks which are larger than E, then done. Otherwise, let $a(x) = a_2(x)$ and go to step 2.

Using the Gaussian approximation procedure in Algorithm 1, $\hat{u}_k(\tau_{k-1}, x)$ in (28) can be decomposed into a finite number of Gaussian distributions, and the Kolmogorov equation (25) with Gaussian initial condition is solved in terms of ODEs by Theorem 3. In summary, the algorithm to compute $\tilde{u}_k(t, x)$ is listed in Algorithm 2 [30].

4. Time-varying Yau filtering system

Now we discuss the general time-varying filtering system (1) and furthermore, we consider the time-varying Yau filtering system, i.e.:

$$
(32)\qquad f(x,t) = L(t)x + l(t) + \nabla_x \phi(t,x),
$$

where $L(t) = (l_{ij}(t)), 1 \le i, j \le n, l^T(t) = (l_1(t), \cdots l_n(t))$ and $\phi(t,x)$ is a C^∞ function w.r.t. x on $\mathbb{R}^n$.

Algorithm 2 Compute $\tilde{u}_k(t,x)$

1: Choose the total computing time T, Δt and the parameter α in Algorithm 1. Let $N = \frac{T}{\Delta t}$, and partition the time interval $[0,T]$ by $\{0 = \tau_0 < \tau_1 < \tau_2 < \ldots < \tau_N = T\}$.
2: **for** $k = 1 : N$ **do**
3: Using Algorithm 1, suppose $\hat{u}_k(\tau_{k-1},x)$ is decomposed into $\sum_{i=1}^{N(k)} c_{k,i}G(\mu_{k,i},\sigma_{k,i})$.
4: For each Gaussian distribution $G(\mu_{k,i},\sigma_{k,i})$, suppose the solution of (28) with initial condition $G(\mu_{k,i},\sigma_{k,i})$ is $\hat{u}_{k,i}(\tau_k,x)$. Solving (31), we obtain $\hat{u}_{k,i}(\tau_k,x)$. Then $\hat{u}_k(\tau_k,x) = \sum_{i=1}^{N(k)} c_{k,i}\hat{u}_{k,i}(\tau_k,x)$.
5: From (27), we have $\tilde{u}_k(\tau_k,x) = e^{-\phi(x)}\hat{u}_k(\tau_k,x)$.
6: By (26), we obtain $\hat{u}_{k+1}(\tau_k,x)$.
7: **end for**

Both the limited direct method [8] and general direct method [9] are devoted to solve the time-varying Yau filtering system. And the main difference is that the limited direct method relies on the strong assumption 2 w.r.t. the system while the general direct method only needs the basic assumptions w.r.t. the system. We shall introduce these two methods in the following two subsections and both of them follows Proposition 1, i.e., they seek to solve (13).

4.1. Limited direct method for time-varying Yau filtering system

Proposition 2 (Proposition 2 in [8])**.** *Suppose $\tilde{u}_k(t,x)$ is the solution to (13) in the interval $[\tau_{k-1},\tau_k]$, $k = 1,2,\cdots N$ and $f(x,t)$ is of the form (32). Let*

$$\tilde{u}_k(t,x) = e^{\phi(t,x)}\tilde{v}_k(t,x), \tag{33}$$

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

$$\begin{cases} \dfrac{\partial \tilde{v}_k}{\partial t}(t,x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n} G_{ij}(t)\dfrac{\partial^2 \tilde{v}_k}{\partial x_i \partial x_j}(t,x) - (Lx+l)^T\nabla\tilde{v}_k(t,x) \\ \qquad\qquad - \dfrac{1}{2}q(t,x)\tilde{v}_k(t,x), \\ \tilde{v}_1(0,x) = \sigma_0(x)e^{-\phi(0,x)}, \\ \tilde{v}_k(\tau_{k-1},x) = \exp\left[h^T(x,\tau_{k-1})S^{-1}(\tau_{k-1})(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] \\ \qquad\qquad \cdot\, \tilde{v}_{k-1}(\tau_{k-1},x), \;\; k = 2,3,\cdots N, \end{cases} \tag{34}$$

where

$$
\begin{aligned}
q(t,x) = &\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2\phi}{\partial x_i\partial x_j}(t,x) + \nabla_x\phi^T(t,x)G(t)\nabla_x\phi(t,x) \\
&+ 2(Lx+l)^T\nabla_x\phi(t,x) \\
&+ \sum_{p,l=1}^{n} S_{pl}^{-1}(t)h_p(x,t)h_l(x,t) + 2tr(L).
\end{aligned} \tag{35}
$$

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

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

$$
G(t) = F(t)F^T(t) \tag{36}
$$

according to Cholesky decomposition.

