Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Filtering problem with nonlinear observations and drift terms equal to gradient vector field plus affine vector field

Stephen S.-T. Yau*† , IEEE Fellow
Department of Mathematics, Statistics
and Computer Science (M/C 249)
University of Illinois at Chicago
851 South Morgan Street
Chicago, IL 60607-7045, U.S.A.
yau@uic.edu

Lixing Jia
Department of Mathematics
and Computer Science
Chicago State University
Chicago, IL 60628
ljia@csu.edu

Joint Conference on Decision and Control, and European Control Conference
December 12-15, 2005, Seville, Spain

***Abstract*— For all known finite dimensional filters, one always needs the condition that the observation terms are degree one polynomial. On the other hand, in many practical examples, e.g. tracking problem, the observation terms may be nonlinear. In this paper, we solve filtering problems with nonlinear observation terms and drift term equal to gradient vector field plus affine vector field, which includes Kalman-Bucy filter, Benes filter and Yau filter as special cases.**

1. Introduction

In 1961, Kalman-Bucy first established the finite dimensional filters for linear filtering system with Gaussian initial distribution. In the sixties and early seventies, the basic approach to nonlinear filtering theory was via the "innovations method" originally proposed by Kailath and subsequently rigorously developed by Fujisaki, Kallianpur, and Kunita in 1972. As pointed out by Mitter, the difficulty with this approach is that the innovations process is not, in general, explicit computable. In view of this weakness, Brockett and Mitter proposed independently the idea of using estimation algebras to construct finite-dimensional nonlinear filters. The idea is to imitate the Wei-Norman approach of using the Lie algebraic method to solve the DMZ equation, which the unnormalized conditional probability of the state must satisfy. Perhaps the most important merit of the Lie algebra approach is the following. As long as the estimation algebra is finite dimensional, not only the finite dimensional filter can be constructed explicitly, but also the filter so constructed is universal in the sense of Chaleyat-Maurel and Michel [3]. In [21], [14], and [18] Yau proves that the number of sufficient statistics in the Lie algebra method, which is required in the computation of conditional probability density, is linear in n, where n is the dimension of the state space. Recently Stephen Yau [14] and his co-workers [11], [13], [4], [19], [18], and [5] have completely classified all finite dimensional estimation algebras of maximal rank. In particular, they have proved that all the observation terms $h_i(x)$, $1 \leq i \leq m$ must be degree one polynomials.

Recently, a new direct method was introduced to study the linear filtering and exact filtering systems with arbitrary initial condition for which f, g and h in (2.1) below are independent of time (cf. [20], [21], [17], [16]). This approach offers several advantages. It is easy and the derivation no longer needs controllability and observability. Thus, the algorithm is universal for any linear filtering system. Furthermore, it eliminates the necessity of integrating n first-order linear partial differential equations, as was the case in the Lie algebra method. Finally, the number of sufficient statistics required to compute the conditional probability density of the state in this direct method is n. In all these direct methods in [20], [21], [16], and [17] they need to assume that all the observation terms $h_i(x)$, $1 \leq i \leq m$, are degree one polynomials.

In our previous Asian J. Math. paper [24], we have proved the existence and decay estimates of the solution to the DMZ equation under the assumption that $f(x)$ and $h(x)$ in (2.1) below have linear growth. In this paper, we use the theory developed in [24] to show that the real time computation of DMZ equation can be reduced to numerical solution of Kolmogorov equation if $f(x)$ and $h(x)$ have linear growth. Similar results under a much stronger assumption that $f(x)$ and $h(x)$ are bounded functions were treated by various authors including Bensoussan, Glowinski and Rascanu [1], Elliott and Glowinski [7], Florchinger and Legland [8], Mikulevicious and Rozovskii [9]. Unlike our results, however their results cannot cover Kalman-Bucy filters. Theorem 4.2 of this paper says that if the drifts $(f(x))$ is an affine vector field plus gradient vector field and the observation terms $(h(x))$ are nonlinear with linear growths, then the Kolmogorov equation can be solved in real time.

For all known finite dimensional filters, one always needs the condition that the observation terms are degree one polynomial. On the other hand, in many practical examples, e.g. tracking problem, the observation terms may be nonlinear. Our new method in this paper can treat filtering problems with nonlinear observation terms in the first time, which

*Research partially supported by U.S. Army Research Office.
†Ze-Jiang Professor of East China Normal University.

includes Kalman-Bucy filter, Benes filter and Yau filter as a special case.

This paper is organized as follows. In section 2 we shall set up the notations and recall the basic filtering problem. In section 3, we shall show that real time computation of DMZ equation can be reduced to off time computation of Kolmogorov equation. An explicit algorithm of such a reduction is provided. In section 4, we show that if the drift is an affine vector field plus gradient vector field and the observation terms are nonlinear with linear growths, then the Kolmogorov equation can be solved in real time via a system of ODEs. Consequently, the nonlinear filtering problem with nonlinear observations with linear growth and drift term equal to gradient vector field plus affine vector field can be solved in real time and memoryless manner. In section 5, we give a conclusion of this paper.

