SIAM J. CONTROL OPTIM.
Vol. 44, No. 3, pp. 1019–1039

SOLUTION OF FILTERING PROBLEM WITH NONLINEAR OBSERVATIONS*

STEPHEN S.-T. YAU† AND SHING-TUNG YAU‡

Abstract. For all known finite-dimensional filters, one always needs the conditon 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 includes Kalman–Bucy filter as a special case.

Key words. filtering problem, nonlinear observations, real time computation, DMZ equation, Kolmogorov equation

AMS subject classifications. 93E10, 93E11, 60G35

DOI. 10.1137/S0363012902411970

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 [10]. As pointed out by Mitter [13], the difficulty with this approach is that the innovations process is not, in general, explicitly computable. In view of this weakness, Brockett [2] and Mitter [13] 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 [4]. In [23], [17], and [20] 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 [17] and Tam, Wong, and Yau [14], [16], [5], [21], [20], and [6] 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.

However, in the Wei–Norman approach, one has to know explicitly the basis as vector space of the estimation algebra in order to reduce the DMZ equation to a finite system of ordinary differential equations, Kolmogorov equation, and several first-order linear partial differential equations. Classically, one knows the explicit

*Received by the editors July 24, 2002; accepted for publication (in revised form) December 8, 2004; published electronically September 20, 2005. Research partially supported by U.S. Army Research Office and NSF.

http://www.siam.org/journals/sicon/44-3/41197.html

†Department of Mathematics, Statistics and Computer Science (M/C 249), University of Illinois at Chicago, 851 South Morgan Street, Chicago, IL 60607-7045 (yau@uic.edu). Ze-Jiang Professor of East China Normal University.

‡Department of Mathematics, Harvard University, Cambridge, MA 02138 (yau@math.harvard.edu).

basis for the estimation algebra only in the case that it has maximal rank. Typically people assume that the linear system is controllable and observable. 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) are independent of time (cf. [22], [23], [19], [18]). 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 the direct methods in [22], [23], [18], and [19] they need to assume that all the observation terms $h_i(x)$, $1 \leq i \leq m$, are degree one polynomials.

In [26], 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) have linear growth. In this paper, we use the theory developed in [26] to show that the real time computation of the 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 [8], Florchinger and LeGland [9], Mikulevicious and Rozovskii [12]. Unlike our results, however, their results cannot cover Kalman–Bucy filters. Theorem 4.2 of this paper says that if the drifts $(f(x))$ are affine 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 includes Kalman–Bucy 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 the DMZ equation can be reduced to off time computation of the Kolmogorov equation. An explicit algorithm of such a reduction is provided. In the appendix, we give a rigorous proof that the solution of our algorithm converges to the solution of the DMZ equation in pointwise and L^2 sense. In section 4, we show that if the drifts are linear 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 linear drifts and nonlinear observations with linear growth 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 the robust DMZ-equation:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,x) = L_0u(t,x) + \displaystyle\sum_{i=1}^{m} y_i(t)[L_0, L_i]u(t,x) \\ \qquad\qquad + \dfrac{1}{2}\displaystyle\sum_{i,j=1}^{m} y_i(t)y_j(t)[[L_0, L_i], L_j]u(t,x) \\ u(0,x) = \sigma_0(x), \end{cases} \tag{2.4}$$

where $[\cdot\,,\,\cdot]$ is the Lie bracket as described in [14]. It is easy to show [24] that (2.4) is equivalent to the following time-varying partial differential equation:

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

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 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 proposition.

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

$$\frac{\partial \widetilde{u}}{\partial t}(t,x) = \frac{1}{2}\Delta \widetilde{u}(t,x) - \sum_{i=1}^{n} f_i(x)\frac{\partial \widetilde{u}}{\partial x_i}(t,x) - \left(\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}(x) + \frac{1}{2}\sum_{i=1}^{m} h_i^2(x)\right)\widetilde{u}(t,x) \tag{3.1}$$

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 $y(\tau_1)$,

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

Proof. It is straightforward to show that

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

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

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

Suppose that $u(t,x)$ is the solution of the 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 $\tau_{i-1} \le t \le \tau_i$:

$$
(3.4)\quad \begin{cases} \dfrac{\partial u_i}{\partial t}(t,x) = \dfrac{1}{2}\Delta u_i(t,x) + \displaystyle\sum_{\ell=1}^{n}\left(-f_\ell(x) + \sum_{j=1}^{m} y_j(\tau_{i-1})\frac{\partial h_j}{\partial x_\ell}(x)\right)\frac{\partial u_i}{\partial x_\ell}(t,x) \\ \qquad -\left(\displaystyle\sum_{\ell=1}^{n}\frac{\partial f_\ell}{\partial x_\ell}(x) + \frac{1}{2}\sum_{\ell=1}^{m} h_\ell^2(x) - \frac{1}{2}\sum_{j=1}^{m} y_j(\tau_{i-1})\Delta h_j(x)\right. \\ \qquad +\displaystyle\sum_{j=1}^{m}\sum_{\ell=1}^{n} y_j(\tau_{i-1}) f_\ell(x)\frac{\partial h_j}{\partial x_\ell}(x) \\ \qquad \left. -\dfrac{1}{2}\displaystyle\sum_{p=1}^{n}\sum_{j,\ell=1}^{m} y_j(\tau_{i-1}) y_\ell(\tau_{i-1})\frac{\partial h_j}{\partial x_p}(x)\frac{\partial h_\ell}{\partial x_p}(x)\right) u_i(t,x) \\ u_i(\tau_{i-1},x) = u_{i-1}(\tau_{i-1},x). \end{cases}
$$

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 Appendices A and B, we shall show that in both pointwise and L^2 sense

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

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:

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

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

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

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:

$$
(3.8)\quad \begin{cases} \dfrac{\partial \widetilde{u}_i}{\partial t}(t,x) = \dfrac{1}{2}\Delta\widetilde{u}_i(t,x) - \displaystyle\sum_{j=1}^{n} f_j(x)\frac{\partial \widetilde{u}_i}{\partial x_j}(t,x) - \left(\sum_{j=1}^{n}\frac{\partial f_j}{\partial x_j}(x) + \frac{1}{2}\sum_{j=1}^{m} h_j^2(x)\right)\widetilde{u}_i(t,x) \\ \widetilde{u}_i(\tau_{i-1},x) = e^{\sum_{j=1}^{m}(y_j(\tau_{i-1}) - y_j(\tau_{i-2}))h_j(x)}\,\widetilde{u}_{i-1}(\tau_{i-1},x), \end{cases}
$$

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 the 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

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

Observe that in our algorithm at step i (Lemma B.2), 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 affine drift,

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

where ℓ_{ij}, ℓ_i are constants, and nonlinear observation

$$\sum_{i=1}^m h_i^2(x) = \sum_{i,j=1}^n q_{ij}x_ix_j + \sum_{i=1}^n q_ix_i + q_0, \tag{4.2}$$

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

We first remark that if $h_i(x)$, $1 \le i \le m$, are nonlinear observation with linear growths as follows:

$$h_i^2(x) \le m(1+|x|^2), \qquad 1 \le i \le m-1, \tag{4.3}$$

where M is a constant, and

$$h_m^2(x) = (m-1)M(1+|x|^2) - \sum_{i=1}^{m-1} h_i^2(x), \tag{4.4}$$

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 $\tau_1 \le t \le \tau_2$,

$$\begin{cases} \dfrac{\partial \widetilde{u}}{\partial t}(t,x) = \dfrac{1}{2}\Delta \widetilde{u}(t,x) - \displaystyle\sum_{j=1}^n f_j(x)\frac{\partial \widetilde{u}}{\partial x_j}(t,x) - \left(\sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(x) + \frac{1}{2}\sum_{j=1}^m h_j^2(x)\right)\widetilde{u}(t,x) \\ \widetilde{u}(0,x) = \phi(x). \end{cases} \tag{4.5}$$

It is well known that any $\phi(x)$ is well approximated by finite linear combination of Gaussians of the form $\alpha_1 G_1 + \cdots + \alpha_p G_p$, where α_is are real numbers and G_is are Gaussian distributions. Let $\widetilde{u}_i$ be the solution of (4.5) with initial distribution G_i. Since (4.5) is a linear partial differential equation, it follows that the solution of (4.5) is of the form $\alpha_1\widetilde{u}_1 + \cdots + \alpha_p\widetilde{u}_p$. Therefore it remains to solve (4.5) with Gaussian initial distribution. Theorem 4.2 gives an explicit solution of (4.5) with linear drift (4.1), nonlinear observation (4.2), and Gaussian initial distribution in terms of solutions of ODEs.