Assumption 2. *$L(t)$ in (32) can be expressed as follows:*

$$
L(t) = G(t)\Omega(t) + \frac{dF(t)}{dt}F^{-1}(t) \tag{37}
$$

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

Theorem 4 (Kolmogorov equation → Schrödinger equation, Theorem 1 in [8])**.** *Under Assumption 1-2, suppose $\tilde{v}_k(t,x)$ is a solution of (34) and let*

$$
\tilde{v}_k(t,x) = e^{x^T D(t)x}v_k(t,z), \tag{38}
$$

where

$$
\begin{aligned}
z &= B(t)x + b(t), \\
B(t) &= F^{-1}(t), \\
b(t) &= \int_0^t B(s)l(s)ds,
\end{aligned} \tag{39}
$$

and

$$
D(t) = \frac{1}{2}\Omega(t). \tag{40}
$$

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

(41)
$$\begin{cases} \dfrac{\partial v_k}{\partial t}(t,z) = \dfrac{1}{2}\Delta v_k(t,z) - \dfrac{1}{2}\tilde{q}\left(t, F(t)z - F(t)b(t)\right) v_k(t,z) \\ v_1(0,x) = \sigma_0(F(0)x)\exp\left[-\phi\left(0, F(0)x\right) - \left(F(0)x\right)^T D(0)\left(F(0)x\right)\right] \\ v_k(\tau_{k-1},x) = \exp\left[h^T(F(\tau_{k-1})x - F(\tau_{k-1})b(\tau_{k-1}), \tau_{k-1})S^{-1}(\tau_{k-1}) \right. \\ \qquad \left. \cdot\,(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] v_{k-1}(\tau_{k-1},x),\ k = 2,3,\cdots N \end{cases}$$

where

$$\begin{aligned} \tilde{q}(t,x) =& q(t,x) + 2x^T\frac{dD(t)}{dt}x - tr\left(G(t)(D(t)+D^T(t))\right) \\ & - x^T\left(D(t)+D^T(t)\right)G(t)\left(D(t)+D^T(t)\right)x \\ & + 2(L(t)x+l)^T\left(D(t)+D^T(t)\right)x. \end{aligned} \tag{42}$$

If $q(t,x)$ in (41) is quadratic in x, then it is called Schrödinger equation. Though it feels very restrictive, it includes Kalman-Bucy [18] and Beneš [5] filtering.

Assumption 3. *$\tilde{q}(t,x)$ defined in (35) is quadratic 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 (35).

Since $\tilde{q}(t,x)$ is quadratic in x by (42) under Assumption 3. Thus we can assume that

$$\tilde{q}(t,x) = x^T Q(t)x + p^T(t)x + r(t). \tag{43}$$

Theorem 5 (Theorem 3 in [8])**.** *Under Assumption 1-3, the solution $v_k(t,z)$ in $\tau_{k-1} \le t \le \tau_k$ of (41) is given by*

$$v_k(t,x) = \int_{\mathbb{R}^n} K(t,x,y)v_k(\tau_{k-1},y)dy, \tag{44}$$

where

$$K(t,x,y) = (2\pi(t-\tau_{k-1}))^{-n/2} \cdot \exp\Big\{ -\frac{|x-y|^2}{2(t-\tau_{k-1})} + x^T\tilde{A}(t-\tau_{k-1})x + x^T\tilde{B}(t-\tau_{k-1})y + y^T\tilde{C}(t-\tau_{k-1})y + \tilde{D}^T(t-\tau_{k-1})x + \tilde{E}^T(t-\tau_{k-1})y + s(t-\tau_{k-1})\Big\}, \tag{45}$$

