NEW ALGORITHMS IN REAL TIME SOLUTION OF THE NONLINEAR FILTERING PROBLEM*

STEPHEN S. T. YAU†

Abstract. It is well known that the filtering theory has important applications in both military and commercial industries. The Kalman–Bucy filter has been used in many areas such as navigational and guidance systems, radar tracking, solar mapping, and satellite orbit determination. However, the Kalman–Bucy filter has limited applicability because of the linearity assumptions of the drift term and observation term as well as the Gaussian assumption of the initial value. Therefore there has been an intensive interest in solving the nonlinear filtering problem. The central problem of nonlinear filtering theory is to solve the DMZ equation in real time and memoryless way. In this paper, we shall describe three methods to solve the DMZ equation: Brockett-Mitter estimation algebra method, direct method, and new algorithm method. The first two methods are relatively easy to implement in hardware and can solve a large class of nonlinear filtering problems. We shall present the recent advance in the third method which solves all the nonlinear filtering problems in a real-time manner in theory.

1. Introduction. It is well known that the filtering theory has important applications in both military and commercial industries. The Kalman–Bucy filter has been used in many areas such as navigational and guidance systems, radar tracking, solar mapping, and satellite orbit determination. However, the Kalman–Bucy filter has limited applicability because of the linearity assumptions of the drift term and observation term as well as the Gaussian assumption of the initial value. Therefore there has been an intensive interest in solving the nonlinear filtering problem. The nonlinear filtering problem involves the estimation of a stochastic process $x = \{x_t\}$ (called the signal or state process) that cannot be observed directly. Information containing x is obtained from observations of a related process $y = \{y_t\}$ (the observation process). The goal of nonlinear filtering is to determine the conditional density $\rho(t, x)$ of x_t given the observation history of $\{y_s \colon 0 \leqslant s \leqslant t\}$. In the late 1960s, Duncan [Du], Mortensen [Mo] and Zakai [Za] independently derived the Duncan–Mortensen–Zakai (DMZ) equation for the nonlinear filtering theory which the conditional probability density $\rho(t, x)$ must satisfy. The central problem of nonlinear filtering theory is to solve the DMZ equation in real time and memoryless way. In this paper, we shall describe three methods to solve the DMZ equation: Brockett–Mitter estimation algebra method, direct method, and new algorithm method. The first two methods are relatively easy to implement in hardware and can solve a large class of nonlinear filtering problems. We shall present the recent advance in the third method which solves all the nonlinear filtering problems in a real-time manner theoretically.

The first method began in the late seventies. Brockett and Clark [Br-Cl], Brockett [Br], and Mitter [Mi] proposed the idea of using estimation algebras to construct finite-dimensional nonlinear filters. There is an excellent survey paper by Marcus [Ma] of earlier results before 1984. In 1994, Chiou and Yau [Chi-Ya] introduced the concept of maximal rank estimation algebra. In a series of papers [Ya1], [Ch-Ya1], [Ch-Ya2], [H-Y-C], [W-Y-H], [Y-W-W], and [Ya-Hu3], Yau and his co-workers have classified all finite-dimensional estimation algebras with maximal rank. The self-contained proof of the classification of finite-dimensional estimation algebras of maximal rank can be found in [Ya3], [Ya4]. Ten years ago, Wong and Yau [Wo-Ya] wrote a survey paper of results before 1998, which was dedicated to Brockett on the occasion of his 60th birthday. In this survey paper, Wong and Yau had a detailed discussion on the importance of the papers [T-W-Y] and [D-T-W-Y] in understanding the structure of exact estimation algebra. Therefore we shall not discuss the classification of maximal rank finite dimensional estimation algebras, which was done before 1998, in this paper.

Despite the success of the classification of finite-dimensional estimation algebras with maximal rank, the problem of classification of non-maximal rank finite-dimensional estimation algebras is still wide open except for the case of state space dimension 2 which was finished recently by Wu and Yau [Wu-Ya] and some construction of non-maximal rank finite-dimensional estimation algebras by Rasoulian and Yau [Ra-Ya]. There are several reasons to tackle the problem of classification of non-maximal rank finite-dimensional estimation algebras. The first one is that the theory of estimation algebra provides a systematic tool to deal with questions concerning finite-dimensional filters. It has led to a number of new results concerning finite-dimensional filters and to a deeper understanding of the structure of nonlinear filtering in general. It explains convincingly in [Bu-Pa] and [H-M-S] why it is easy to find exact recursive filters for linear dynamical models, while it is very hard to handle the cubic sensor problem. The second reason is that the solution of this problem may give us many new classes of finite-dimensional estimation algebras and hence many new classes of finite-dimensional filters. More importantly, the finite-dimensionality of the estimation algebra guarantees the explicit construction of the finite-dimensional recursive filter, and the filter so constructed is universal in the sense of [Ch-Mi]. Due to the difficulty of the problem, Brockett suggested that one should understand the low-dimensional estimation algebras first. In [Ya-Ra] Yau and Rasoulian have classified estimation algebras of dimension at most four. In [C-C-Y], Chiou, Chiueh and Yau gave a structure theorem for estimation algebras of dimension 5.