2. Basic Filtering Problem

The filtering problem considered here is based on the following signal observation model in Itô form:

$$\begin{cases} dx(t) = f(x(t))dt + g(x(t))dv(t) & x(0) = x_0 \\ dy(t) = h(x(t))dt + dw(t) & y(0) = 0 \end{cases} \tag{2.1}$$

in which x, v, y and w are, respectively, $\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^m$, and $\mathbb{R}^m$ valued processes and v and w independent, standard Brownian processes. We further assume that $n = p$; f, g, and h are vector-valued, orthogonal matrix-valued and vector-valued C^∞ smooth functions. We shall refer to $x(t)$ as the state and $y(t)$ as the observation at time t.

Let $\rho(t, x)$ denote the conditional probability density of the state given the observation $\{y(s) : 0 \le s \le t\}$. It is well known that $\rho(t, x)$ is given by normalizing a function $\sigma(t, x)$ that satisfies the following DMZ equation in Fisk-Stratonovich form:

$$\begin{cases} d\sigma(t,x) & = L_0\sigma(t,x)dt + \sum_{i=1}^m L_i\sigma(t,x)dy_i(t) \\ \sigma(0,x) & = \sigma_0(x) \end{cases} \tag{2.2}$$

where

$$L_0 = \frac{1}{2}\sum_{i=1}^n \frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^n f_i(x)\frac{\partial}{\partial x_i} - \sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) - \frac{1}{2}\sum_{i=1}^m h_i^2(x),$$

and for $i = 1, \dots, m$, L_i is the zero-degree differential operator of multiplication by h_i and σ_0 is the probability density of the initial point x_0.

Davis introduced a new unnormalized density

$$u(t,x) = \exp\left(-\sum_{i=1}^m h_i(x)y_i(t)\right)\sigma(t,x). \tag{2.3}$$

He reduced (2.2) to the following time-varying partial differential equation which is called robust DMZ-equation

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) & = L_0 u(t,x) + \sum_{i=1}^m y_i(t)[L_0, L_i]u(t,x) \\ & \quad + \frac{1}{2}\sum_{i,j=1}^m y_i(t)y_j(t)[[L_0,L_i],L_j]u(t,x) \\ u(0,x) & = \sigma_0(x) \end{cases} \tag{2.4}$$

where $[\cdot\ ,\ \cdot]$ is the Lie bracket as described in [11]. It is easy to show [22] that (2.4) is equivalent to the following time-varying partial differential equation:

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\sum_{i=1}^n \frac{\partial^2 u}{\partial x_i^2}(t,x) + \sum_{i=1}^n \big(-f_i(x) \\ + \sum_{j=1}^m y_j(t)\frac{\partial h_j}{\partial x_i}(x)\big)\frac{\partial u}{\partial x_i}(t,x) \\ -\big[\sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^m h_i^2(x) \\ -\frac{1}{2}\sum_{i=1}^m y_i(t)\Delta h_i(x) \\ + \sum_{i=1}^m\sum_{j=1}^n y_i(t)f_j(x)\frac{\partial h_i}{\partial x_j}(x) \\ -\frac{1}{2}\sum_{i,j=1}^m\sum_{k=1}^n y_i(t)y_j(t)\frac{\partial h_i}{\partial x_k}(x)\frac{\partial h_j}{\partial x_k}(x)\big]u(t,x) \\ u(0,x) = \sigma_0(x) \end{cases} \tag{2.5}$$

In this paper we shall solve the filtering problem in the case $f_i(x)$, $1 \le i \le n$, are degree one polynomials and $h_j(x)$, $1 \le j \le m$, may be nonlinear with linear growth, i.e. $|h_j(x)| \le C(1 + |x|)$ for some constant C.

3. Reduction from robust DMZ equation to Kolmogorov equation

The fundamental problem of nonlinear filtering theory is how to solve the robust DMZ equation (2.5) in real time and in memoryless manner. In this section, we shall describe our algorithm which achieves this goal for a large class of filtering system with arbitrary initial distribution by reducing it to solve Kolmogorov equation. Our algorithm is based on the following observation.