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*

$$\tilde{A}_{\nu+1} = \frac{2}{\nu+3}\sum_{i=0}^{\nu}\tilde{A}_i\tilde{A}_{\nu-i} - \frac{1}{2(\nu+3)}\sum_{j=0}^{\nu}\sum_{i=0}^{j}F_i^TQ_{j-i}F_{\nu-j}, \tag{46}$$

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

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

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

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

$$
\begin{aligned}
s_{\nu+1} = & \frac{1}{2(\nu+1)} \sum_{i=-1}^{\nu+1} \tilde{D}_i^T \tilde{D}_{\nu-i} + \frac{1}{\nu+1} tr(\tilde{A}_\nu) \\
& - \frac{1}{2(\nu+1)} \left[\sum_{i=0}^{\nu} \sum_{j=0}^{i} \sum_{m=0}^{j} \sum_{l=0}^{m} b_l^T F_{m-l}^T Q_{j-m} F_{i-j} b_{\nu-i} \right. \\
& \left. - \sum_{j=0}^{\nu} \sum_{i=0}^{j} p_i^T F_{j-i} b_{\nu-j} + r_\nu \right],
\end{aligned}
$$

with

$$
\begin{aligned}
& \tilde{A}_{-1} = \tilde{C}_{-1} = -\frac{1}{2} I, \ \tilde{B}_{-1} = I, \\
& \tilde{D}_{-1} = \tilde{E}_{-1} = s_{-1} = 0, \\
& \tilde{A}_0 = \tilde{B}_0 = \tilde{C}_0 = \tilde{D}_0 = \tilde{E}_0 = s_0 = 0.
\end{aligned} \tag{47}
$$

To implement the proposed direct method numerically, we need to truncate the higher order, 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 (46) where M is the assumed order. The numerical procedure of direct method for nonlinear filtering problem (1) is listed in TABLE 3 [8].

Algorithm 3 Limited direct method for time-varying (1)

1: **Initialization**: give $T_0, T, \Delta, \sigma_0(x), M \geq 0$
2: Calculate $N = (T - T_0)/\Delta$
3: Calculate $F(t), B(t), b(t), D(t)$ by (36), (37), (39), (40)
4: Calculate $Q(t), p(t), r(t)$ by (35), (42), (43)
5: Calculate $\tilde{A}_\nu, \tilde{B}_\nu, \tilde{C}_\nu, \tilde{D}_\nu, \tilde{E}_\nu, s_\nu, 0 \leq \nu < M$ by (46)
6: **for** $k = 1 : N$ **do**
7:    Calculate $v_k(t_{k-1}, x), v_k(t_k, x)$ by (41), (44)
8:    Calculate $\tilde{v}_k(t_k, x), \tilde{u}_k(t_k, x)$ by (38), (33)
9:    Calculate $u_k(t_k, x), \sigma(t_k, x)$ by (12), (6)
10:    Calculate $\hat{x}_{t_k}$
11:    Assign $k := k + 1$
12: **end for**

4.2. General direct method for time-varying Yau filtering system

The aim of the general direct method is to solve (13) without Assumption 2.

Proposition 3 ([9]). *Under the Assumption 1, and let $\tilde{u}_k(t,x)$ be the solution of* (13) *in $[\tau_{k-1},\tau_k]$, $k=1,2,\cdots N$, $f(x,t)$ satisfies (32). Let*

$$\tilde{u}_k(t,x) = e^{\bar{\phi}(t,x)}\tilde{\psi}_k(t,x), \tag{48}$$

where $\bar{\phi}(t,x)$ satisfies $\nabla_x\bar{\phi}(t,x) = G^{-1}(t)\nabla_x\phi(t,x)$, then $\tilde{\psi}_k(t,x)$ satisfies the following equation:

$$\begin{cases} \dfrac{\partial\tilde{\psi}_k}{\partial t}(t,x) = \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{n} G_{ij}(t)\dfrac{\partial^2\tilde{\psi}_k}{\partial x_i\partial x_j}(t,x) \\ \qquad\qquad - (Lx+l)^T\nabla\tilde{\psi}_k(t,x) - \dfrac{1}{2}\bar{q}(t,x)\tilde{\psi}_k(t,x), \\ \tilde{\psi}_1(0,x) = \sigma_0(x)e^{-\bar{\phi}(0,x)}, \\ \tilde{\psi}_k(\tau_{k-1},x) = \exp\left[h^T(x,\tau_{k-1})S^{-1}(\tau_{k-1})(y_{\tau_{k-1}} - y_{\tau_{k-2}})\right] \\ \qquad\qquad \cdot\tilde{\psi}_{k-1}(\tau_{k-1},x), \ k=2,3,\cdots N, \end{cases} \tag{49}$$