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

$$\begin{cases} \dfrac{\partial \widetilde{u}}{\partial t}(t,x) = \dfrac{1}{2}\Delta \widetilde{u}(t,x) - \displaystyle\sum_{j=1}^n f_j(x)\frac{\partial \widetilde{u}}{\partial x_j}(t,x) - \left(\sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(x) + \frac{1}{2}\sum_{j=1}^m h_j^2(x)\right)\widetilde{u}(t,x) \\ \widetilde{u}(\tau_1,x) = \exp[x^T A(\tau_1)x + B^T(\tau_1)x + C(\tau_1)], \end{cases} \tag{4.6}$$

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

$$\widetilde{u}(t,x) = \exp(x^T Ax + B^T x + C), \tag{4.7}$$

where $A = A^T = (A_{ij}(t))$ is a $n \times n$ matrix valued function of t, $B^T = (B_1(t), \dots, B_n(t))$ 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*:*

$$\begin{cases} \dfrac{dA}{dt}(t) = 2A^2(t) - A(t)L - L^T A(t) - \dfrac{1}{2}Q \\ A(t)|_{t=\tau_1} = A(\tau_1) \end{cases} \tag{4.8}$$

$$\begin{cases} \dfrac{dB^T}{dt}(t) = 2B^T(t)A(t) - B^T(t)L - 2\ell^T A(t) - \dfrac{1}{2}q \\ B^T(t)|_{t=\tau_1} = B^T(\tau_1) \end{cases} \tag{4.9}$$

$$\begin{cases} \dfrac{dC}{dt}(t) = tr\, A(t) + \dfrac{1}{2}B^T(t)B(t) - \ell^T B(t) - \dfrac{1}{2}q_0 - tr\, L \\ C(t)|_{t=\tau_1} = C(\tau_1), \end{cases} \tag{4.10}$$

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

Proof. Differentiating (4.7) with respect to t and x, respectively, we get the following equations:

$$\begin{aligned} \frac{\partial \widetilde{u}}{\partial t} &= \left(x^T \frac{dA}{dt} x + \frac{dB^T}{dt} x + \frac{dC}{dt} \right) \widetilde{u} \\ \nabla \widetilde{u} &= [(A + A^T)x + B] e^{x^T Ax + B^T x + C} \\ \Delta \widetilde{u} &= \{2trA + [(A + A^T)x + B]^T [(A + A^T)x + B]\} e^{x^T Ax + B^T x + C} \\ &= [x^T (AA^T + A^T A + 2A^2)x + 2B^T (A + A^T)x + 2trA + B^T B]\widetilde{u} \\ \sum_{j=1}^n f_j(x) \frac{\partial \widetilde{u}}{\partial x_j} &= (Lx + \ell)^T \nabla \widetilde{u} \\ &= [x^T (A^T + A)Lx + (B^T L + \ell^T A + \ell^T A^T)x + \ell^T B]\widetilde{u}, \end{aligned} \tag{4.11}$$

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

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

where $Q = (q_{ij}), q^T = (q_1, \dots, q_n)$.

Therefore the R.H.S. of (4.6) is given by

$$\begin{aligned} &\frac{1}{2}\Delta \widetilde{u}(t, x) - \sum_{j=1}^n f_j(x) \frac{\partial \widetilde{u}}{\partial x_j}(t, x) - \left(\sum_{j=1}^n \frac{\partial f_j}{\partial x_j}(x) + \frac{1}{2} \sum_{j=1}^m h_j^2(x) \right) \widetilde{u}(t, x) \\ &= \left[x^T \left(\frac{1}{2}AA^T + \frac{1}{2}A^T A + A^2 \right) x + B^T (A + A^T)x + trA + \frac{1}{2}B^T B \right] \widetilde{u} \\ &\quad -[x^T (A^T + A)Lx + (B^T L + \ell^T A + \ell^T A^T)x + \ell^T B]\widetilde{u} \\ &\quad - \left(\frac{1}{2} x^T Q x + \frac{1}{2} q^T x + \frac{1}{2} q_0 + trL \right) \widetilde{u}(t, x) \\ &= \left[x^T \left(\frac{1}{2}AA^T + \frac{1}{2}A^T A + A^2 - A^T L - AL - \frac{1}{2}Q \right) x + (B^T A + B^T A^T - B^T L \right. \\ &\quad \left. -\ell^T A - \ell^T A^T - \frac{1}{2}q^T)x + trA + \frac{1}{2}B^T B - \ell^T B - \frac{1}{2}q_0 - trL \right] \widetilde{u}. \end{aligned} \tag{4.12}$$

By comparing (4.11) and (4.12), we get (4.8), (4.9), and (4.10), which are necessary and sufficient conditions for (4.7) to be a solution of (4.6). □

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 linear drift and nonlinear observations in real time and memoryless manner. We first show that the solution of the DMZ equation can be obtained by solving the Kolmogorov equation. We also show that the Kolmogorov equation can be solved via solutions of systems of ODEs if the summation of observations is a quadratic polynomial (cf. (4.2)).

Appendix A: Pointwise Convergence of (3.5). By changing variables from x_i to $\sqrt{2}x_i$ and by letting

$$\overline{u}(t,x) = u\left(t, \frac{x}{\sqrt{2}}\right), \tag{A.1}$$

we get

$$\begin{aligned}
\frac{\partial \overline{u}}{\partial t}(t,x) &= \frac{\partial u}{\partial t}\left(t, \frac{x}{\sqrt{2}}\right),\\
\frac{\partial \overline{u}}{\partial x_i}(t,x) &= \frac{1}{\sqrt{2}}\frac{\partial u}{\partial x_i}\left(t, \frac{x}{\sqrt{2}}\right),\\
\frac{\partial^2 \overline{u}}{\partial x_i^2}(t,x) &= \frac{1}{2}\frac{\partial^2 u}{\partial x_i^2}\left(t, \frac{x}{\sqrt{2}}\right).
\end{aligned}$$

Hence the robust DMZ equation becomes

$$\frac{\partial\, \overline{u}}{\partial t}(t,x) = \Delta\overline{u}(t,x) + \sum_{i=1}^{m}\overline{f}_i(t,x)\frac{\partial\, \overline{u}}{\partial x_i}(t,x) - \overline{V}(t,x)\overline{u}(t,x), \tag{A.2}$$

where

$$\overline{f}_i(t,x) = \sqrt{2}\left[-f_i\left(\frac{x}{\sqrt{2}}\right) + \sum_{j=1}^{m} y_j(t)\frac{\partial h_j}{\partial x_i}\left(\frac{x}{\sqrt{2}}\right)\right] \tag{A.3}$$

$$\begin{aligned}
\overline{V}(t,x) = {}& \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i}\left(\frac{x}{\sqrt{2}}\right) + \frac{1}{2}\sum_{i=1}^{m} h_i^2\left(\frac{x}{\sqrt{2}}\right) - \sum_{i=1}^{m} y_i(t)\Delta h_i\left(\frac{x}{\sqrt{2}}\right)\\
&+ \sum_{i=1}^{m}\sum_{j=1}^{n} y_i(t) f_j\left(\frac{x}{\sqrt{2}}\right)\frac{\partial h_i}{\partial x_j}\left(\frac{x}{\sqrt{2}}\right)\\
&- \frac{1}{2}\sum_{i,j=1}^{m}\sum_{k=1}^{n} y_i(t)y_j(t)\frac{\partial h_i}{\partial x_k}\left(\frac{x}{\sqrt{2}}\right)\frac{\partial h_j}{\partial x_k}\left(\frac{x}{\sqrt{2}}\right).
\end{aligned} \tag{A.4}$$

For any $\tau > 0$, we shall consider the following parabolic equations on $[0,\tau]\times\mathbb{R}^n$.

$$\begin{cases}
\dfrac{\partial\, \overline{u}}{\partial t}(t,x) = \Delta\overline{u}(t,x) + \displaystyle\sum_{i=1}^{n}\overline{f}_i(t,x)\frac{\partial\, \overline{u}}{\partial x_i}(t,x) - \overline{V}(t,x)\overline{u}(t,x)\\
\overline{u}(0,x) = \overline{\psi}(x)
\end{cases} \tag{A.5}$$

$$\text{(A.6)} \qquad \begin{cases} \dfrac{\partial \widetilde{u}}{\partial t}(t,x) = \Delta \widetilde{u}(t,x) + \displaystyle\sum_{i=1}^{n} \widetilde{f}_i(0,x) \dfrac{\partial \widetilde{u}}{\partial x_i}(t,x) - \widetilde{V}(0,x)\widetilde{u}(t,x) \\ \widetilde{u}(0,x) = \widetilde{\psi}(x), \end{cases}$$

where $\widetilde{f}_i(0,x) := \overline{f}_i(0,x)$ and $\widetilde{V}(0,x) := \overline{V}(0,x)$ are obtained from $\overline{f}_i(t,x)$ and $\overline{V}(t,x)$ by freezing the time variable at 0. For simplicity, we shall assume that the first, second, and third derivatives of $h(x)$ are bounded.