The second approach to solve the nonlinear filtering problem is the direct method introduced by Yau and Yau [Ya-Ya1], [Ya-Ya2] and generalized by Hu and Yau [Hu-Ya], [Ya-Hu1], [Ya-Hu2]. Recall that in the Wei–Norman approach of using the estimation algebras to solve the DMZ equation one has to know explicitly the basis of a vector space of the estimation algebra in order to reduce the DMZ equation to

a finite system of ordinary differential equation, Kolmogorov equation, and several first-order linear partial differential equations. Classically one knows the explicit basis for the estimation algebra only in the case that it has maximal rank. The new direct method offers several advantages. It is easy and the derivation no longer needs the maximal rank condition of the estimation algebras. Thus, the algorithm is universal for any Yau filtering model [Ch] which includes the linear filtering model and the Benés filtering model as special cases. Furthermore, it eliminates the necessity of integrating n first-order linear partial differential equations, as was the case in the estimation algebra method. Finally, the number of sufficient statistics required to compute the conditional probability density of the state in this direct method is n.

The main purpose of this paper is to describe the third approach which enables us to solve the nonlinear filtering problem completely. The filtering problem considered here is based on the following continuous signal observation model described by the vector Ito stochastic differential equation

$$(1.1) \qquad \begin{cases} dx_t = f(x_t, t)\, dt + G(x_t, t)\, dv_t, \\ dy_t = h(\lambda_t, t)\, dt + dw_t \end{cases}$$

where x_t and f are n-vectors, G is an $n \times r$ matrix, and v_t is an r-vector Brownian motion process with $E[dv_t dv_t^T] = Q(t)dt$, y_t and h are m-vectors and w_t is an m-vector Brownian motion process with $E[dw_t dw_t^T] = R(t)dt$ and $R(t) > 0$. We assume that $\{v_t, t \geqslant 0\}$, $\{w_t, t \geqslant 0\}$, and x_0 are independent. Recently Yau and Yau [Ya-Ya3], [Ya-Ya4] studied the above model with the following assumptions: (1) f, G and h are functions of the state x_t only; (2) G is an orthogonal matrix, i.e., $GG^T = I$; (3) $\{v_t, t \geqslant 0\}$ and $\{w_t, t \geqslant 0\}$ are standard Brownian motion processes, i.e., $R(t) = Q(t) = I$. In 2000, Yau and Yau [Ya-Ya3] proposed a novel algorithm to solve the DMZ equation in real time and memoryless way under the assumptions that the drift term of the state equation, and its first and second derivatives, and the drift term of the observation equation and its first derivatives, have linear growth. They showed that the solution obtained from their algorithm converges to the true solution of the DMZ equation. Although this approach is very successful, so far it cannot handle the famous cubic sensor problem in engineering, which is well known that there is no finite dimensional filter.

The breakthrough in the subject is the recent paper [Ya-Ya4] which appeared in SIAM Control and Optimization. Yau and Yau showed that under very mild conditions (which essentially say that the growth of the observation $|h|$ is greater than the growth of the drift $|f|$), the DMZ equation admits a unique nonnegative weak solution u which can be approximated by a solution u_R of the DMZ equation on the ball B_R with $u_R|_{\partial B_R} = 0$. The error of this approximation is bounded by a function of R which tends to zero as R goes to infinity. The solution u_R can in turn be approximated efficiently by an algorithm depending only on solving the

observation-independent Kolmogorov equation on B_R. In theory, their algorithm can solve practically all engineering problems in real time. Specifically, they show that the solution obtained from their algorithms converges to the solution of the DMZ equation in the L^1-sense. Equally important, they have a precise error estimate of this convergence, which is important in numerical computation. Most recently, Yan and Yau [Yan-Ya] generalize the Yau–Yau algorithm so that it can solve the general filtering model without the above assumptions (1), (2) and (3). Basically for the general filtering model, we can again solve the robust DMZ equation by means of the Kolmogorov equation which is independent of observation

In section 2, we shall recall some basic concepts and results in nonlinear filtering theory. In section 3, we shall describe the recent results on finite dimensional non-maximal rank estimation algebras. In section 4, we shall discuss the direct method and explain how it solves the DMZ equation arising from the Yau filtering model. In section 5, we shall describe the last approach which solves theoretically all the engineering problems in nonlinear filtering.

We are very grateful to Professor Brockett for his continuous support on this research project. It is our great pleasure to dedicate this paper to him on the occasion of his 70th birthday. We wish him a healthy and active life for another seventy years!

2. Some Basic Concepts and Results. The filtering problem considered in §3, §4 and part of §5 is based on the signal observation model

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

Here x, v, y and w are respectively $\mathbb{R}^n$, $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{R}^m$ valued processes, and v and w have components which are independent, standard Brownian processes. We assume that f, h are C^∞ smooth, and g is an orthogonal matrix. We refer to $x(t)$ as the state of the system at time t and to $y(t)$ as the observation at time t.

Let $\rho(t, x)$ denote the conditional probability density of the state given the observation $\{y(t) \colon 0 \leqslant s \leqslant t\}$. It is well known (see, e.g., [Da-Ma]) that $\rho(t, x)$ is given by normalizing $\sigma(t, x)$, which satisfies the DMZ equation:

(2.2)
$$\left.\begin{aligned} d\sigma(t, x) = L_0 \sigma(t, x)\, dx + \sum_{i=1}^{m} L_i \sigma(t, x)\, dy_i(t) \\ \sigma(0, x) = \sigma_0 \end{aligned}\right\}$$

where

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

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

Equation (2.2) is a stochastic partial differential equation (with probability space a space of paths $\{y\}$) and as such a solution is in principle only defined apart from a set of measure zero. On the other hand, actual observations will always consist of piecewise smooth sample paths $y(t)$ and the class of all such paths is of measure zero. Thus there arises the question whether there exists a version of (2.2) which can be interpreted pathwise for all $y(t)$ and for which the solution of (2.2) for piecewise smooth $y(t)$ carry (approximate) information. This means that in real applications, we are interested in constructing robust state estimators from observed sample paths with some property of robustness. Davis [Da] studied this problem and proposed some robust algorithms. In our case, his basic idea reduces to defining a new unnormalized density

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

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

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

$$u(0, x) = \sigma_0(x).$$

Here we have used the following notation.