Theorem 3.1: For any τ_1, τ_2 with $\tau_1 < \tau_2$, $\widetilde{u}(t,x)$ satisfies the following Kolmogorov equation

$$\begin{aligned} \frac{\partial \widetilde{u}}{\partial t}(t,x) &= \frac{1}{2}\Delta\widetilde{u}(t,x) - \sum_{i=1}^n f_i(x)\frac{\partial \widetilde{u}}{\partial x_i}(t,x) \\ &\quad -\Big(\sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^m h_i^2(x)\Big)\widetilde{u}(t,x) \end{aligned} \tag{3.1}$$

for $\tau_1 \le t \le \tau_2$ if and only if

$$u(t,x) = e^{-\sum_{i=1}^m y_i(\tau_1)h_i(x)}\widetilde{u}(t,x)$$

satisfies the robust DMZ equation with observation being freezed at $y(\tau_1)$

$$\begin{aligned} &\frac{\partial u}{\partial t}(t,x) = \frac{1}{2}\Delta u(t,x) \\ &+ \sum_{i=1}^n \big(-f_i(x) + \sum_{j=1}^m y_j(\tau_1)\frac{\partial h_j}{\partial x_i}(x)\big)\frac{\partial u}{\partial x_i}(t,x) \\ &-\big(\sum_{i=1}^n \frac{\partial f_i}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^m h_i^2(x) \\ &-\frac{1}{2}\sum_{i=1}^m y_i(\tau_1)\Delta h_i(x) \\ &+ \sum_{i=1}^m\sum_{j=1}^n y_i(\tau_1)f_j(x)\frac{\partial h_i}{\partial x_j}(x) \\ &-\frac{1}{2}\sum_{k=1}^n\sum_{i,j=1}^m y_i(\tau_1)y_j(\tau_1)\frac{\partial h_i}{\partial x_k} \\ &\qquad (x)\frac{\partial h_j}{\partial x_k}(x)\big)u(t,x) \end{aligned} \tag{3.2}$$

Proof: It is straight forward to show that

$$
\begin{aligned}
&e^{\sum_{i=1}^{m} y_i(\tau_1)h_i(x)}\big[-\tfrac{\partial}{\partial t}+\tfrac{1}{2}\Delta+\sum_{i=1}^{n}(-f_i(x)\\
&+\sum_{j=1}^{m} y_j(\tau_1)\tfrac{\partial h_j}{\partial x_i}\big)\tfrac{\partial}{\partial x_i}\\
&-\big(\sum_{i=1}^{n}\tfrac{\partial f_i}{\partial x_i}(x)+\tfrac{1}{2}\sum_{i=1}^{m} h_i^2(x)\\
&-\tfrac{1}{2}\sum_{i=1}^{m} y_i(\tau_1)\Delta h_i(x)\\
&+\sum_{i=1}^{m}\sum_{j=1}^{n} y_i(\tau_1)f_j(x)\tfrac{\partial h_i}{\partial x_j}(x)\\
&-\tfrac{1}{2}\sum_{k=1}^{n}\sum_{i,j=1}^{m} y_i(\tau_1)y_j(\tau_1)\tfrac{\partial h_i}{\partial x_k}\tfrac{\partial h_j}{\partial x_k}\big)\big]u(t,x)\\
&=-\tfrac{\partial \widetilde{u}}{\partial t}(t,x)+\tfrac{1}{2}\Delta\widetilde{u}(t,x)\\
&-\sum_{i=1}^{n} f_i(x)\tfrac{\partial \widetilde{u}}{\partial x_i}(t,x)\\
&-\big(\sum_{i=1}^{n}\tfrac{\partial f_i}{\partial x_i}(x)+\tfrac{1}{2}\sum_{i=1}^{m} h_i^2(x)\big)\widetilde{u}(t,x)
\end{aligned}
\tag{3.3}
$$

Proposition (3.1) follows immediately from (3.3). ■

We remark that (3.2) is obtained from robust DMZ equation by freezing the observation $y(t)$ to $y(\tau_1)$. Basing on Proposition (3.1), we shall formulate our algorithm to solve robust DMZ equation and we shall show in the appendix that the solution of our algorithm approximate the solution of robust DMZ equation very well in both pointwise sense and L^2-sense.

Suppose that $u(t,x)$ is the solution of robust DMZ equation and we want to compute $u(\tau,x)$. Let $\mathcal{P}_k=\{0=\tau_0<\tau_1<\tau_2<\cdots<\tau_k=\tau\}$ be a partition of $[0,\tau]$. Let $u_i(t,x)$ be a solution of the following partial differential equation for $\tau_{i-1}\le t\le\tau_i$.