where

$$\begin{aligned} \bar{q}(t,x) =& -\sum_{i,j=1}^{n} G_{ij}(t)\frac{\partial^2\bar{\phi}}{\partial x_i\partial x_j}(t,x) + \nabla_x\bar{\phi}^T(t,x)G(t)\nabla_x\bar{\phi}(t,x) \\ &+ 2(Lx+l)^T\nabla_x\bar{\phi}(t,x) + 2\sum_{i=1}^{n}\frac{\partial^2\phi(t,x)}{\partial^2 x_i^2} + 2\frac{\partial\bar{\phi}(t,x)}{\partial t} \\ &+ \sum_{p,l=1}^{n} S_{pl}^{-1}(t)h_p(x,t)h_l(x,t) + 2tr(L). \end{aligned} \tag{50}$$

Theorem 6 ([9]). *Under the Assumption 1, and $\tilde{\psi}_k(t,x)$ is the solution of (49), let*

$$\tilde{\psi}_k(t,x) = \psi_k(t,z), \tag{51}$$

where

$$\begin{aligned} z &= B(t)x, \\ B(t) &= F^{-1}(t). \end{aligned} \tag{52}$$

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

$$
\left\{
\begin{aligned}
\frac{\partial \psi_k}{\partial t}(t,z) &= \frac{1}{2}\Delta\psi_k(t,z) - \frac{1}{2}\bar{q}\,(t,F(t)z)\,\psi_k(t,z) \\
&\qquad - \left[\left(\frac{dB}{dt}B^{-1} + BLB^{-1}\right)z + Bl\right]^T \nabla\psi_k(t,z), \\
\psi_1(0,z) &= \sigma_0(F(0)z)\exp\left(-\bar{\phi}(0,F(0)z)\right), \\
\psi_k(\tau_{k-1},z) &= \exp\big[h^T(F(\tau_{k-1})z,\tau_{k-1})S^{-1}(\tau_{k-1}) \\
&\qquad \cdot (y_{\tau_{k-1}} - y_{\tau_{k-2}})\big]\psi_{k-1}(\tau_{k-1},z), \\
&\qquad k = 2,3,\cdots N.
\end{aligned}
\right. \tag{53}
$$

Define

$$
\tilde{\bar{q}}(t,z) := \bar{q}(t,F(t)z), \tag{54}
$$

and rewrite (53) as

$$
\left\{
\begin{aligned}
\frac{\partial \psi_k}{\partial t}(t,x) &= \frac{1}{2}\Delta\psi_k(t,x) - \frac{1}{2}\tilde{\bar{q}}(t,x)\psi_k(t,x) \\
&\qquad - \left[\left(\frac{dB}{dt}B^{-1} + BLB^{-1}\right)x + Bl\right]^T \nabla\psi_k(t,x) \\
\psi_1(0,x) &= \sigma_0(F(0)x)\exp\left(-\bar{\phi}(0,F(0)x)\right) \\
\psi_k(\tau_{k-1},x) &= \exp\big[h^T(F(\tau_{k-1})x,\tau_{k-1})S^{-1}(\tau_{k-1}) \\
&\qquad \cdot (y_{\tau_{k-1}} - y_{\tau_{k-2}})\big]\psi_{k-1}(\tau_{k-1},x), \\
&\qquad k = 2,3,\cdots N.
\end{aligned}
\right. \tag{55}
$$

Assumption 4. *$\tilde{\bar{q}}(t,x)$ in (54) is quadratic w.r.t. x.*

It follows naturally $\tilde{\bar{q}}(t,x)$ can be rewritten as

$$
-\frac{1}{2}\tilde{\bar{q}}(t,x) = x^TQ(t)x + p^T(t)x + r(t), \tag{56}
$$

where $Q(t)$ is a $n\times n$ symmetric matrix, $p(t)$ is a $n\times 1$ vector and $r(t)$ is a scalar.