DEFINITION 2.1. *If X and Y are differential operators, the Lie bracket of X and Y, $[X, Y]$ is defined by $[X, Y]\phi = X(Y\phi) - Y(X\phi)$ for any C^∞ function ϕ.*

Recall that a real vector space $\mathcal{F}$, with an operation $\mathcal{F} \times \mathcal{F} \to \mathcal{F}$ denoted $(x, y) \to [x, y]$ (called the Lie bracket of x and y) is called a Lie algebra if the following axioms are satisfied:

(i) The Lie bracket operation is bilinear

(ii) $[x, x] = 0$ for all $x \in \mathcal{F}$

(iii) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \quad (x, y, z \in \mathcal{F}).$

DEFINITION 2.2. *The estimation algebra E of a filtering model (2.1) is defined as the Lie algebra generated by $\{L_0, L_i, \ldots, L_m\}$ denoted by $\{L_0, L_1, \ldots, L_m\}_{L.A.}$. E is said to be an estimation algebra of maximal rank if, for any $1 \leqslant i \leqslant n$, there exists a constant c_i such that $x_i + c_i$ is in E.*

EXAMPLE 2.1. *Consider the filtering model (2.1) with $n = m = 2$ and $f \equiv 0$.*

CASE 1. *$h_1 = x_1$, $h_2 = x_2$. In this case the estimation algebraa E generated by $\{L_0, x_1, x_2\}$ is an estimation algebra of maximal rank. In fact, E is of dimension 6. It consists of elements 1, x_1, x_2, D_1, D_2, and L_0.*

CASE 2. *$h_1 = x_1$, $h_2 = 0$. In this case the estimatin algebra E generated by*

$\{L_0, x_1\}$ *is not of maximal rank. In fact, E is of dimension* 4. *It consists of elements* 1, x_1, D_1 *and* L_0.

DEFINITION 2.3. *Wong's matrix of a filtering model* (2.1) *is a* $n \times n$ *matrix* $\Omega = (\omega_{ij})$ *defined by*

$$(2.5) \qquad \omega_{ij} = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \qquad 1 \leqslant i, j \leqslant n$$

We remark that clearly Ω is a skew symmetric matrix with the cyclic conditions

$$\frac{\partial \omega_{jk}}{\partial x_i} + \frac{\partial \omega_{ki}}{\partial x_j} + \frac{\partial \omega_{ij}}{\partial x_k} = 0, \qquad \forall\, 1 \leqslant i, j, k \leqslant n.$$

Define

$$(2.6) \qquad D_i = \frac{\partial}{\partial x_i} - f_i \quad \text{and} \quad \eta = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{m} h_i^2.$$

Then

$$(2.7) \qquad L_0 = \frac{1}{2} \Big(\sum_{i=1}^{n} D_i^2 - \eta \Big)$$

The following theorem due to Ocone [Oc] is the first result which allows us to understand what kind of functions can appear in a finite dimensional estimation algebra.

THEOREM 2.1 (Ocone [Oc]). *Let E be a finite-dimensional estimation algebra. If a function ξ is in E, then ξ is a polynomial of degree at most two.*

It follows by (2.6) that the underdetermined partial differential equation

$$(2.8) \qquad \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F$$

provides a characterization of the realization of such nonlinear filtering models. Therefore, it is of primary interest to investigate the solution and solution properties for this class of equations. In (2.8), $f_1, \ldots, f_n$ and F are C^∞ functions on $\mathbb{R}^n$. F is given and $f_i, \ldots, f_n$ are treated as unknown. Although there is only one equation with n unknowns, (2.8) may not have solutions. The following important result can be found in [Ya2].

THEOREM 2.2 (See [Ya2]). *Let $F(x_1, \ldots, x_m)$ be a C^∞ function on $\mathbb{R}^n$. Suppose that there exists a path $c \colon \mathbb{R} \to \mathbb{R}^n$ and $\delta > 0$ such that $\lim_{t \to \infty} \|c(t)\| = \infty$ and $\lim_{t \to \infty} \sup_{B_\delta(c(t))} F = -\infty$, where $B_\delta(c(t)) = \{x \in \mathbb{R}^n \colon \|x - c(t)\| < \delta\}$. Then there are no C^∞ functions $f_1, f_2, \ldots, f_n$ on $\mathbb{R}^n$ satisfying* (2.8).

The following results are applications of Theorem 2.2 that will be used in the classification of estimation algebras.

THEOREM 2.3 (See [Wu-Ya]). *Let $F(x_1, \ldots, x_n)$ be a polynomial on $\mathbb{R}^n$, let $c(t) = (c_1(t), \ldots, c_n(t))$ be a polynomial path (i.e., $c_i(t)$'s are polynomials in t), and let $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$. Then $F(c(t)+\varepsilon)$ is a polynomial in t, i.e., $F(c(t)+\varepsilon) = \sum_{j=0}^{d} a_j(\varepsilon)t^j$, where $a_j(\varepsilon)$'s are polynomials of $\varepsilon_1, \ldots, \varepsilon_n$. If there exists a path $c(t)$ such that the coefficient $a_d(t)$ of the leading term of $F(c(t) + \varepsilon)$ is a negative constant and $d \geqslant 1$, there are no C^∞ functions $f_1, f_2, \ldots, f_n$ on $\mathbb{R}^n$ satisfying the equation* (2.8).