$$
\left\{
\begin{aligned}
&\tfrac{\partial u_i}{\partial t}(t,x)=\tfrac{1}{2}\Delta u_i(t,x)+\sum_{\ell=1}^{n}\big(-f_\ell(x)\\
&+\sum_{j=1}^{m} y_j(\tau_{i-1})\tfrac{\partial h_j}{\partial x_\ell}(x)\big)\tfrac{\partial u_i}{\partial x_\ell}(t,x)\\
&-\big(\sum_{\ell=1}^{n}\tfrac{\partial f_\ell}{\partial x_\ell}(x)+\tfrac{1}{2}\sum_{\ell=1}^{m} h_\ell^2(x)\\
&-\tfrac{1}{2}\sum_{j=1}^{m} y_j(\tau_{i-1})\Delta h_j(x)\\
&+\sum_{j=1}^{m}\sum_{\ell=1}^{n} y_j(\tau_{i-1})f_\ell(x)\tfrac{\partial h_j}{\partial x_\ell}(x)\\
&-\tfrac{1}{2}\sum_{p=1}^{n}\sum_{j,\ell=1}^{m} y_j(\tau_{i-1})y_\ell(\tau_{i-1})\\
&\tfrac{\partial h_j}{\partial x_p}(x)\tfrac{\partial h_\ell}{\partial x_p}(x)\big)u_i(t,x)\\
&u_i(\tau_{i-1},x)=u_{i-1}(\tau_{i-1},x).
\end{aligned}
\right.
\tag{3.4}
$$

Define the norm of the partition $\mathcal{P}_k$ by $|\mathcal{P}_k|=\sup_{1\le i\le k}\{|\tau_i-\tau_{i-1}|\}$. In the appendix, we shall show that in both pointwise sense and L^2 sense

$$
u(\tau,x)=\lim_{|\mathcal{P}_k|\to 0} u_k(\tau,x). \tag{3.5}
$$

Therefore it remains to describe an algorithm to compute $u_k(\tau_k,x)$. By Proposition 3.1, $u_1(\tau_1,x)$ can be computed by $\widetilde{u}_1(\tau_1,x)$ where $\widetilde{u}_1(t,x)$ for $0\le t\le\tau_1$ satisfies the following Kolmogorov equation

$$
\left\{
\begin{aligned}
&\tfrac{\partial \widetilde{u}_1}{\partial t}(t,x)=\tfrac{1}{2}\Delta\widetilde{u}_1(t,x)-\sum_{j=1}^{n} f_j(x)\tfrac{\partial \widetilde{u}_1}{\partial x_j}(x)\\
&\quad-\big(\sum_{j=1}^{n}\tfrac{\partial f_j}{\partial x_j}(x)+\tfrac{1}{2}\sum_{i=1}^{m} h_j^2(x)\big)\widetilde{u}_1(t,x)\\
&\quad\widetilde{u}_1(0,x)=\sigma_0(x)e^{\sum_{j=0}^{m} y_j(0)h_j(x)}=\sigma_0(x).
\end{aligned}
\right.
\tag{3.6}
$$

In fact, by the uniqueness solution of Kolmogorov equation, we have

$$
u_1(t,x)=\widetilde{u}_1(t,x),\qquad 0\le t\le\tau_1. \tag{3.7}
$$

In general, Proposition 3.1 tells us that for $i\ge 2$, $u_i(\tau_i,x)$ can be computed by $\widetilde{u}_i(\tau_i,x)$ where $\widetilde{u}_i(t,x)$ for $\tau_{i-1}\le t\le\tau_i$ satisfies the following Kolmogorov equation

$$
\left\{
\begin{aligned}
&\tfrac{\partial \widetilde{u}_i}{\partial t}(t,x)=\tfrac{1}{2}\Delta\widetilde{u}_i(t,x)-\sum_{j=1}^{n} f_j(x)\tfrac{\partial \widetilde{u}_i}{\partial x_j}(t,x)\\
&\quad-\big(\sum_{j=1}^{n}\tfrac{\partial f_j}{\partial x_j}(x)+\tfrac{1}{2}\sum_{j=1}^{m} h_j^2(x)\big)\widetilde{u}_i(t,x)\\
&\widetilde{u}_i(\tau_{i-1},x)=e^{\sum_{j=1}^{m}(y_j(\tau_{i-1})-y_j(\tau_{i-2}))h_j(x)}\\
&\qquad\widetilde{u}_{i-1}(\tau_{i-1},x)
\end{aligned}
\right.
\tag{3.8}
$$

where the last initial condition comes from

$$
\begin{aligned}
\widetilde{u}_i(\tau_{i-1},x) &= u_i(\tau_{i-1},x)e^{\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)}\\
&= u_{i-1}(\tau_{i-1},x)e^{\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)}\\
&= e^{\sum_{j=1}^{m}(y_j(\tau_{i-1})-y_j(\tau_{i-2}))h_j(x)}\widetilde{u}_{i-1}(\tau_{i-1},x).
\end{aligned}
$$

In fact, we have

$$
u_i(\tau_i,x)=e^{-\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)}\widetilde{u}_i(\tau_i,x). \tag{3.9}
$$

In view of (2.3), (3.5) and (3.9), we have the following theorem.