Theorem 7. *Under Assumption 1 and Assumption 4, we consider the fol-*

lowing equation:

$$
\left\{
\begin{aligned}
\frac{\partial \psi_k}{\partial t}(t,x) &= \frac{1}{2}\Delta\psi_k(t,x) - \frac{1}{2}\tilde{\tilde{q}}(t,x)\psi_k(t,x) \\
&\qquad - \left[\left(\frac{dB}{dt}B^{-1} + BLB^{-1}\right)x + Bl\right]^T \nabla\psi_k(t,x) \\
\psi_k(\tau_{k-1},x) &= \exp\left\{x^T A(\tau_{k-1})x + b^T(\tau_{k-1})x + c(\tau_{k-1})\right\},
\end{aligned}
\right. \tag{57}
$$

where $A(\tau_{k-1})$ is a $n \times n$ symmetric matrix, $b(\tau_{k-1})$ is a $n \times 1$ column vector, $x^T = (x_1, x_2, \ldots, x_n)$ is a row vector and $c(\tau_{k-1})$ is a scalar. Then the solution of (57) is of the following form:

$$
\psi_k(t,x) = \exp\left\{x^T A(t)x + b^T(t)x + c(t)\right\}, \tag{58}
$$

where $A(t)$ is a $n \times n$ matrix function w.r.t. t which is symmetric, $b(t)$ is a $n \times 1$ column vector function w.r.t. t and $c(t)$ is a scalar function w.r.t. t, and they satisfy the following ODEs:

$$
\begin{aligned}
\frac{dA(t)}{dt} &= 2A^2(t) - 2A(t)D(t) + Q(t), \\
\frac{db^T(t)}{dt} &= 2b^T(t)A(t) - b^T(t)D(t) - 2d^T(t)A(t) + p^T(t), \\
\frac{dc(t)}{dt} &= trA(t) + \frac{1}{2}b^T b(t) - d^T(t)b(t) + r(t),
\end{aligned} \tag{59}
$$

where

$$
D(t) = \frac{dB}{dt}B^{-1} + BLB^{-1}, d(t) = B(t)l(t). \tag{60}
$$

Comparing Theorem 5 and Theorem 7, it is known that we eliminate Assumption 2 and only keep that basic assumptions, i.e., Assumption 1 and Assumption 3 (or 4).

Theorem 7 requires that the initial value $\psi_k(\tau_{k-1}, x)$ at every τ_{k-1} must be gaussian. According to the gaussian approximation algorithm in Algorithm 1, non-gaussian function can be approximated by the sum of several gaussian functions and we can use Theorem 7 for every gaussian function. The general direct method is summarized in Algorithm 4.

5. Conclusion

In this survey, we first give the general framework of the global method and then introduce four kinds of direct methods. The key ingredient for direct

Algorithm 4 General direct method for time-varying (1)

1: **Initialization**: give $T, \Delta t, \sigma_0(x)$ and the parameter α in Algorithm 1. Let $N = \frac{T}{\Delta t}$, and $\{0 = \tau_0 < \tau_1 < \tau_2 < ... < \tau_N = T\}$.
2: **for** $k = 1 : N$ **do**
3: Using Algorithm 1 to get the Gaussian approximation $\psi_k(\tau_{k-1}, x) \approx \sum_{i=1}^{N(k)} \alpha_{k,i}\mathcal{N}(\mu_{k,i}, \sigma_{k,i})$.
4: For each Gaussian distribution $\mathcal{N}(\mu_{k,i}, \sigma_{k,i})$, suppose the solution of (57) with initial condition $\mathcal{N}(\mu_{k,i}, \sigma_{k,i})$ is $\psi_{k,i}(\tau_k, x)$. Solving (59), we obtain $\psi_{k,i}(\tau_k, x)$. Then $\psi_k(\tau_k, x) = \sum_{i=1}^{N(k)} \alpha_{k,i}\hat{u}_{k,i}(\tau_k, x)$.
5: Calculate $\psi_{k+1}(\tau_k, x)$ by $\psi_k(\tau_k, x)$ and (55).
6: Calculate $\tilde{\psi}_k(t_k, x), \tilde{u}_k(t_k, x)$ by (38), (33).
7: Calculate $u_k(t_k, x), \sigma(t_k, x)$ by (12), (6).
8: Calculate $\rho(t_k, x)$ by (4).
9: Calculate the conditional expectation of the state x_{t_k}.
10: **end for**

method is how to solve the DMZ equation and these direct methods solve the DMZ equation by different transformations and approximations. In the work of general direct method, we extend the classic direct method to the most general time-varying systems.