The goal of this appendix is to prove that if $\widetilde{\psi}(x)$ is close to $\overline{\psi}(x)$ uniformly in x, then $\widetilde{u}(\tau,x)$ is close to $\overline{u}(\tau,x)$ uniformly in x. From (A.5) and (A.6), we deduce that

(A.7)

$$\begin{aligned} \frac{\partial(\overline{u}-\widetilde{u})}{\partial t}(t,x) &= \Delta(\overline{u}-\widetilde{u})(t,x) + \sum_{i=1}^{n} \overline{f}_i(t,x) \frac{\partial(\overline{u}-\widetilde{u})}{\partial x_i}(t,x) - \overline{V}(t,x)(\overline{u}-\widetilde{u})(t,x) \\ &\quad + \sum_{i=1}^{n} (\overline{f}_i(t,x) - \widetilde{f}_i(0,x)) \frac{\partial \widetilde{u}}{\partial x_i}(t,x) - (\overline{V}(t,x) - \widetilde{V}(0,x))\widetilde{u}(t,x) \\ &= (\Delta - \overline{V}(t,x))(\overline{u}-\widetilde{u})(t,x) + \sum_{i=1}^{n} \overline{f}_i(t,x) \frac{\partial(\overline{u}-\widetilde{u})}{\partial x_i}(t,x) + G_\tau(t,x), \end{aligned}$$

where

$$\text{(A.8)} \qquad G_\tau(t,x) = \sum_{i=1}^{n} (\overline{f}_i(t,x) - \widetilde{f}_i(0,x)) \frac{\partial \widetilde{u}}{\partial x_i}(t,x) - (\overline{V}(t,x) - \widetilde{V}(0,x))\widetilde{u}(t,x).$$

LEMMA A.1. *There exists a nonnegative function $\alpha(t,x,y)$ such that*

$$\text{(A.9)} \qquad \begin{cases} \dfrac{\partial \alpha}{\partial t}(t,x,y) = \Delta_x \alpha(t,x,y) - \displaystyle\sum_{i=1}^{n} \overline{f}_i(\tau - t, x) \dfrac{\partial \alpha}{\partial x_i}(t,x,y) \\ \qquad\qquad - \left[\overline{V}(\tau - t, x) + \displaystyle\sum_{i=1}^{n} \dfrac{\partial \overline{f}_i}{\partial x_i}(\tau - t, x) \right] \alpha(t,x,y) \\ \alpha(0,x,y) = \delta_y(x), \qquad \int_x \alpha(0,x,y)dx = 1, \end{cases}$$

where $\int_x$ denotes the integration with respect to x variable.

Proof. Let $\beta_n(x,y)$ be a sequence of Gaussian with

$$\text{(A.10)} \qquad \int_x \beta_n(x,y)dx = 1 \quad \text{and} \quad \lim_{n\to\infty} \beta_n(x,y) = \delta_y(x).$$

In view of [26], there exists a solution $\alpha_n(t,x,y)$ with initial condition $\alpha_n(0,x,y) = \beta_n(x,y)$. By maximal principle, $\alpha(t,x,y) \geq 0$ for all $t \geq 0$. We shall take $\alpha(t,x,y) = \lim_{n\to\infty} \alpha_n(t,x,y)$. ∎

THEOREM A.2. *Let $w(t,x) = \overline{u}(t,x) - \widetilde{u}(t,x)$, where $\overline{u}$ and $\widetilde{u}$ are the solutions of the parabolic equations* (A.5) *and* (A.6)*, respectively. Let $\alpha(t,x,y)$ be the nonnegative function in* Lemma A.1. *Then*

$$w(\tau,y) = \int_x \alpha(\tau,x,y)w(0,x)dx + \int_0^\tau \int_x \alpha(t,x,y)G_\tau(t,x)dx,$$

where $G_\tau(t,x)$ *is given in* (A.8).

Proof.

$$\text{(A.11)}\quad \int_0^\tau \frac{d}{dt}\int_x \alpha(\tau-t,x,y)w(t,x)dx = -\int_0^\tau\int_x \frac{\partial\alpha}{\partial t}(\tau-t,x,y)w(t,x)dx + \int_0^\tau\int_x \alpha(\tau-t,x,y)\frac{\partial w}{\partial t}(t,x)dx$$

$$\begin{aligned}
\text{L.H.S. of (A.11)} &= w(\tau,y) - \int_x \alpha(\tau,x,y)w(0,x)dx\\
\text{R.H.S. of (A.11)} &= -\int_0^\tau\int_x \Delta_x\alpha(\tau-t,x,y)w(t,x)dx\\
&\quad + \int_0^\tau\int_x \sum_{i=1}^n \overline{f}_i(t,x)\frac{\partial\alpha}{\partial x_i}(\tau-t,x,y)w(t,x)dx\\
&\quad + \int_0^\tau\int_x \left[\overline{V}(t,x)+\sum_{i=1}^n\frac{\partial\overline{f}_i}{\partial x_i}(t,x)\right]\alpha(\tau-t,x,y)w(t,x)dx\\
&\quad + \int_0^\tau\int_x \alpha(\tau-t,x,y)\frac{\partial w}{\partial t}(t,x)dx\\
&= \int_0^\tau\int_x \alpha(\tau-t,x,y)\left[\frac{\partial w}{\partial t}(t,x) - \Delta w(t,x) - \sum_{i=1}^n \overline{f}_i(t,x)\frac{\partial w}{\partial x_i}(t,x)\right.\\
&\qquad\qquad \left. + \overline{V}(t,x)w(t,x)\right]dx\\
&= \int_0^\tau\int_x \alpha(\tau-t,x,y)G_\tau(t,x)dx. \qquad \text{by (A.7)}
\end{aligned}$$

In the above computation, we have used the fact proved in [26] that $\alpha(t,x,y)$ has Gaussian decay in x. ∎

PROPOSITION A.3. *Let* $\alpha(t,x,y)$ *be the nonnegative function in* Lemma A.1. *Suppose that* $\overline{V}(t,x)\geq -c_1$ *for some positive constant* c_1. *Then*

$$\text{(A.12)}\qquad \int_x \alpha(\tau,x,y)dx \leq e^{c_1\tau}.$$

Proof.

$$\begin{aligned}
e^{c_1 t}\frac{d}{dt}&\left(e^{-c_1 t}\int_x\alpha(t,x,y)dx\right) = -c_1\int_x\alpha(t,x,y)dx + \int_x\frac{\partial\alpha}{\partial t}(t,x,y)dx\\
&= -c_1\int_x\alpha(t,x,y)dx + \int_x\Delta_x\alpha(t,x,y)dx - \int_x\sum_{i=1}^n\overline{f}_i(\tau-t,x)\frac{\partial\alpha}{\partial x_i}(t,x,y)dx\\
&\quad - \int_x\left[\overline{V}(\tau-t,x)+\sum_{i=1}^n\frac{\partial\overline{f}_i}{\partial x_i}(\tau-t,x)\right]\alpha(t,x,y)dx\\
&= -c_1\int_x\alpha(t,x,y)dx - \int_x\overline{V}(\tau-t,x)\alpha(t,x,y)dx\\
&= -\int_x[\overline{V}(\tau-t,x)+c_1]\alpha(t,x,y)dx \leq 0.
\end{aligned}$$

It follows that $e^{-c_1 t}\int_x \alpha(t,x,y)dx$ is a decreasing function of t and (A.12) follows. ◻

THEOREM A.4. *With the assumption of* Proposition A.3, *let $w(t,x) = \overline{u}(t,x) - \widetilde{u}(t,x)$, where $\overline{u}$ and $\widetilde{u}$ are the solutions of the parabolic equations* (A.5) *and* (A.6), *respectively. If τ is small and $w(0,x)$ is small uniformly in τ, then $w(\tau,x)$ is small uniformly in x. More precisely, we have*

$$\sup_{y\in\mathbb{R}^n} |w(\tau,y)| \le e^{c_1\tau}\sup_{x\in\mathbb{R}^n}|w(0,x)| + \tau e^{c_1\tau}\sup_{\substack{x\in\mathbb{R}^n\\ 0\le t\le\tau}}|G_\tau(t,x)|, \tag{A.13}$$

where $G_\tau(t,x)$ is given in (A.8).