THEOREM 2.4 (See [Wu-Ya]). *Let $F(x_1, \ldots, x_n)$ be a degree $d \geqslant 1$ polynomial on $\mathbb{R}^n$. The homogeneous degree d part of F is denoted by $F_d = \sum_{|i|=d} a_i x_1^{i_1} \ldots x_n^{i_n}$, where $i = (i_1, \ldots, i_n)$. If there exist n numbers $b_1, \ldots, b_n$ such that $F_d(b_1, \ldots, b_n) < 0$, then there are no C^∞ functions $f_1, f_2, \ldots, f_n$ on $\mathbb{R}^n$ satisfying* (2.8).

COROLLARY 2.1. *Let $F(x_1, \ldots, x_n)$ be a polynomial on $\mathbb{R}^n$. If the degree of F is odd, there are no C^∞ functions $f_1, f_2, \ldots, f_n$ in $\mathbb{R}^n$ satisfying* (2.8).

COROLLARY 2.2. *Let $F(x_1, \ldots, x_n)$ be a degree $d \geqslant 1$ polynomial on $\mathbb{R}^n$. If there exist C^∞-functions $f_1, f_2, \ldots, f_n$ on $\mathbb{R}^n$ satisfying* (2.8), *then the homogeneous degree d part of F is a nonnegative function.*

In fact, the above theorems can be further generalized to the case when F is a polynomial in $x_1, \ldots, x_r$ variables with C^∞ coefficient functions $x_{r+1}, \ldots, x_n$. For example, we have the following theorem.

THEOREM 2.5 (See [Wu-Ya]). *Let d and $r \leqslant n$ be two positive integers and*

$$F(x_1, \ldots, x_n) = \sum_{|i| \leqslant d} a_i(x_{r+1}, \ldots, x_n) x_1^{i_1}, \ldots x_r^{i_r},$$

where $i = (i_1, \ldots, i_r)$, and where a_i's are C^∞ functions in $x_{r+1}, \ldots, x_n$ variables. The homogeneous degree d part in $x_1, \ldots, x_r$ variables of F is denoted by $F_d = \sum_{|i|=d} a_i(x_{r+1}, \ldots, x_n) x_1^{i_1} \ldots x_r^{i_r}$. If there exist n numbers $b_1, \ldots, b_n$ such that $F_d(b_1, \ldots, b_n) < 0$, then there are no C^∞ functions $f_1, f_2, \ldots, f_n$ on $\mathbb{R}^n$ satisfying (2.8).

3. Finite-Dimensional Estimation Algebras of Non-Maximal Rank.

In this section, we shall study finite-dimensional estimation algebras of non-maximal rank. We shall first give general construction of finite-dimensional estimation algebras of non-maximal rank. This construction gives rise to a new class of finite-dimensional nonlinear filters which are not discussed previously. Presumably most of the finite-dimensional estimation algebras of non-maximal rank are arising in this way.

3.1. Construction of a finite-dimensional estimation algebra of non-maximal rank.

Suppose that E is the finite-dimensional estimation algebra associated with the filtering model (2.1). Consider the following deterministically decoupled filtering system:

$$(3.1) \qquad \left. \begin{array}{l} d\tilde{x}(t) = \tilde{f}(\tilde{x}(t))\, dt + \tilde{g}(\tilde{x}(t))\, d\tilde{v}(t), \quad \tilde{x}(0) = \tilde{x}_0 \\ dy(t) = h(\tilde{x}(t))\, dt + d\tilde{w}(t), \quad y(0) = 0 \end{array} \right\}$$

Here $\tilde{x} = (x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+k})$, $\tilde{f}(\tilde{x}(t)) = (f_1(x_1, \ldots, x_n), \ldots, f_n(x_1, \ldots, x_n),$
$f_{n+1}(x_{n+1}, \ldots, x_{n+k}), \ldots, f_{n+k}(x_{n+1}, \ldots, x_{n+k}))$, $\tilde{g}(\tilde{x}(t)) = $ orthogonal matrix,
$h(\tilde{x}(t)) = h(x_1, \ldots, x_n)$, and $\tilde{v}$ and $\tilde{w}$ have components which are independent, standard Brownian processes.

Let $\tilde{E}$ be the estimation algebra associated with (3.1). We shall show that $\tilde{E}$ is isomorphic to E as a Lie algebra. Observe that

$$\tilde{\omega}_{ij} = \frac{\partial \tilde{f}_j}{\partial x_i} - \frac{\partial \tilde{f}_i}{\partial x_j} = \begin{cases} \dfrac{\partial f_j}{\partial x_i}(x) - \dfrac{\partial f_i}{\partial x_j}(x) = \omega_{ij} & \text{if } 1 \leqslant i, j \leqslant n \\[2mm] 0 \quad \text{if } i \geqslant n, j \leqslant n \text{ or } i \leqslant n, j \geqslant n, \\[2mm] \dfrac{\partial f_j}{\partial x_i}(x_{n+1}, \ldots, x_{n+k}) \\[2mm] \qquad\qquad - \dfrac{\partial f_i}{\partial x_j}(x_{n+1}, \ldots, x_{n+k}), \ i, j \geqslant n, \end{cases}$$

$$\tilde{L}_0 = \frac{1}{2}\Big(\sum_{i=1}^{n+k} D_i{}^2 - \tilde{\eta} \Big),$$

where

$$D_{n+i} = \frac{\partial}{\partial x_{n+i}} - f_{n+i}(x_{n+1}, \ldots, x_{n+k}), \quad 1 \leqslant i \leqslant k,$$

$$\tilde{\eta}(\tilde{x}) = \eta(x) + \sum_{i=n+1}^{n+k} \frac{\partial f_j}{\partial x_i}(x_{n+1}, \ldots, x_{n+k}) + \sum_{i=n+1}^{n+k} f_i{}^2(x_{n+1}, \ldots, x_{n+k}).$$