Acknowledgements

We are grateful to the editor and the referees for their helpful and detailed comments on our manuscript.

References

[1] J. L. Anderson and S. L. Anderson, *A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts*. Monthly Weather Review, **127**(12):2741–2758, 1999.

[2] J. L. Anderson, *An ensemble adjustment Kalman filter for data assimilation*. Monthly Weather Review, **129**(12):2884–2903, 2001.

[3] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, *A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking*. IEEE Transactions on Signal Processing, **50**(2):174–188, 2002.

[4] A. Bain and D. Crisan, *fundamentals of stochastic filtering*. Stochastic Modelling and Applied Probability, Springer, Berlin, volume 60, 2009. MR2454694

[5] V. Beneš, *Exact finite dimensional filters for certain diffusions with nonlinear drift*. Stochastics, **5**:65–92, 1981. MR0643062

[6] R. W. Brockett and J. M. C. Clark, *The geometry of the conditional density equation*. Analysis & Optimisation of Stochastic Systems, pages 299–309, 1980. MR0592992

[7] J. Chen, *On ubiquity of Yau filters*. In Proceedings of the American Control Conference, Baltimore, MD, pages 252–254, June 1994.

[8] X. Q. Chen, X. Luo and S. S.-T. Yau, *Direct method for time-varying nonlinear filtering problems*. IEEE Transactions on Aerospace and Electronic Systems, **53**(2):630–639, 2017.

[9] X. Q. Chen, J. Shi and S. S.-T. Yau, *Real-time solution of time-varying yau filtering problems via direct method and gaussian approximation*. IEEE Transactions on Automatic Control, **64**(4):1648–1654, Apr. 2019. MR3936439

[10] W. L. Chiou and S. S.-T. Yau, *Finite-dimensional filters with nonlinear drift. II. Brockett's problem on classification of finite-dimensional estimation algebras*. SIAM Journal on Control and Optimization, **32**(1):297–310, 1994. MR1255972

[11] A. Doucet, N. D. Freitas and N. Gordon, *An introduction to sequential Monte Carlo methods*. Sequential Monte Carlo Methods in Practice, Springer, New York, pages 3–14, 2001. MR1847784

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

[13] G. Evensen, *The Ensemble Kalman Filter: theoretical formulation and practical implementation*. Ocean Dynamics, **53**(4):343–367, 2003.

[14] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, *Novel approach to nonlinear/non-gaussian bayesian state estimation*. Radar & Signal Processing IEE Proceedings F, **140**(2):107–113, 1993.

[15] G.-Q. Hu and S. S.-T. Yau, *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):50–57, 2002. MR2696815

[16] A. H. Jazwinski, *Stochastic processes and filtering theory*. Academic Press, 1970.

[17] R. E. Kalman, *A new approach to linear filtering and prediction problems*. Journal of Basic Engineering, **82**:35–45, 1960.

[18] R. E. Kalman and R. S. Bucy, *New results in linear filtering and prediction theory*. Transactions of the Asme-Journal of Basic Engineering, **83**:95–108, 1961. MR0234760

[19] S. Lakshmivarahan and D. J. Stensrud, *Ensemble kalman filter application to meteorological data assimilation*. IEEE Control Systems Magazine, **29**(3):34–46, 2009. MR2518191

[20] X. Luo and S. S.-T. Yau, *Hermite spectral method to 1-D forward Kolmogorov equation and its application to nonlinear filtering problems*. IEEE Transactions on Automatic Control, **58**(10):2495–2507, 2013. MR3106057

[21] X. Luo and S. S.-T. Yau, *Complete real time solution of the general nonlinear filtering problem with out memory*. IEEE Transactions on Automatic Control, **58**(10):2563–2578, 2013. MR3106062

[22] X. Luo, *On recent advance of nonlinear filtering theory: emphases on global approaches*. Pure and Applied Mathematics Quarterly, **10**(4):685–721, 2014. MR3324765