Theorem 3.2: The unnormalized density σ can be computed via solution $\widetilde{u}$ of Kolmogorov equation (3.8). More specifically,

$$
\sigma(\tau,x)=\lim_{|\mathcal{P}_k|\to 0}\widetilde{u}_k(\tau_k,x) \tag{3.10}
$$

Proof:

$$
\begin{aligned}
\sigma(\tau,x) &= u(\tau,x)\exp\left(\sum_{i=1}^{m} h_i(x)y_i(\tau)\right) && \text{by (2.3)}\\
&= \lim_{|\mathcal{P}_k|\to 0} u_k(\tau,x)\exp\left(\sum_{i=1}^{m} h_i(x)y_i(\tau)\right), && \text{by (3.5)}
\end{aligned}
$$

where $\mathcal{P}_k=\{0=\tau_0<\tau_1<\cdots<\tau_k=\tau\}$.

In view of (3.9), we have

$$
\begin{aligned}
\sigma(\tau,x) &= \lim_{|\mathcal{P}_k|\to 0} e^{-\sum_{i=1}^{m} y_j(\tau_{k-1})h_j(x)}\\
&\qquad \widetilde{u}_k(\tau,x)e^{\sum_{i=1}^{m} h_i(x)y_i(\tau)}\\
&= \lim_{|\mathcal{P}_k|\to 0}\widetilde{u}_k(\tau,x)
\end{aligned}
$$

■

Observe that in our algorithm at step i, we only need the observation at time τ_{i-1} and τ_{i-2}. We do not need any other previous observation data. Observe also that the Kolmogorov equation (3.8) is uniform for all time steps and it depends on observation $y(t)$ only via initial condition.

4. Filtering problem with nonlinear observations

Consider the filtering system (2.1) with drift equal to affine vector field plus gradient vector field

$$f_i(x) = \sum_{j=1}^{n} \ell_{ij} x_j + \ell_i + \frac{\partial \phi}{\partial x_i}, \qquad 1 \le i \le n, \tag{4.1}$$

where ℓ_{ij}, ℓ_i are constants, ϕ is a C^∞ function on $\mathbb{R}^n$, and nonlinear observation satisfying the following equation

$$\begin{aligned}\sum_{i=1}^{m} h_i^2(x) + \Delta\phi + |\nabla\phi|^2 + 2\nabla\phi\cdot(f-\nabla\phi) \\ = \sum_{i,j=1}^{n} q_{ij}x_ix_j + \sum_{i=1}^{n} q_i x_i + q_0,\end{aligned} \tag{4.2}$$

where $q_{ij} = q_{ji}$, q_i, q_0 are constants.

We first remark that there are many examples which satisfy equation (4.2). For example, let M be a positive constant such that

$$h_i^2(x) \le M(1+|x|^2), \qquad 1 \le i \le m-1, \tag{4.3}$$
$$|f(x) - \nabla\phi(x)| \le M(1+|x|^2) \tag{4.4}$$
$$|\Delta\phi(x)| \le M(1+|x|^2) \tag{4.5}$$
$$|\nabla\phi(x)|^2 \le M(1+|x|^2) \tag{4.6}$$

$$\begin{aligned}h_m^2(x) = (m+5)M(1+|x|^2) - \sum_{i=1}^{m-1} h_i^2(x) \\ -\Delta\phi(x) - |\nabla\phi(x)|^2 - 2\nabla\phi(x)\cdot(f(x)-\nabla\phi(x).\end{aligned} \tag{4.7}$$

Then condition (4.2) is satisfied. The purpose of this section is to prove the following theorem.

Theorem 4.1: The unnormalized density of the filtering system (2.1) with affine drift (4.1), nonlinear observation (4.2) and Gaussian initial distribution can be computed in real time in a memoryless way.

In view of Theorem 3.2, in order to solve the nonlinear filtering problem with conditions (4.1), (4.2) it suffices to solve the following Kolmogorov equation in real time. For $\tau_1 \le t \le \tau_2$,

$$\begin{cases} \frac{\partial \widetilde{u}}{\partial t}(t,x) = \frac{1}{2}\Delta\widetilde{u}(t,x) - \sum_{j=1}^{n} f_j(x)\frac{\partial \widetilde{u}}{\partial x_j}(t,x) \\ -\big(\sum_{j=1}^{n}\frac{\partial f_j}{\partial x_j}(x) + \frac{1}{2}\sum_{j=1}^{m} h_j^2(x)\big)\widetilde{u}(t,x) \\ \widetilde{u}(0,x) = \psi(x). \end{cases} \tag{4.8}$$

We first simplify (4.8) by introducing the following transformation. Let

$$\begin{aligned}\widetilde{u} &= e^{\phi(x)}\widetilde{\widetilde{u}},\ \nabla\widetilde{u} = e^{\phi(x)}\widetilde{\widetilde{u}}\nabla\phi + e^{\phi(x)}\nabla\widetilde{\widetilde{u}} && (4.9)\\ \Delta\widetilde{u} &= e^{\phi(x)}\widetilde{\widetilde{u}}\Delta\phi + |\nabla\phi|^2 e^{\phi(x)}\widetilde{\widetilde{u}} && (4.10)\\ &\quad +2e^{\phi(x)}\nabla\phi\cdot\nabla\widetilde{\widetilde{u}} + e^{\phi(x)}\Delta\widetilde{\widetilde{u}}. \end{aligned}$$