LEMMA 3.1. *If F is a differential operator in $x_1, \ldots, x_n$ variables of order r, then $[F, \tilde{\eta}(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+k})] = [F, \eta(x_1, \ldots, x_n)]$ is a differential operator of order $r-1$ in $x_1, \ldots, x_n$ variables.*

LEMMA 3.2. *Let F be a differential operator of order r in $x_1, \ldots, x_n$ variables. Then for $n+1 \leqslant j \leqslant n+k$, $[D_j{}^2, F] = 0$.*

LEMMA 3.3. *For any differential operator F in $x_1, \ldots, x_n$ variables, $[\tilde{L}_0, F] = [L_0, F]$.*

THEOREM 3.1 (Rasoulian–Yau [Ra-Ya]). *The estimation algebra $\tilde{E}$ associated with the filtering model (3.1) is isomorphic to the estimation algebra E associated with the filtering model (2.1), and $\tilde{E}$ consists of a basis such that all elements in this basis are differential operators in $x_1, \ldots, x_n$ variables except $\tilde{L}_0$. Furthermore, $\tilde{\Omega} = \Big(\frac{\partial \tilde{f}_j}{\partial x_i} - \frac{\partial \tilde{f}_i}{\partial x_j} \Big)$ is given by $\tilde{\Omega} = \big(\begin{smallmatrix} \Omega & 0 \\ 0 & A \end{smallmatrix} \big)$, where Ω is the $n \times n$ matrix $\Big(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} \Big)$ associated with (2.1) and A is a $k \times k$ matrix with (i, j)-entry $\Big(\frac{\partial f_{n+j}}{\partial x_{n+i}} - \frac{\partial f_{n+i}}{\partial x_{n+j}} \Big)(x_{n+1}, \ldots, x_{n+k})$.*

REMARK 3.1. (1) *The finite-dimensional estimation algebra $\tilde{E}$ is of non-maximal rank if $k > 0$. Observe that A is quite arbitrary.* (2) *We would like to emphasize that the orthogonal matrix $\tilde{g}(x_1, \ldots, x_n, x_{n+1}, \ldots, x_{n+k})$ is arbitrary and is not necessarily*

of the form

$$\begin{pmatrix} g(x_1, \ldots, x_n) & 0 \\ 0 & g(x_{n+1}, \ldots, x_{n+k}) \end{pmatrix}.$$

So (3.1) *is not a direct sum of two filtering systems. This is a central point of Theorem 3.1.*

3.2. Classification of finite-dimensional estimation algebras of dimension less than or equal to four.

Let $E = \{L_0, h_1, h_2, \ldots, h_m\}_{L.A.}$ be the estimation algebra associated with the filtering model (2.1). By Ocone's theorem, $h_1, h_2, \ldots, h_m$ are polynomials at most degree two. Let $h_1, h_2, \ldots, h_k$, $k \leqslant m$ be linearly independent and the other $h_i{}'$s, $k+1 \leqslant i \leqslant m$ be linear combinations of $h_1, \ldots, h_k$. If $k = 1$ and $h_1 = $ constant, it is obvious that E is 1-dimensional if and only if $h_1 = 0$, and E is 2-dimensional if and only if $h_1 = 1$. Thus we have the following proposition.

PROPOSITION 3.1. *Let E be the estimation algebra associated with the filtering model* (2.1). *Then*

(1) $\dim E = 1$ *if and only if* $E = \langle L_0 \rangle$

(2) $\dim E = 2$ *if and only if* $E = \langle L_0, 1 \rangle$.

THEOREM 3.2 (Yau–Rasoulian [Ya-Ra]). *Let E be the estimation algebra associated with the filtering model* (2.1). *Then* $\dim E \neq 3$.

The following theorem gives a structure of 4-dimensional estimation algebras.

THEOREM 3.3 ([Ya-Ra]). *Suppose that state space of the filtering model* (2.1) *is of a dimension greater than one. Then the observation function $h(x)$ is linear and the linear span of $\nabla h_1, \ldots, \nabla h_m$ is 1-dimensional. Assume $h_1(x) = x_1$. Then the 4-dimensional estimation algebra has a basis given by 1, x_1, $D_1 = \frac{\partial}{\partial x_1} - f_1$ and $L_0 = \frac{1}{2}\left(\sum_{i=1}^n D_i^2 - \eta\right)$. Moreover, $\omega_{12} = \omega_{13} = \cdots = \omega_{1n} = 0$, $[L_0, x_1] = D_1$, $[D_1, x_1] = 1$, $[L_0, D_1] = \frac{1}{2}\frac{\partial \eta}{\partial x_1} = \alpha x_1 + \beta$. where α, β are constants. Also, $\eta = \alpha x_1{}^2 + 2\beta x_1 + g(x_2, \ldots, x_n)$, where $g(x_2, \ldots, x_n)$ is in $C^\infty(\mathbb{R}^{n-1})$. In particular, $f_1, \ldots, f_n$ have to satisfy the equation*

$$(3.2) \qquad \sum_{i=1}^n \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^n f_i{}^2 = (\alpha - 1)x_1^2 + 2\beta x_1 + g(x_2, \ldots, x_n)$$

where $\alpha \geqslant 1$.