[23] X. Luo, Y. Jiao, Yang W.-L. Chiou and S. S.-T. Yau, *A novel suboptimal method for solving polynomial filtering problems*. Automatica, **62**:26–31, 2015. MR3423968

[24] X. Luo, S. Yau and S. S.-T. Yau, *Time-dependent Hermite-Galerkin spectral method and its applications*. Applied Mathematics and Computation, **264**(C):378–391, 2015. MR3351619

[25] X. Luo, Y. Jiao, Yang W.-L. Chiou and S. S.-T. Yau, *Novel suboptimal filter via higher order central moments*. IEEE Transactions on Aerospace and Electronic Systems, **52**(4):2030–2038, 2016.

[26] S. K. Mitter, *On the analogy between mathematical problems of nonlinear filtering and quantum physics*. Ricerche di Automatica, **10**(2):163–216, 1979. MR0614260

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

[28] B. L. Rozovsky, *Stochastic partial differential equations arising in nonlinear filtering problems*. Uspekhi Mat. Nauk, **27**:213–214, 1972. MR0397864

[29] S. Sarkka, *On unscented Kalman filtering for state estimation of continuous-time nonlinear systems*. IEEE Transactions on Automatic Control, **52**(9):1631–1641, 2007. MR2352439

[30] J. Shi, Z. Y. Yang and S. S.-T. Yau, *Direct method for Yau filter system with nonlinear observations*. International Journal of Control, **91**(3):678–687, 2018. MR3760421

[31] J. Shi and S. S.-T. Yau, *Finite dimensional estimation algebras with state dimension 3 and rank 2, I: Linear structure of Wong matrix*. SIAM Journal on Control and Optimization, **55**(6):4227–4246, 2017. MR3738843

[32] J. Shi, X. Q. Chen, W. H. Dong and S. S.-T. Yau, *New classes of finite dimensional filters with non-maximal rank*. IEEE Control Systems Letters, **1**(2):233–237, 2017.

[33] E. A. Wan, and R. Van Der Merwe, *The unscented Kalman filter for nonlinear estimation*. Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000, AS-SPCC, pages 153–158, 2000.

[34] S. T. Yau, *Recent results on nonlinear filtering: new class of finite dimensional filters*. In Proceedings of the Decision and Control, Honolulu, HI, pages 231–233, Dec. 1990.

[35] S. S.-T. Yau, and S.-T. Yau, *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, pages 1221–1225, Dec. 14–16, 1994.

[36] S. S.-T. Yau, *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**:181–203, 1994. MR1298555

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

[38] S. T. Yau and G.-Q. Hu, *Finite-dimensional filters with nonlinear drift x: explicit solution of DMZ equation*. IEEE Transactions on Automatic Control, **46**(1):142–148, 2001. MR1809477

[39] S. S.-T. Yau and Y. T. Lai, *Explicit solution of DMZ equation in nonlinear filtering via solution of ODE*. IEEE Transactions on Automatic Control, **48**:505–508, 2003. MR1962262

[40] S. T. Yau and S. S.-T. Yau, *Nonlinear filtering and time varying Schrödinger equation*. IEEE Transactions on Aerospace and Electronic Systems, **40**(1):284–292, 2004.

[41] S. S.-T. Yau and G.-Q. Hu, *Classification of finite-dimensional estimation algebras of maximal rank with arbitrary state-space dimension and mitter conjecture*. International Journal of Control, **78**:689–705, 2005. MR2152444

[42] S. S.-T. Yau, *Solution of filtering problem with nonlinear observations*. Math. Nachr., **10**(3–4):187–196, 2006.

[43] S. S.-T. Yau and S.-T. Yau, *Real time solution of nonlinear filtering problem without memory II*. SIAM Journal on Control and Optimization, **47**(1):163–195, 2008. MR2373467

[44] M. Zakai, *On the optimal filtering of diffusion processes*. Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete, **11**:230–243, 1969.

Xiuqiong Chen
Yau Mathematical Sciences Center
Tsinghua University
Beijing, 100084
China
E-mail address: chenxiuqiong0828@163.com

Wenhui Dong
Department of Mathematical Sciences
Tsinghua University
Beijing, 100084
China
E-mail address: dwh15@mails.tsinghua.edu.cn

Received March 15, 2019