Proof. In view of (A.3), (A.4), and (A.8), we have

$$\begin{aligned}
G_\tau(t,x) &= \sum_{i=1}^n (\overline{f}_i(t,x) - \widetilde{f}_i(0,x))\frac{\partial\widetilde{u}}{\partial x_i}(t,x) - (\overline{V}(t,x) - \widetilde{V}(0,x))\widetilde{u}(t,x)\\
&= \sum_{i=1}^n \sqrt{2}\sum_{j=1}^m (y_j(t) - y_j(0))\frac{\partial h_j}{\partial x_i}\left(\frac{x}{\sqrt{2}}\right)\frac{\partial\widetilde{u}}{\partial x_i}(t,x)\\
&\quad + \left[-\sum_{i=1}^m (y_i(t)-y_i(0))\Delta h_i\left(\frac{x}{\sqrt{2}}\right) + \sum_{i=1}^m\sum_{j=1}^n (y_i(t)-y_i(0)) f_j\left(\frac{x}{\sqrt{2}}\right)\frac{\partial h_i}{\partial x_j}\left(\frac{x}{\sqrt{2}}\right)\right.\\
&\quad \left. -\frac{1}{2}\sum_{i,j=1}^m\sum_{k=1}^n (y_i(t)y_j(t) - y_i(0)y_j(0))\frac{\partial h_i}{\partial x_k}\left(\frac{x}{\sqrt{2}}\right)\frac{\partial h_j}{\partial x_k}\left(\frac{x}{\sqrt{2}}\right)\right]\widetilde{u}(t,x).
\end{aligned}$$

Therefore if τ is small, then $G_\tau(t,x)$ is uniformly small in x for $0\le t\le\tau$, because both $\widetilde{u}(t,x)$ and $\frac{\partial\widetilde{u}}{\partial x_i}(t,x)$ have Gaussian decay by [26]. The estimate (A.13) follows readily from Theorem A.2. ◻

Now we consider the global situation. For a fixed $T>0$, we want to find the solution $\overline{u}(t,x)$ of the following parabolic equation on $[0,T]\times\mathbb{R}^n$:

$$\begin{cases} \dfrac{\partial\,\overline{u}}{\partial t}(t,x) = \Delta\overline{u}(t,x) + \displaystyle\sum_{j=1}^n \overline{f}_j(t,x)\frac{\partial\,\overline{u}}{\partial x_j}(t,x) - \overline{V}(t,x)\overline{u}(t,x)\\ \overline{u}(0,x) = \overline{\psi}(x). \end{cases} \tag{A.14}$$

Let $\{0<\tau_1<\tau_2<\cdots<\tau_k = T\}$ be a partition of $[0,T]$. Let $\widetilde{u}_i(t,x)$ be the solution of the following parabolic equation on $[\tau_{i-1},\tau_i]\times\mathbb{R}^n$:

$$\begin{cases} \dfrac{\partial\widetilde{u}_i}{\partial t}(t,x) = \Delta\widetilde{u}_i(t,x) + \displaystyle\sum_{j=1}^n \widetilde{f}_j(\tau_{i-1},x)\frac{\partial\widetilde{u}_i}{\partial x_j}(t,x) - \widetilde{V}(\tau_{i-1},x)\widetilde{u}_i(t,x)\\ \widetilde{u}_i(\tau_{i-1},x) = \widetilde{u}_{i-1}(\tau_{i-1},x), \end{cases} \tag{A.15}$$

where $\widetilde{u}_1(0,x) = \overline{\psi}(x)$; $\widetilde{f}_j(\tau_{i-1},x)$ and $\widetilde{V}(\tau_{i-1},x)$ are functions independent of t and equal to $\overline{f}_j(\tau_{i-1},x)$ and $\overline{V}(\tau_{i-1},x)$, respectively.

LEMMA A.5. *Fix T, let $G_{\tau_i}(t,x) = \sum_{j=1}^n(\overline{f}_j(t,x) - \widetilde{f}_j(\tau_{i-1},x))\frac{\partial\widetilde{u}_i}{\partial x_j}(t,x) - (\overline{V}(t,x) - \widetilde{V}(\tau_{i-1},x))\widetilde{u}_i(t,x)$. For any given $\epsilon>0$, we can choose k sufficiently large so that*

$$\sup_{1\le i\le n}\ \sup_{\tau_{i-1}\le t\le\tau_i}\ \sup_{x\in\mathbb{R}^n}|G_{\tau_i}(t,x)|\le\epsilon.$$

Proof. This follows from the proof of Theorem A.4. ◻

We are now ready to prove the main theorem in this appendix.

THEOREM A.6. *Let $\overline{u}(t,x)$ and $\widetilde{u}_k(t,x)$ be the solutions of* (A.14) *and* (A.15)*, respectively. For any $\epsilon > 0$, let k be sufficiently large so that* Lemma A.5 *holds. Then*

$$|\overline{u}(T,x) - \widetilde{u}_k(T,x)| \leq \epsilon T e^{c_1 T},$$

where c_1 is the constant in Proposition A.3.

Proof. In view of $\widetilde{u}_1(0,x) = \overline{\psi}(x) = \overline{u}(0,x)$ and Theorem A.4, we have

$$|\overline{u}(\tau_{1,x}) - \widetilde{u}_1(\tau_{1,x})| \leq \tau_1 e^{c_1\tau_1} \sup_{\substack{x\in\mathbb{R}^n \\ 0\leq t\leq \tau_1}} |G_{\tau_1}(t,x)|.$$

By Theorem A.4 and induction, we have

$$\begin{aligned}
|\overline{u}(\tau_{2,x}) - \widetilde{u}_2(\tau_{2,x})| &\leq \tau_1 e^{c_1\tau_1} e^{c_1(\tau_2-\tau_1)} \sup_{\substack{x\in\mathbb{R}^n \\ 0\leq t\leq \tau_1}} |G_{\tau_1}(t,x)| \\
&\quad + (\tau_2 - \tau_1) e^{c_1(\tau_2-\tau_1)} \sup_{\substack{x\in\mathbb{R}^n \\ \tau_1\leq t\leq \tau_2}} |G_{\tau_2}(t,x)| \\
|\overline{u}(\tau_{k,x}) - \widetilde{u}_k(\tau_{k,x})| &\leq \tau_1 e^{c_1\tau_k} \sup_{\substack{x\in\mathbb{R}^n \\ 0\leq t\leq \tau_1}} |G_{\tau_1}(t,x)| + (\tau_2-\tau_1) e^{c_1(\tau_k-\tau_1)} \sup_{\substack{x\in\mathbb{R}^n \\ \tau_1\leq t\leq \tau_2}} G_{\tau_2}(t,x)| \\
&\quad + \cdots + (\tau_i - \tau_{i-1}) e^{c_1(\tau_k-\tau_{i-1})} \sup_{\substack{x\in\mathbb{R}^n \\ \tau_{i-1}\leq t\leq \tau_i}} |G_{\tau_i}(t,x)| \\
&\quad + \cdots + (\tau_k - \tau_{k-1}) e^{c_1(\tau_k-\tau_{k-1})} \sup_{\substack{x\in\mathbb{R}^n \\ \tau_{k-1}\leq t\leq \tau_k}} |G_{\tau_k}(t,x)| \\
&\leq \epsilon[\tau_1 e^{c_1\tau_k} + (\tau_2-\tau_1) e^{c_1(\tau_k-\tau_1)} + \cdots + (\tau_i-\tau_{i-1}) e^{c_1(\tau_k-\tau_{i-1})} \\
&\quad + \cdots + (\tau_k - \tau_{k-1}) e^{c_1(\tau_k-\tau_{k-1})}] \\
&\leq \epsilon[\tau_1 + (\tau_2-\tau_1) + \cdots + (\tau_i - \tau_{i-1}) + \cdots + (\tau_n - \tau_{n-1})] e^{c_1 T} \\
&= \epsilon T e^{c_1 T}. \qquad \square
\end{aligned}$$

THEOREM A.7. *Fix $T > 0$, let $\mathcal{P}_n = \{0 < \tau_1 < \tau_2 < \cdots < \tau_k = T\}$ be a partition of $[0,T]$. Let $\overline{u}(t,x)$ be the solution of the following parabolic equation on $[0,T]\times\mathcal{R}^n$:*

$$\begin{cases} \dfrac{\partial\,\overline{u}}{\partial t}(t,x) = \Delta\overline{u}(t,x) + \displaystyle\sum_{j=1}^n \overline{f}_j(t,x) \frac{\partial\,\overline{u}}{\partial x_j}(t,x) - \overline{V}(t,x)\overline{u}(t,x) \\ \overline{u}(0,x) = \psi(x). \end{cases}$$