Then by direct computation, we have

$$\begin{cases} \frac{\partial \widetilde{\widetilde{u}}}{\partial t} = \frac{1}{2}\Delta\widetilde{\widetilde{u}} - \sum_{i=1}^{n}(\sum_{j=1}^{n}\ell_{ij}x_j + \ell_i)\frac{\partial \widetilde{\widetilde{u}}}{\partial x_i} \\ \quad -[\frac{1}{2}\Delta\phi + \frac{1}{2}|\nabla\phi|^2 \\ \quad + \sum_{i=1}^{n}\frac{\partial\phi}{\partial x_i}(\sum_{j=1}^{n}\ell_{ij}x_j + \ell_i) \\ \quad + \sum_{i=1}^{n}\ell_{ii} + \frac{1}{2}\sum_{i=1}^{m}h_i^2(x)]\widetilde{\widetilde{u}} \\ \widetilde{\widetilde{u}}(0,x) = \psi(x)e^{-\phi(x)} \end{cases} \tag{4.11}$$

It is well known that any $\psi(x)e^{-\phi(x)}$ is well approximated by finite linear combination of Gaussians of the form $\alpha_1 G_1 + \cdots + \alpha_p G_p$ where α_i's are real numbers and G_i's are Gaussian distributions. Let $\widetilde{\widetilde{u}}_i$ be the solution of (4.11) with initial distribution G_i. Since (4.11) is a linear partial differential equation, it follows that the solution of (4.11) is of the form $\alpha_1\widetilde{\widetilde{u}}_1 + \cdots + \alpha_p\widetilde{\widetilde{u}}_p$. Therefore it remains to solve (4.11) with Gaussian initial distribution. The following Theorem 4.2 gives an explicit solution of (4.11) with nonlinear observation assumption (4.2), and Gaussian initial distribution in terms of solutions of ODEs.

Theorem 4.2: Consider the filtering system (2.1) with nonlinear drift (4.1), nonlinear observation (4.2), and Kolmogorov equation. For $\tau_1 \le t \le \tau_2$,

$$\begin{cases} \frac{\partial \widetilde{u}}{\partial t}(t,x) = \frac{1}{2}\Delta\widetilde{u}(t,x) \\ \qquad - \sum_{j=1}^{n} f_j(x)\frac{\partial \widetilde{u}}{\partial x_j}(t,x) \\ \qquad + \frac{1}{2}\sum_{j=1}^{m} h_j^2(x)\big)\widetilde{u}(t,x) \\ \widetilde{u}(\tau_1,x) = \exp[x^T A(\tau_1)x \\ \qquad + B^T(\tau_1)x + C(\tau_1)] \end{cases} \tag{4.12}$$

where $A(\tau_1) = (A_{ij}(\tau_1))$ is a $n\times n$ matrix, $B^T(\tau_1) = (B_1(\tau_1),\ldots,B_n(\tau_1))$, $x^T = (x_1,\ldots,x_n)$ are $1\times n$ matrix and $C(\tau_1)$ is a scalar. Then the solution of (4.6) is of the following form

$$\widetilde{u}(t,x) = \exp(x^T A x + B^T x + C) \tag{4.13}$$

where $A = A^T = (A_{ij}(t))$ is a $n\times n$ matrix valued function of t, $B^T = (B_1(t),\ldots,B_n(t))$ is a $1\times n$ matrix valued function of t, and $C(t)$ is a scalar function of t. Moreover $A(t)$, $B^T(t)$ and $C(t)$ satisfy the following system of nonlinear ODEs:

$$\begin{cases} \frac{dA}{dt}(t) = 2A^2(t) - A(t)L - L^T A(t) - \frac{1}{2}Q \\ A(t)\Big|_{t=\tau_1} = A(\tau_1) \end{cases} \tag{4.14}$$

$$\begin{cases} \frac{dB^T}{dt}(t) = 2B^T(t)A(t) - B^T(t)L \\ \qquad -2\ell^T A(t) - \frac{1}{2}q \\ B^T(t)\Big|_{t=\tau_1} = B^T(\tau_1) \end{cases} \tag{4.15}$$

$$\begin{cases} \frac{dC}{dt}(t) = tr\,A(t) + \frac{1}{2}B^T(t)B(t) - \ell^T B(t) \\ \qquad -\frac{1}{2}q_0 - tr\,L \\ C(t)\Big|_{t=\tau_1} = C(\tau_1) \end{cases} \tag{4.16}$$

where $L = (\ell_{ij})$, $Q = (q_{ij})$, $1 \le i,j \le n$, $\ell^T = (\ell_1,\ldots,\ell_n)$, $q^T = (q_1,\ldots,q_n)$ as in (4.1) and (4.2).