3.3. Structure theorem for five-dimensional estimation algebras.

The following lemmas plays an important role in understanding estimation algebra of non-maximal rank.

LEMMA 3.4 (Chiou–Chiueh–Yau [C-C-Y]). *For any $1 \leqslant l \leqslant n$, if γ_i, $i = 1, \ldots, l$, are polynomials in $x_1, \ldots, x_l$ with coefficients in C^∞ functions of $x_{l+1}, \ldots, x_n$ satis-*

fying

$$
\text{(3.3)} \qquad \frac{\partial \gamma_j}{\partial x_i} + \frac{\partial \gamma_i}{\partial x_j} = 0 \ \text{for all } 1 \leqslant i, j \leqslant l,
$$

then each γ_i is necessarily of the form

$$
\text{(3.4)} \qquad \gamma_i = \sum_{1 \leqslant j \leqslant l} c_i^j(x_{l+1}, \ldots, x_n) x_j + d_i(x_{l+1}, \ldots, x_n)
$$

where $c_i^j(x_{l+1}, \ldots, x_n)$ and $d_i(x_{l+1}, \ldots, x_n)$ are C^∞ functions and $c_i^j = -c_j^i$.

LEMMA 3.5. *If $\dim E = 5$, then E cannot contain two linear independent degree one polynomials.*

THEOREM 3.4 (Chiou–Chiueh–Yau [C-C-Y]). *If $\dim E = 5$, then E cannot contain any degree two polynomial.*

Theorem 3.4, above, together with Theorem 3.3 asserts the following Mitter conjecture is true for estimation algebras with dimension at most five.

MITTER CONJECTURE 1. *Let E be a finite-dimensional estimation algebra. Then any function in E is a polynomial of degree at most one.*

Now we are ready to state the structure theorem of five-dimensional estimation algebras.

THEOREM 3.5 (Chiou–Chiueh–Yau [C-C-Y]). *Suppose that the state space of the filtering model (2.1) is of dimension at least two. Then the five-dimensional estimation algebra is isomorphic to to Lie algebra generated by L_0 and an observation function $h = x_1$ with a basis given by 1, x_1, $D_1 = \frac{\partial}{\partial x_1} - f_1(x_1, \ldots, x_n)$, $Y_1 = [L_0, D_1] = \sum_{i=1}^{n} \omega_{i1} D_i + \frac{1}{2} \frac{\partial \eta}{\partial x_1}$, $L_0 = \frac{1}{2}\left(\sum_{i=1}^{n} D_i^2 - \eta\right)$. Moreover, the following holds:*

(1) *$\omega_{1i} \neq 0$ for some $i = 2, \ldots, n$ and each ω_{1i} is of the form*

$$
\text{(3.5)} \qquad \omega_{1i} = \sum_{k=2}^{n} e_{ik} x_k + e_i \quad \text{for } 2 \leqslant i \leqslant n,
$$

$$
\text{(3.6)} \qquad e_{ij} = -e_{ji}, \quad 2 \leqslant i, j \leqslant n,
$$

   where e_{ij} and e_i are constants.

(2) *η is of the form*

$$
\text{(3.7)} \qquad \eta = \left(\sum_{j=2}^{n} \omega_{ij}^2 + c_3\right) x_1^{\,2} + \beta(x_2, \ldots, x_n) x_1 + \gamma(x_2, \ldots, x_n),
$$

   where $c_3 \geqslant 1$ is a constant and $\beta(x_2, \ldots, x_n)$ and $\gamma(x_2, \ldots, x_n)$ are C^∞ functions

(3) *There exists a constant c_1 such that*

$$
\text{(3.8)} \qquad \sum_{j=1}^{n} \omega_{1j} \omega_{ji} + \frac{1}{2} \frac{\partial^2 y}{\partial x_i \partial x_1} = c_1 \omega_{1i}, \ 2 \leqslant i \leqslant n
$$

(4) *There exists constants c_0 and c_2 such that*

$$(3.9) \qquad -\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial\omega_{1j}}{\partial x_i}\omega_{ji} + \frac{1}{2}\sum_{j=1}^{n}\omega_{1j}\frac{\partial\eta}{\partial x_j} = c_0 x_1 + \frac{c_1}{2}\frac{\partial\eta}{\partial x_1} + c_2.$$

In particular, $f_1, \ldots, f_n$ have to satisfy the following equation:

$$(3.10) \qquad \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n}f_i{}^2 = \Big(\sum_{j=2}^{n}\omega_{1j}^2 + c_3 - 1\Big)x_1{}^2$$

$$+ \beta(x_2, \ldots, x_n)x_1 + \gamma(x_2, \ldots, x_n).$$

Moreover, the five-dimensional estimation algebra has the following multiplication table.