Let $\overline{u}_i(t,x)$ be the solution of the following parabolic equation on $[\tau_{i-1},\tau_i]\times\mathcal{R}^n$:

$$\begin{cases} \dfrac{\partial\widetilde{u}_i}{\partial t}(t,x) = \Delta\widetilde{u}_i(t,x) + \displaystyle\sum_{j=1}^n \widetilde{f}_j(\tau_{i-1},x) \frac{\partial\widetilde{u}_i}{\partial x_j}(t,x) - \widetilde{V}(\tau_{i-1},x)\widetilde{u}_i(t,x) \\ \widetilde{u}_i(\tau_{i-1},x) = \widetilde{u}_{i-1}(\tau_{i-1},x), \end{cases}$$

where $\widetilde{u}_i(0,x) = \psi(x)$ *and* $\widetilde{f}_j(\tau_{i-1},x) = \overline{f}_j(\tau_{i-1},x)$, $\widetilde{V}(\tau_{i-1},x) = \overline{V}(\tau_{i-1},x)$ *are obtained from* $\overline{f}_j(t,x)$ *and* $\overline{V}(t,x)$ *by freezing time variable at* τ_{i-1}. *Then*

$$\overline{u}(\tau,x) = \lim_{|\mathcal{P}_k|\to 0} \widetilde{u}_k(\tau_k,x) \text{ uniformly in } x.$$

Appendix B: L^2 Convergence of (3.5). In Appendix A we have shown that the solution $\widetilde{u}(t,x)$ of (A.6) is uniformly close to the solution $\overline{u}(t,x)$ of (A.5) for $0 \le t \le T$ if $\widetilde{\psi}(x) = \widetilde{u}(0,x)$ is uniformly close to $\overline{\psi}(x) = \overline{u}(0,x)$. In this section, we shall show that $\widetilde{u}(t,x)$ is also close to $\overline{u}(t,x)$ in L^2-sense, if $\overline{\psi}(x)$ is close to $\widetilde{\psi}(x)$ in L^2 sense. We first recall the following lemma.

LEMMA B.1. *If* $\frac{d\alpha}{dt}(t) \le c\alpha(t)+\beta(t)$, *where c is a constant, then* $e^{-ct}\alpha(t)-\alpha(0) \le \int_0^t e^{-cs}\beta(s)ds$.

Let $\overline{f}_{2R}$, $\widetilde{f}_{2R}$, $\overline{V}_{2R}$, and $\widetilde{V}_{2R}$ be the functions obtained by multiplying $\overline{f}, \widetilde{f}, \overline{V}$, and $\widetilde{V}$, respectively, by a cut off function σ which is equal to one in the ball of radius $R \ge 1$ and equal to zero outside a ball of radius $2R$. We can choose σ such that

$$|\nabla\sigma(x)| \le \frac{4}{1+|x|} \text{ and } |\Delta\sigma(x)| \le \frac{4}{1+|x|^2}. \tag{B.1}$$

Consider the following equations:

$$\frac{\partial \overline{u}_{2R}}{\partial t} = \Delta\overline{u}_{2R} + \sum_{i=1}^n (\overline{f}_{2R})_i \frac{\partial \overline{u}_{2R}}{\partial x_i} - \overline{V}_{2R}\overline{u}_{2R} \tag{B.2}$$

$$\frac{\partial \widetilde{u}_{2R}}{\partial t} = \Delta\widetilde{u}_{2R} + \sum_{i=1}^n (\widetilde{f}_{2R})_i \frac{\partial \widetilde{u}_{2R}}{\partial x_i} - \widetilde{V}_{2R}\widetilde{u}_{2R} \tag{B.3}$$

in the ball B_{2R} of radius $2R$ with the Neumann condition, where $(f_{2R})_i$ and $(\widetilde{f}_{2R})_i$ denote the ith components of f_{2R} and $\widetilde{f}_{2R}$, respectively. Let $\overline{\psi}_{2R}(x) = \overline{\psi}(x)\sigma(x)$ and $\widetilde{\psi}_{2R}(x) = \widetilde{\psi}(x)\sigma(x)$ to be the initial conditions of (B.2) and (B.3), respectively. Then (B.2) and (B.3) have unique solutions, respectively, for $t \in [0,\infty)$ with Neumann condition on $\partial B_{2R} \times (0,T]$.

LEMMA B.2. *Assume that* (4.1)–(4.3) *hold and the first, second, and third derivatives of* $h_i(x)$ *are bounded. Let* $\widetilde{c}$ *and* δ *be positive constants such that* $\widetilde{\widetilde{c}} := \widetilde{c}+\delta < \frac{5}{254}$. *Choose* τ *and* ϵ *suitably small with* $\tau+\epsilon<\delta$. *Then the following conclusions hold for any* $0 \le t \le \tau$ *for both* $\rho \epsilon \{\overline{\rho},\widetilde{\rho}\}$, $u \in \{\overline{u},\widetilde{u}\}$, *and where* $\overline{\rho}(t,x) = \frac{\widetilde{c}(1+|x|^2)}{t+\epsilon}$, $\widetilde{\rho}(t,x) = \frac{\widetilde{\widetilde{c}}(1+|x|^2)}{t+\epsilon}$:

(i) $\displaystyle\int_{\{t\}\times B_{2R}} e^{\overline{\rho}}\overline{u}_{2R}^2 \le \int_{\{0\}\times B_{2R}} e^{\overline{\rho}}\overline{u}_{2R}^2$

(ii) $\displaystyle\int_{\{t\}\times B_{2R}} e^{\overline{\rho}}|\nabla\overline{u}_{2R}|^2 \le \int_{\{0\}\times B_{2R}} e^{\overline{\rho}}|\nabla\overline{u}_{2R}|^2 + \int_0^t \int_{B_{2R}} e^{\overline{\rho}(s,x)}|\overline{u}_{2R}(s,x)|^2$

(iii) $\displaystyle\int_{\{t\}\times B_{2R}} e^{\overline{\rho}}|\Delta\overline{u}_{2R}|^2 \le \int_{\{0\}\times B_{2R}} e^{\overline{\rho}}|\Delta\overline{u}_{2R}|^2$

$$+O\Bigg(\int_{[0,t]\times B_{2R}} e^{\overline{\rho}}|\nabla\overline{\rho}|^2|\overline{f}_{2R}|^2|\nabla\overline{u}_{2R}|^2+\int_{[0,t]\times B_{2R}} e^{\overline{\rho}}|\nabla\overline{f}_{2R}|^2|\nabla\overline{u}_{2R}|^2$$
$$+\int_{[0,t]\times B_{2R}} e^{\overline{\rho}}|\overline{f}_{2R}||\nabla\overline{u}_{2R}|^2|\Delta\overline{f}_{2R}|+\int_{[0,t]\times B_{2R}} e^{\overline{\rho}}|\overline{f}_{2R}|^4|\nabla\overline{u}_{2R}|^2$$
$$+\int_{[0,t]\times B_{2R}} e^{\overline{\rho}}|\nabla(\overline{V}_{2R}\overline{u}_{2R})|^2+\int_{[0,t]\times B_{2R}} e^{\overline{\rho}}|\nabla\overline{u}_{2R}|^2\left(\sum_{i=1}^n\frac{\partial(\overline{f}_{2R})_i}{\partial x_i}\right)^2\Bigg).$$

Moreover, the following inequalities hold for both $\{\overline{\rho},\overline{f},\overline{V}\}$, *or* $\{\overline{\rho},\widetilde{f},\widetilde{V}\}$ *or* $\{\widetilde{\rho},\overline{f},\overline{V}\}$ *or* $\{\widetilde{\rho},\widetilde{f},\widetilde{V}\}$ *if* δ *is small enough,*

(iv) $\displaystyle \frac{\partial\overline{\rho}}{\partial t}+2|\nabla\overline{\rho}|^2-\sum_{i=1}^n\overline{f}_i\frac{\partial\overline{\rho}}{\partial x_i}-\sum_{i=1}^n\frac{\partial\overline{f}_i}{\partial x_i}-2\overline{V}\le 0.$