Proof: Differentiating (4.13) with respect to t and x respectively, we get the following equations.

$$\begin{aligned}
&\frac{\partial\widetilde{u}}{\partial t} = \left(x^T\frac{dA}{dt}x + \frac{dB^T}{dt}x + \frac{dC}{dt}\right)\widetilde{u}\\
&\nabla\widetilde{u} = [(A+A^T)x+B]e^{x^TAx+B^Tx+C}\\
&\Delta\widetilde{u} = \{2trA + [(A+A^T)x+B]^T\\
&[(A+A^T)x+B]\}e^{x^TAx+B^Tx+C}\\
&= [x^T(AA^T+A^TA+2A^2)x \qquad\qquad (4.17)\\
&+2B^T(A+A^T)x+2trA+B^TB]\widetilde{u}\\
&\sum_{j=1}^n f_j(x)\frac{\partial\widetilde{u}}{\partial x_j} = (Lx+\ell)^T\nabla\widetilde{u}\\
&= [x^T(A^T+A)Lx+(B^TL+\ell^TA\\
&+\ell^TA^T)x+\ell^TB]\widetilde{u}
\end{aligned}$$

where $L=(\ell_{ij}), \ell^T=(\ell_1,\ldots,\ell_n)$

$$\left(\sum_{j=1}^m\frac{\partial f_j}{\partial x_j}(x)+\frac{1}{2}\sum_{j=1}^m h_j^2(x)\right)\widetilde{u}(t,x) = \\
= \left(\frac{1}{2}x^TQx+\frac{1}{2}q^Tx+\frac{1}{2}q_0+trL\right)\widetilde{u}(t,x)$$

where $Q=(q_{ij}), q^T=(q_1,\ldots,q_n)$. Therefore the R.H.S. of (4.12) is given by

$$\begin{aligned}
& \frac{1}{2}\Delta\widetilde{u}(t,x)-\sum_{j=1}^n f_j(x)\frac{\partial\widetilde{u}}{\partial x_j}(t,x)\\
& -(\sum_{j=1}^n\frac{\partial f_j}{\partial x_j}(x)+\frac{1}{2}\sum_{j=1}^m h_j^2(x))\widetilde{u}(t,x)\\
= \quad & [x^T\left(\frac{1}{2}AA^T+\frac{1}{2}A^TA+A^2\right)x\\
& +B^T(A+A^T)x+trA+\frac{1}{2}B^TB]\widetilde{u}\\
& -[x^T(A^T+A)Lx+(B^TL+\ell^TA \qquad\qquad (4.18)\\
& +\ell^TA^T)x+\ell^TB]\widetilde{u}\\
& -(\frac{1}{2}x^TQx+\frac{1}{2}q^Tx+\frac{1}{2}q_0+trL)\widetilde{u}(t,x)\\
= \quad & [x^T(\frac{1}{2}AA^T+\frac{1}{2}A^TA+A^2-A^TL\\
& -AL-\frac{1}{2}Q)x+(B^TA+B^TA^T\\
& -B^TL-\ell^TA-\ell^TA^T-\frac{1}{2}q^T)x+trA\\
& +\frac{1}{2}B^TB-\ell^TB-\frac{1}{2}q_0-trL]\widetilde{u}
\end{aligned}$$

By comparing (4.17) and (4.18), we get equations (4.14), (4.15) and (4.16) which are necessary and sufficient conditions for (4.13) to be a solution of (4.12). ∎

5. Conclusion

All the known finite dimensional filters require observation terms linear in nature. In this paper we have solved the nonlinear filtering problem with nonlinear drift and nonlinear observations in real time and memoryless manner. We first show that the solution of DMZ equation can be obtained by solving Kolmogorov equation. We also show that the Kolmogorov equation can be solved via solutions of systems of ODEs if the condition (4.2) is satisfied.

References

[1] A. Bensoussan, R. Glowinski and A. Rascanu, "Approximation of the Zakai equation by splitting up method", in Stochastic Systems and Optimization (Warsaw, 1988) (J. Zabczyk, ed.), Springer-Verlag, (1989), 257-265.

[2] R. W. Brockett, "Nonlinear systems and nonlinear estimation theory," in *The Mathematics of Filtering and Identification and Applications*, M. Hazewinkel and J. C. Willems, Eds. Dordrecht: Reidel, 1981.

[3] M. Chaleyat-Maurel and D. Michel, Des resultats de non-existence de filtre de dimension finie, *Stochastics*, 13 (1984), 83-102.