E	1	x_1	D_1	Y_1	L_0
1	0	0	0	0	0
x_1	0	0	-1	0	$-D_1$
D_1	0	1	0	c_3	$-Y_1$
Y_1	0	0	$-c_3$	0	$-c_0 x_1 - c_1 Y_1 - c_2 - c_3 D_1$
L_0	0	D_1	Y_1	$c_0 x_1 + c_1 Y_1 + c_2 + c_3 D_1$	0

COROLLARY 3.1. *Suppose that the state space of the filtering model* (2.1) *is of dimension at least two. If ω_{1j}, $2 \leqslant j \leqslant n$, are constants, then the five-dimensional estimation algebra is isomorphic to a Lie algebra generated by L_0 and x_1 with a basis given by 1, x_1, $D_1 = \frac{\partial}{\partial x_1} - f_1(x_1, \ldots x_n)$, $Y_1 = [L_0, D_1] = \sum_{i=1}^{n}\omega_{i1}D_i + \frac{1}{2}\frac{\partial\eta}{\partial x_1}$, $L_0 = \frac{1}{2}\big(\sum_{i=1}^{n}D_i^2 - \eta\big)$. Moreover, the following holds:*

(1) *η is of the form*

$$(3.11) \qquad \eta = ax_1{}^2 + \beta(x_2, \ldots, x_n)x_1 + \gamma(x_2, \ldots, x_n)$$

  for some constant $a > 1$ and C^∞ functions $\beta(x_2, \ldots, x_n)$ and $\gamma(x_2, \ldots, x_n)$.

(2) *There exists constant d_1 such that*

$$(3.12) \qquad \frac{1}{2}\frac{\partial\beta}{\partial x_i}(x_2, \ldots, x_n) = d_1\omega_{1i} + \sum_{j=1}^{n}\omega_{1j}\omega_{ij} \text{ for } 2 \leqslant i \leqslant n.$$

(3) *There exist constants d_2 and d_3 such that*

$$(3.13) \qquad \sum_{j=1}^{n}\omega_{1j}\frac{\partial\beta}{\partial x_j}(x_2, \ldots, x_n) = d_2$$

$$(3.14) \qquad \sum_{j=1}^{n}\omega_{1j}\frac{\partial\gamma}{\partial x_j}(x_2, \ldots, x_n) = d_1\beta(x_2, \ldots, x_n) + d_3.$$

We shall give an example of five-dimensional estimation algebras by using Theorem 3.5. Example 3.1 below is a new class of finite-dimensional estimation algebras.

EXAMPLE 3.1. *Consider the filtering model*

$$\begin{cases} dx(t) = f(x(t))\,dt + g(x(t))\,dv(t) \\ dy(t) = h(x(t))\,dt + dw(t) \end{cases}$$

where $f_1 = ax_1$, $f_2 = bx_1x_3$, $f_3 = -bx_1x_2$, $f_i = g_i(x_4, \ldots, x_n)$, $4 \leqslant i \leqslant n$, $h(x) = x_1$, *and* a, b *are nonzero constants. Then*

$$\omega_{12} = \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} = bx_3, \quad \omega_{13} = \frac{\partial f_3}{\partial x_1} - \frac{\partial f_1}{\partial x_3} = -bx_2$$

$$\omega_{23} = \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} = -2bx_1, \quad \omega_{1j} = \frac{\partial f_j}{\partial x_1} - \frac{\partial f_1}{\partial x_j} = 0, \quad 4 \leqslant i \leqslant n.$$

$$\sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} = a^2 x_1^2 + b^2 x_1^2 x_3^2 + b^2 x_1^2 x_2^2$$

$$+ \sum_{i=4}^{n} g_i^2(x_4, \ldots, x_n) + a + \sum_{i=4}^{n} \frac{\partial g_i}{\partial x_i}(x_4, \ldots, x_n)$$

$$= \Big(\sum_{i=1}^{n} \omega_{1i}^2 + a^2 \Big) x_1^2 + \sum_{i=4}^{n} g_i^2(x_4, \ldots, x_n) + a$$

$$+ \sum_{i=4}^{n} \frac{\partial g_i}{\partial x_i}(x_4, \ldots, x_n).$$

It is easy to check that ω_{1i}, $1 \leqslant i \leqslant n$, *satisfy* (3.5), η *is of the form* (3.7) *with* $c_3 = a^2 + 1$ *satisfying* (3.8) *and* (3.9) *with* $c_0 = 2b^2$, $c_1 = c_2 = 0$. *The estimation algebra* E *is five-dimensional with basis* $\{1, x_1, D_1, Y_1 = bx_3D_2 - bx_2D_3 + (a^2 + 1 + b^2x_2^2 + b^2x_3^2)x_1, L_0\}$.

3.4. Classification of estimation algebras with state dimension 2.

In this subsection, we shall fix the state dimension 2 and consider all finite-dimensional estimation algebras.

THEOREM 3.6 (Wu–Yau [Wu-Ya]). *Suppose that the state dimension of the filterings model* (2.1) *is two. Assume* $\dim E < \infty$ *and* $Y = p(x)D_2 + function$ *is an element in* E. *Then* p *is a polynomial in* x_1, x_2, *of degree at most one.*

THEOREM 3.7 (Wu–Yau [Wu-Ya]). *Suppose that the state space dimension of the filtering model* (2.1) *is two. If* $\dim E < \infty$ *and* $\phi \in E$, *then* ϕ *is a polynomial of degree at most one.*

THEOREM 3.8 (Wu–Yau [Wu-Ya]). *Suppose that the state dimension of the filtering model* (2.1) *is two. If* E *has linear rank* 1 *and is finite-dimensional, then the Wong matrix* Ω *has constant entries*

We are now read to classify all finite-dimensional estimation algebras with state dimension 2. First assume E has linear rank 1. By Theorem 3.7, h_i's must be degree at most 1 polynomial in x_1. By Theorem 3.8, ω_{12} is a constant.