Proof. (i), (ii), and (iii) follow from Lemma 1.3 of [26] by setting $\epsilon_1=\frac{1}{5}$ in that lemma. In equality (iv), it follows from

$$\frac{\partial\widetilde{\rho}}{\partial t}+2|\nabla\widetilde{\rho}|^2-\sum_{i=1}^n\overline{f}_i\frac{\partial\widetilde{\rho}}{\partial x_i}-\sum_{i=1}^n\frac{\partial\overline{f}_i}{\partial x_i}-2\overline{V}$$
$$\le\left[\frac{-\,\widetilde{\widetilde{c}}\,(1-8\,\widetilde{\widetilde{c}})}{(t+\epsilon)^2}+\frac{2c\,\widetilde{\widetilde{c}}}{t+\epsilon}+(n+2)c\right](1+|x|)^2$$

as $1-8\,\widetilde{\widetilde{c}}\ge 0$. □

PROPOSITION B.3. *Consider the parabolic differential equations* (A.5) *and* (A.6). *Let* ϕ *be any smooth function defined on* $\mathbb{R}^n$ *with compact support contained in a domain* Ω. *Let* $\overline{\rho}$ *be any smooth function on* $\mathbb{R}_+\times\mathbb{R}^n$ *satisfying*

$$\frac{\partial\overline{\rho}}{\partial t}+2|\nabla\overline{\rho}|^2-\sum_{i=1}^n\overline{f}_i\frac{\partial\overline{\rho}}{\partial x_i}-\sum_{i=1}^n\frac{\partial\overline{f}_i}{\partial x_i}-2\overline{V}\le 0. \tag{B.4}$$

Then

$$\frac{d}{dt}\int_{\{t\}\times\Omega}\phi^2e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2\le\int_{\{t\}\times\Omega}\phi^2e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2+10\int_{\{t\}\times\Omega}e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2|\nabla\phi|^2$$
$$+4\int_{\{t\}\times\Omega}e^{\overline{\rho}}\left|\sum_{i=1}^n\overline{f}_i\frac{\partial\phi}{\partial x_i}\right|^2(\overline{u}-\widetilde{u})^2+4\int_{\{t\}\times\Omega}e^{\overline{\rho}}\phi^2\left|\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\overline{\rho}}{\partial x_i}\right|^2\widetilde{u}^2$$
$$+2\int_{\{t\}\times\Omega}e^{\overline{\rho}}\phi^2\widetilde{u}|\overline{f}-\widetilde{f}|^2+4\int_{\{t\}\times\Omega}e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{V}-\widetilde{V}|^2$$
$$+2\int_{\{t\}\times\Omega}e^{\overline{\rho}}\phi^2|\overline{f}-\widetilde{f}|^2\widetilde{u}^2+4\int_{\{t\}\times\Omega}e^{\overline{\rho}}\phi^2\widetilde{u}^2\left|\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}-\frac{\partial\widetilde{f}_i}{\partial x_i}\right)\right|^2. \tag{B.5}$$

Proof. From (A.5) and (A.6), we deduce that

$$\frac{\partial(\overline{u}-\widetilde{u})}{\partial t}=\Delta(\overline{u}-\widetilde{u})+\sum_{i=1}^n f_i\frac{\partial(\overline{u}-\widetilde{u})}{\partial x_i}-V(u-\widetilde{u})+\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\widetilde{u}}{\partial x_i}-(\overline{V}-\widetilde{V})\widetilde{u}. \tag{B.6}$$

Then using (B.6) and integrating by part, we obtain

$$\begin{aligned}
\frac{d}{dt}\int_{\{t\}\times\Omega}\phi^2(\overline{u}-\widetilde{u})^2 e^{\overline{\rho}} \le & \int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})^2\left(\frac{\partial\overline{\rho}}{\partial t}+2|\nabla\overline{\rho}|^2-\sum_{i=1}^n \overline{f}_i\overline{\rho}_i-\sum_{i=1}^n\frac{\partial \overline{f}_i}{\partial x_i}-2\overline{V}\right)\\
&-\frac{1}{2}\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2|\nabla(\overline{u}-\widetilde{u})|^2+8\int_{\{t\}\times\Omega} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2|\nabla\phi|^2\\
&-2\int_{\{t\}\times\Omega}\phi e^{\overline{\rho}}\left(\sum_{i=1}^n\overline{f}_i\phi_i\right)(\overline{u}-\widetilde{u})^2-4\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi\left[\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\phi_i\right]\widetilde{u}(\overline{u}-\widetilde{u})\\
&-2\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\overline{\rho}}{\partial x_i}\widetilde{u}(\overline{u}-\widetilde{u})+2\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|f-\widetilde{f}|^2\\
&-2\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})\widetilde{u}\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}-\frac{\partial\widetilde{f}_i}{\partial x_i}\right)-2\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})\widetilde{u}(\overline{V}-\widetilde{V}).
\end{aligned}\tag{B.7}$$

In view of (B.4), (B.7) implies that

$$\begin{aligned}
\frac{d}{dt}\int_{\{t\}\times\Omega}\phi^2 e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2 \le\ & 8\int_{\{t\}\times\Omega} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2|\nabla\phi|^2+4\int_{\{t\}\times\Omega} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2\left|\sum_{i-1}^n\overline{f}_i\frac{\partial\phi}{\partial x_i}\right|^2\\
&+\frac{1}{4}\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})^2+\frac{1}{4}\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})^2\\
&+4\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2\left|\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\overline{\rho}}{\partial x_i}\right|^2\widetilde{u}^2+2\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{f}-\widetilde{f}|^2\\
&+\frac{1}{4}\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})^2+4\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{V}-\widetilde{V}|^2\\
&+4\left[\frac{1}{2}\int_{\{t\}\times\Omega} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2|\nabla\phi|^2+\frac{1}{2}\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2|\overline{f}-\widetilde{f}|^2\widetilde{u}^2\right]\\
&+2\left[\frac{1}{8}\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2(\overline{u}-\widetilde{u})^2+2\int_{\{t\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2\left|\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}-\frac{\partial\widetilde{f}_i}{\partial x_i}\right)\right|^2\right].
\end{aligned}$$

Inequality (B.5) follows immediately. □

The following theorem states that when τ is sufficiently small and $\overline{\psi}(x)$ close to $\widetilde{\psi}(x)$ in L^2-sense, then the solution $\widetilde{u}(t,x)$ of (A.6) approximates the solution $\overline{u}(t,x)$ of (A.5) well in L^2-sense.

THEOREM B.4. *Consider the parabolic differential equation* (A.5) *and* (A.6). *Assume that* (4.1)–(4.3) *hold and the first, second, and third derivatives of* $h_i(x)$ *are bounded. Let* $\widetilde{c}$ *and* δ *be positive constants such that* $\widetilde{\widetilde{c}} := \widetilde{c}+\delta < \frac{5}{254}$. *Let*

$$\overline{\rho}(t,x)=\frac{\widetilde{c}(1+|x|^2)}{t+\epsilon},\qquad \widetilde{\rho}(t,x)=\frac{\widetilde{\widetilde{c}}\ (1+|x|^2)}{t+\epsilon}.$$

Suppose that

$$\int_{\mathbb{R}^n} e^{\overline{\rho}(0,x)}(|\overline{\psi}(x)|^2+|\nabla\widetilde{\psi}(x)|^2+|\Delta\overline{\psi}(x)|^2)<\infty$$
$$\int_{\mathbb{R}^n} e^{\overline{\rho}(0,x)}(|\widetilde{\psi}(x)|^2+|\nabla\widetilde{\psi}(x)|^2+|\Delta\widetilde{\psi}(x)|^2)<\infty.$$

Choose τ and ϵ suitably small so that $\tau+\epsilon<\delta$ and the conclusions of Lemma B.2 *hold. Suppose that for $0\le t\le\tau$,*

$$|\overline{f}(t,x)-\widetilde{f}(t,x)|\le\widetilde{\epsilon}_1 c(1+|x|) \tag{B.8}$$

$$\left|\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}(t,x)-\frac{\partial\widetilde{f}_i}{\partial x_i}(t,x)\right)\right|\le\widetilde{\epsilon}_1 c \tag{B.9}$$

$$|\overline{V}(t,x)-\widetilde{V}(t,x)|\le\widetilde{\epsilon}_1 c(1+|x|^2) \tag{B.10}$$

$$\int_{\mathbb{R}^n}e^{\overline{\rho}(0,x)}|\overline{\psi}(x)-\widetilde{\psi}(x)|^2\le\widetilde{\epsilon}_2. \tag{B.11}$$

Then

$$\begin{aligned}\int_{\{t\}\times\mathbb{R}^n}e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2&\le\widetilde{\epsilon}_2e^t+16\widetilde{\epsilon}_1^2c^2\widetilde{c}^2\frac{t}{\epsilon(t+\epsilon)}e^td_1+24t\widetilde{\epsilon}_1^2c^2e^td_1\\&\le\widetilde{\epsilon}_2e^\tau+\widetilde{\epsilon}_1^2\tau e^\tau c_1,\end{aligned}$$

where $d_1=\int_{\mathbb{R}^n}e^{\overline{\rho}(0,x)}(\widetilde{\psi}(x))^2$, $c_1=\frac{16c^2\widetilde{c}^2d_1}{\epsilon^2}+24c^2d_1$, and c is a constant for linear growth of $\nabla\overline{V}$ and $\nabla\widetilde{V}$, i.e., $|\nabla\overline{V}(t,x)|\le c(1+|x|)$, and $|\nabla\widetilde{V}(t,x)|\le c(1+|x|)$.