[4] J. Chen and S.S.-T. Yau, Finite dimensional filters with nonlinear drift VI: Linear structure of Ω matrix, *Math. Control Signals Systems*, 9 (1996), 370-385.

[5] W. L. Chiou and S.S.-T. Yau, Finite dimensional filters with nonlinear drift II: Brockett's problem on classification of finite dimensional estimation algebra, *SIAM J. Control and Optimization*, vol. 32, no. 1 (1994), 297-310.

[6] M. H. A. Davis, "On a multiplicative functional transformation arising in nonlinear filtering theory," *Z. Wahrsch Verw. Gebiete*, vol. 54, (1980), 125-139.

[7] R. J. Elliott and R. Glowinski, Approximations to solutions of the Zakai filtering equation, Stoch. Anal. Appl., 7, (1988), 145-168.

[8] P. Florchinger and F. LeGland, Time-discretization of the Zakai equation for diffusion processes observed in correlated noise, Analysis and Optimization of Systems (A. Bensoussan and J.J. Lions, eds.), Springer-Verlag, Lecture Notes in Inform. Sci. and Control, 144 (4), pp. 228-237, Stoachastics and Stochastics Reports 35 (4), (1991), 233-256.

[9] R. Mikulevicious and B. L. Rozovskii, Separation of observations and parameters in nonlinear filtering, Proceedings of the 32nd IEEE Conference on Decision and Control, San Antonia, Texas, Dec. (1993), 1564-1559.

[10] S. K. Mitter, "On the analogy between mathematical problems of nonlinear filtering and quantum physics," *Ricerche Automat.*, vol. 10, (1979), 163-216.

[11] L.-F. Tam, W. S. Wong, and S. S.-T. Yau, "On a necessary and sufficient condition for finite dimensionality of estimation algebras," *SIAM J. Contr. Optimiz.*, vol. 28, no. 1, (1990), 173-185.

[12] S. S.-T. Yau, "Recent results on nonlinear filtering: New class of finite dimensional filters," in *Proc. 29th Conf. Decision and Control*, Honolulu, HI, Dec. 1990, pp. 231-233.

[13] S. S.-T. Yau, "Finite dimensional filters with nonlinear drift I: A class of filters including both Kalman-Bucy filters and Benes filters," *J. Math. Syst., Estimation and Control*, vol. 4, no. 2, (1994), 181-203.

[14] S. S.-T. Yau, Complete classification of finite-dimensional estimation algebras of maximal rank, International Journal of Control, vol. 76, no. 7 (2003), 657-677.

[15] S. S.-T. Yau and S.-T. Yau, Solution of filtering problem with nonlinear observations, (to appear), SIAM J. Control and Optimization.

[16] S.-T. Yau and G.-Q. Hu, "Direct method without Riccati equation for Kalman-Bucy filtering system with arbitrary initial conditions," in *Proc. 13th World Congr. IFAC*, vol. H, San Francisco, CA, June 30-July 5, 1996, pp. 469-474.

[17] S. S.-T. Yau and G. Q. Hu, Finite dimensional filters with nonlinear drift X: Explicit solution of DMZ equation, IEEE Transactions on Automatic Control, vol. 46, no. 1, (2001), 142-148.

[18] S. S.-T. Yau and G. Q. Hu, Finite dimensional filters with nonlinear drift XIV: Classification of finite-dimensional estimation algebras of maximal rank with arbitrary state space dimension and Mitter conjecture, (to appear) International Journal of Control.

[19] S. S.-T. Yau, X. Wu and W. S. Wong, Hessian Matrix non-decomposition theorem, Mathematical Research Letter, 6 (1999), 1-11.

[20] S. S.-T. Yau and S.-T. Yau, "Finite dimensional filters with nonlinear drift III: Duncan-Mortensen-Zakai equation with arbitrary initial condition for linear filtering system and the Benes filtering system," *IEEE Trans. Aerosp. Elec. Syst.*, vol. 33, pp. 1277-1294, Oct. 1997.

[21] S. S.-T. Yau and S.-T. Yau, "New direct method for Kalman-Bucy filtering system with arbitrary initial condition," in *Proc. 33rd Conf. Decision and Control*, Lake Buena Vista, FL, Dec. 14-16, 1994, pp. 1221-1225.

[22] S. S.-T. Yau and S.-T. Yau, "Explicit solution to a Kolmogorov equation," *Appl. Math. Optimiz. Int. J.*, vol. 34, (1996), 231-266.

[23] S.-T. Yau and S. S.-T. Yau, Real time solution of nonlinear filtering problem without memory I, Mathematical Research Letter, 7 (2000), 671-693.

[24] S.-T. Yau and S. S.-T. Yau, Existence and uniqueness and decay estimates for the time dependent parabolic equation with application to Duncan-Mortensen-Zakai equation, Asian Journal of Mathematics, vol. 2 (1998), 1079-1149.