Proof. Let $R_0\ge1$ and $B_{R_0}^c=\{x\in\mathbb{R}^n:|x|>R_0\}$ and

$$\phi(x)=\begin{cases}1&\text{for }\ |x|\le R_0\\ \frac{\log R-\log|x|}{\log R-\log R_0}&\text{for }\ R_0\le|x|\le R=2R_0\\ 0&\text{for }\ |x|\ge R=2R_0.\end{cases}$$

Let Ω be defined as B_R in Proposition B.3. In view of Lemma B.1 and (B.5), we have

$$\begin{aligned}&e^{-t}\int_{\{t\}\times\Omega}\phi^2e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2-\int_{\{0\}\times\Omega}\phi^2e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2\\&\quad\le10\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2|\nabla\phi|^2+4\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\rho}\left|\sum_{i=1}^n\overline{f}_i\frac{\partial\phi}{\partial x_i}\right|^2(\overline{u}-\widetilde{u})^2\\&\quad+4\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2\left|\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\rho}{\partial x_i}\right|^2\widetilde{u}^2+2\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{f}-\widetilde{f}|^2\\&\quad+4\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{V}-\widetilde{V}|^2+2\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2|\overline{f}-\widetilde{f}|^2\widetilde{u}^2\\&\quad+4\int_0^te^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2\widetilde{u}^2\left|\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}-\frac{\partial\widetilde{f}_i}{\partial x_i}\right)\right|^2\\&\quad\le\frac{10e^{\overline{\rho}(0,R)}}{R_0^2(\log R-\log R_0)^2}\int_0^te^{-s}\int_{\{s\}\times(B_{R_0}^C\cap B_R)}(\overline{u}-\widetilde{u})^2\\&\quad+\frac{4c^2(1+R)^2e^{\overline{\rho}(4R)}}{R_0^2(\log R-\log R_0)^2}\int_0^te^{-s}\int_{\{s\}\times(B_{R_0}^c\cap B_R)}(\overline{u}-\widetilde{u})^2\end{aligned}$$

$$
\begin{aligned}
&+4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\left|\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\overline{\rho}}{\partial x_i}\right|^2\widetilde{u}^2\\
&+4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{f}-\widetilde{f}|^2+4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{V}-\widetilde{V}|^2\\
&+4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2\left|\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}-\frac{\partial\widetilde{f}_i}{\partial x_i}\right)\right|^2. \qquad \text{(B.12)}
\end{aligned}
$$

Observe that (B.11) implies

$$
e^t\int_{\{0\}\times B_R} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2\le\widetilde{\epsilon}_2 e^t. \qquad \text{(B.13)}
$$

By Corollary 4.1 of [26], u and $\widetilde{u}$ decay like Gaussian in x variables. So we shall assume

$$
\max_{x\in\mathbb{R}^n}(|u|,|\widetilde{u}|)\le D_1e^{-D_2|x|^2}\ \text{for } t \text{ small}, \qquad \text{(B.14)}
$$

for some $D_1, D_2 > 0$. In view of the proof of Corollary 4.1 of [26], we can take $D_2\ge\frac{4\widetilde{c}}{\epsilon}+1$ for sufficiently small t

$$
\begin{aligned}
&\frac{[10+4c^2(1+R)^2]e^{\overline{\rho}(0,R)}}{R_0^2(\log R-\log R_0)^2}\int_0^t e^{-s}\int_{\{s\}\times(B_{R_0}^c\cap B_R)}(\overline{u}-\widetilde{u})^2\\
&\le\frac{4[10+4c^2(1+R)^2]te^{\overline{\rho}(0,R)}}{R_0^2(\log R-\log R_0)^2}\int_{B_{R_0}^c\cap B_r}D_1e^{-D_2|x|^2}\\
&\le\frac{4D_1[10+4c^2(1+R)^2]te^{\overline{\rho}(0,R)}}{R_0^2(\log R-\log R_0)^2}\omega_0R^ne^{-D_2R_0^2}\\
&=\frac{4\omega_0D_1R^n[10+4c^2(1+R)^2]t}{R_0^2(\log R-\log R_0)^2}\exp\left(\frac{\overline{c}}{\epsilon}+\left(\frac{4\overline{c}}{\epsilon}-D_2\right)R_0^2\right)\\
&\le\frac{4\omega_0D_1R^n[10+4c^2(1+R)^2]t}{R_0^2(\log R-\log R_0)^2}\exp\left(-R_0^2+\frac{\overline{c}}{\epsilon}\right), \qquad \text{(B.15)}
\end{aligned}
$$

where ω_0 is the volume of the unit ball in $\mathbb{R}^n$. Recall that $|\nabla\overline{\rho}|^2=\frac{4\overline{c}^2|x|^2}{(t+\epsilon)^2}$. Hence (B.8) implies

$$
\begin{aligned}
&4\int_0^t e^{-s}\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2\left|\sum_{i=1}^n(\overline{f}_i-\widetilde{f}_i)\frac{\partial\overline{\rho}}{\partial x_i}\right|^2\widetilde{u}^2\\
&\le4\int_0^t\int_{\{s\}\times\Omega}e^{\overline{\rho}}\phi^2|\overline{f}-\widetilde{f}|^2|\nabla\rho|^2\widetilde{u}^2\\
&\le4\int_0^t\int_{\{s\}\times\Omega}\widetilde{\epsilon}_1^2c^2(1+|x|)^2\frac{4\overline{c}^2|x|^2}{(s+\epsilon)^2}e^{\overline{\rho}}\widetilde{u}^2\\
&\le16\widetilde{\epsilon}_1^2c^2\widetilde{c}^2\int_0^t\frac{1}{(s+\epsilon)^2}\int_{\{s\}\times B_R}e^{\widetilde{\rho}}\widetilde{u}^2. \qquad \text{(B.16)}
\end{aligned}
$$

Similarly, we can prove that

$$4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{f}-\widetilde{f}|^2 \le 16\widetilde{\epsilon}_1^2c^2\int_0^t\int_{\{s\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2 \tag{B.17}$$

$$4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2|\overline{V}-\widetilde{V}|^2 \le 4\widetilde{\epsilon}_1^2c^2\int_0^t\int_{\{s\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2 \tag{B.18}$$

$$4\int_0^t e^{-s}\int_{\{s\}\times\Omega} e^{\overline{\rho}}\phi^2\widetilde{u}^2\left|\sum_{i=1}^n\left(\frac{\partial\overline{f}_i}{\partial x_i}-\frac{\partial\widetilde{f}_i}{\partial x_i}\right)\right|^2 \le 4\widetilde{\epsilon}_1^2c^2\int_0^t\int_{\{s\}\times B_R} e^{\overline{\rho}}\widetilde{u}^2. \tag{B.19}$$

Hence (B.12)–(B.18) and Lemma B.2 imply that

$$\begin{aligned}
&\int_{\{t\}\times B_{R_0}} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2\\
&\le \widetilde{\epsilon}_2e^t+\frac{4te^t\omega_0R^nD_1[10+4c^2(1+R)^2]}{R_0^2(\log R-\log R_0)^2}\exp\left(-R_0^2+\frac{\overline{c}}{\epsilon}\right)\\
&+16e^t\widetilde{\epsilon}_1^2c^2\widetilde{c}^2\int_0^t\frac{1}{(s+\epsilon)^2}\int_{\{s\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2+16e^t\widetilde{\epsilon}_1^2c^2\int_0^t\int_{\{s\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2\\
&+4e^t\widetilde{\epsilon}_1^2c^2\int_0^t\int_{\{s\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2+4\widetilde{\epsilon}_1^2c^2e^t\int_0^t\int_{\{s\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2\\
&\le \widetilde{\epsilon}_2e^t+\frac{4te^t\omega_0D_1R^n[10+4c^2(1+R)^2]}{R_0^2(\log R-\log R_0)^2}\exp\left(-R_0+\frac{\overline{c}}{\epsilon}\right)\\
&+16e^t\widetilde{\epsilon}_1^2c^2\widetilde{c}^2\int_0^t\frac{1}{(s+\epsilon)^2}\int_{\{0\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2\\
&+24e^t\widetilde{\epsilon}_1^2c^2\int_0^t\int_{\{0\}\times B_R} e^{\widetilde{\rho}}\widetilde{u}^2\\
&\le \widetilde{\epsilon}_2e^t+\frac{16e^t\widetilde{\epsilon}_1^2c^2\widetilde{c}^2d_1t}{\epsilon(t+\epsilon)}+24te^t\widetilde{\epsilon}_1^2c^2d_1\\
&+\frac{4te^t\omega_0D_1R^n[10+4c^2(1+R)^2]}{R_0^2(\log R-\log R_0)^2}\exp\left(-R_0^2+\frac{\overline{c}}{\epsilon}\right).
\end{aligned} \tag{B.20}$$

Let $R=2R_0$ go to infinity in (5.20), we obtain the estimate in the statement of Theorem B.4. ◻

Now we are ready to consider the global situation. For a fixed $T>0$, we want to find the solution $\overline{u}(t,x)$ of (A.5).

THEOREM B.5. *Let $\overline{u}(t,x)$ and $\widetilde{u}_i(t,x)$ be the solutions of* (A.14) *and* (A.15), *respectively. For $\widetilde{\epsilon}_1>0$, let $|\mathcal{P}_k|=\sup_i\{|t_i-t_{i-1}|\}$ be sufficiently small so that the following estimates hold:*

$$|\overline{f}(t,x)-\widetilde{f}(\tau_{i-1},x)|\le\widetilde{\epsilon}_1c(1+|x|),\ \text{for}\ \ \tau_{i-1}\le t\le\tau_i, \tag{B.21}$$

$$\left|\sum_{j=1}^n\left(\frac{\partial\overline{f}_j}{\partial x_j}(t,x)-\frac{\partial\widetilde{f}_j}{\partial x_j}(\tau_{i-1},x)\right)\right|\le\widetilde{\epsilon}_1c \tag{B.22}$$

$$\left|\overline{V}(t,x)-\widetilde{V}(\tau_{i-1},x)\right|\le\widetilde{\epsilon}_1c(1+|x|^2). \tag{B.23}$$

Then

$$\int_{\mathbb{R}^n} e^{\overline{\rho}(T,x)}(\overline{u}(T,x)-\widetilde{u}_k(T,x))^2\le\widetilde{\epsilon}_1^2c_1k|\mathcal{P}_k|e^T\le\widetilde{\epsilon}_1^2c_1c_2(T),$$

where $\overline{\rho}(t,x) = \frac{\widetilde{c}(1+|x|^2)}{t+\epsilon}$ *so that the conclusion of* Theorem B.4 *holds,* c_1 *is the constant in* Theorem B.4, *and* $c_2(T)$ *is a constant that depends only on* T.

Proof. In view of $\widetilde{u}_1(0,x) = \psi(x) = \overline{u}(0,x)$ and Theorem B.4, we have

$$\int_{\{\tau_1\}\times\mathbb{R}^n} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2 \le \widetilde{\epsilon}_1^2 \tau_1 e^{\tau_1} c_1$$

$$\int_{\{\tau_2\}\times\mathbb{R}^n} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2 \le \widetilde{\epsilon}_1^2 c_1[\tau_1 e^{\tau_2} + (\tau_2-\tau_1)e^{\tau_2-\tau_1}].$$

By Theorem B.4 and induction, we have

$$\begin{aligned}\int_{\{\tau_k\}\times\mathbb{R}^n} e^{\overline{\rho}}(\overline{u}-\widetilde{u})^2 &= \widetilde{\epsilon}_1^2 c_1[\tau_1 e^{\tau_k} + (\tau_2-\tau_1)e^{\tau_k-\tau_1} + (\tau_3-\tau_2)e^{\tau_k-\tau_2} \\ &\qquad + \cdots + (\tau_k-\tau_{k-1})e^{\tau_k-\tau_{k-1}}] \\ &\le \widetilde{\epsilon}_1^2 c_1 k|\mathcal{P}_k|e^T \\ &\le \widetilde{\epsilon}_1^2 c_1 c_2(T). \qquad \square\end{aligned}$$

As a consequence of Theorem B.5, we have the following L^2-convergent theorem.

THEOREM B.6. *Fix* $T > 0$, *let* $\mathcal{P}_k = \{0 < \tau_1 < \tau_2 < \cdots < \tau_k = T\}$ *be a partition of* $[0,T]$. *Let* $\overline{u}(t,x)$ *be the solution of* (A.14) *on* $[0,T]\times\mathbb{R}^n$. *Let* $\widetilde{u}_i(t,x)$ *be the solution of* (A.15) *on* $[\tau_{i-1},\tau_i]\times\mathbb{R}^n$. *Let* $\overline{\rho}(t,x) = \frac{\widetilde{c}(1+|x|^2)}{t+\epsilon}$ *so that the conclusion of* Theorem B.5 *holds. Then*

$$\lim_{|\mathcal{P}_k|\to 0} \int_{\{T\}\times\mathbb{R}^n} \overline{\rho}(\overline{u}-\widetilde{u}_k)^2 = 0.$$

Acknowledgments. We gratefully acknowledge the referees for their careful reading of our paper and some useful suggestions of improving our presentation.

REFERENCES

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

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

[3] R. W. Brockett and J. M. C. Clark, *The geometry of the conditional density functions,* in Analysis and Optimization of Stochastic Systems, O. L. R. Jacobs, et al., eds., Academic Press, New York, 1980, pp. 299–309.

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

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

[6] 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 Optim., 32 (1994), pp. 297–310.

[7] M. H. A. David, *On a multiplicative functional transformation arising in nonlinear filtering theory,* Z. Wahrsch Verw. Gebiete, 54 (1980), pp. 125–139.

[8] R. J. Elliott and R. Glowinski, *Approximations to solutions of the Zakai filtering equation*, Stochastic Anal. Appl., 7 (1988), pp. 145–168.

[9] 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., Lecture Notes in Inform. Sci. and Control, 144, Springer-Verlag, Berlin, pp. 228–237, Stoachastics Stochastics Rep., 35 (1991), pp. 233–256.

[10] M. Fujisaki, G. Kallianpur and H. Kunita, *Stochastic differential equations for the nonlinear filtering problem*, Osaka J. Math., 1 (1972), pp. 19–40.
[11] R. E. Kalman and R. S. Bucy, *New results in linear filtering and prediction theory,* Trans. ASME Ser. D. J. Basic Engrg., 83 (1961), pp. 95–108.
[12] 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 Antonio, TX, 1993, pp. 1564–1559.
[13] S. K. Mitter, *On the analogy between mathematical problems of nonlinear filtering and quantum physics,* Ricerche Automat., 10 (1979), pp. 163–216.
[14] 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. Control Optim., 28 (1990), pp. 173–185.
[15] S. S.-T. Yau, *Recent results on nonlinear filtering: New class of finite dimensional filters,* in Proceedings of the 29th Conf. Decision and Control, Honolulu, 1990, pp. 231–233.
[16] 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, 4(2) (1994), pp. 181–203.
[17] S. S.-T. Yau, *Brockett's problem on nonlinear filtering theory*, in Lectures on Systems, Control and Information, AMS/IP, Stud. Adv. Math., 17 (2000), pp. 177–212.
[18] 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, vol. II, San Francisco, CA, 1996, pp. 469–474.
[19] S. S.-T. Yau and G. Q. Hu, *Finite dimensional filters with nonlinear drift. X: Explicit solution of DMZ equation*, IEEE Trans. Automat. Control, 46 (2001), pp. 142–148.
[20] 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*, submitted.
[21] S. S.-T. Yau, X. Wu and W. S. Wong, *Hessian Matrix non-decomposition theorem*, Math. Res. Lett., 6 (1999), pp. 663–673.
[22] 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., 33, 1997, pp. 1277–1294.
[23] 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 Conf. Decision and Control, Lake Buena Vista, FL, 1994, pp. 1221–1225.
[24] S. S.-T. Yau and S.-T. Yau, *Explicit solution to a Kolmogorov equation,* Appl. Math. Optim., 34 (1996), pp. 231–266.
[25] S.-T. Yau and S. S.-T. Yau, *Real time solution of nonlinear filtering problem without memory, I*, Math. Res. Lett., 7 (2000), pp. 671–693.
[26] 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 J. Math., 2 (1998), pp. 1079–1149.