(i) $\omega_{12} = 0$, $[L_0, D_1] = \frac{1}{2}\frac{\partial \eta}{\partial x_1} \in E$. Thus, η must be a degree 2 polynomial in x_1 plus a C^∞-function in x_2. $E = \{L_o, x_1, D_1, 1\}$. For example, $f_1 = x_1$, $f_2 = \sin x_2$, $h_1 = x_1$, $\eta = 2x_1{}^2 + 1 + \cos x_2 + \sin^2 x_2$.

(ii) If $\omega_{12} \neq 0$, let

$$(3.15) \qquad A_1 := [L_0, D_1] = \omega_{12}D_2 + \frac{1}{2}\frac{\partial \eta}{\partial x_1} \in E$$

$$(3.16) \qquad A_2 := [D_1, A_1] = -\omega_{12}^2 + \frac{1}{2}\frac{\partial^2 \eta}{\partial x_1{}^2} \in E$$

$$A_3 := [L_0, A_1] = \left(\frac{1}{2}\frac{\partial^2 \eta}{\partial x_1{}^2} - \omega_{12}^2\right)D_1 + \frac{1}{2}\frac{\partial^2 \eta}{\partial x_1 \partial x_2}D_2$$
$$(3.17) \qquad\qquad + \frac{1}{2}\omega_{12}\frac{\partial \eta}{\partial x_2} + \frac{1}{4}\frac{\partial^3 \eta}{\partial x_1{}^3} + \frac{1}{4}\frac{\partial^3 \eta}{\partial x_1 \partial x_2{}^2} \in E$$

By (3.16), $\eta = d_0 x_1{}^3 + d_1 x_1{}^2 + e_2(x_2)x_1 + e_3(x_2)$. By Theorem 2.5, $d_0 = 0$. By (3.17) and $D_1 \in E$.

$$(3.18) \qquad \frac{1}{2}\frac{\partial^2 \eta}{\partial x_1 \partial x_2}D_2 + \frac{1}{2}\omega_{12}\frac{\partial \eta}{\partial x_2} + \frac{1}{4}\frac{\partial^3 \eta}{\partial x_1 \partial x_2{}^2} \in E.$$

By Theorem (3.6), $\frac{\partial^2 \eta}{\partial x_1 \partial x_2} = e_2'(x_2)$ is a polynomial of degree at most 1. Thus e_2 is a polynomial of degree at most 2 in x_2. By substituting e_2 and e_3 into η and removing D_1, x_1, and 1 from (3.15) and (3.18), one has

$$(3.19) \qquad \overline{A}_1 = \omega_{12}D_2 + \frac{1}{2}e_2 \in E$$

$$(3.20) \qquad \overline{A}_3 = \frac{1}{2}e_2'D_2 + \frac{1}{2}\omega_{12}(e_2'x_1 + e_3') \in E$$

If e_2 is a degree 2 polynomial in x_2, then $[\overline{A}_1, \overline{A}_3] - \frac{1}{2}e_2''\overline{A}_1 = \frac{1}{2}\omega_{12}^2 e_3'' - \frac{1}{4}(e_2 e_2'' + e_2'e_2') + \frac{1}{2}\omega_{12}^2 e_2''x_1 \in E$. Since the degree 2 term of $e_2 e_2'' + e_2'e_2'$ will never be zero, $e_3(x_2)$ must be a degree 4 polynomial. Now, consider

$$B_1 = [L_0, \overline{A}_3]$$
$$= \frac{1}{2}e_1''D_2{}^2 + \frac{1}{2}\omega_{12}(e_2''x_1 + e_3'')D_2 + \frac{1}{4}e_2'(e_2'x_1 + e_3') + \frac{1}{4}\omega_{12}e_3'' \in E$$
$$B_2 = [L_0, B_1] = \frac{1}{2}e_2''\left[L_0, D_2{}^2\right] + [L_0, \frac{1}{2}\omega_{12}(e_2''x_1 + e_3'')D_2]$$
$$= \frac{1}{2}\omega_{12}e_3''D_2{}^2 - \frac{1}{2}\omega_{12}e_2''D_2D_1 + \text{function} \in E.$$

From the discussion of the proof of Lemma 4.13 (iii) in [Wu-Ya], e_3''' must be a constant, Contradiction! Hence e_2 must be a degree 1 polynomial.

Consider $\omega_{12}\overline{A}_3 - \frac{1}{2}e_1'\overline{A}_1 = \frac{1}{2}\omega_{12}^2 e_2' x_1 + \frac{1}{2}\omega_{12}^2 e_3' - \frac{1}{4}e_2'e_2 \in E \Rightarrow \frac{1}{2}\omega_{12}^2 e_3' - \frac{1}{4}e_2'e_2$
is independent of x_2 and e_3 must be a degree 2 polynomial. Hence, η is a degree 2 polynomial in x_1 and x_2. E is of dimension 5 and $E = \{L_0, x_1, D_1, D_2 + cx_2, 1\}$.

For example, $f_1 = 5x_1 - 3x_2$, $f_2 = 4x_2$, $h_1 = x_1$. Then $\omega_{12} = 3$ and $\eta = 26x_1{}^2 - 30x_1x_2 + 25x_2{}^2 + 9$. It is easy to show that $E = \{L_0, x_1, D_1, D_2 - 5x_2, 1\}$.

If E has linear rank 0, h_i's must be constants, and $E = \{L_0\}$ or $E = \{L_0, 1\}$.

If E has linear rank 2, E is of maximal rank. The Ω-matrix must have constant entries and $E = \{L_0, x_1, x_2, D_1, D_2, 1\}$ by the theorem of Yau and Hu [Ya-Hu3]. In summary, we have the following.

THEOREM 3.9 (Wu–Yau [Wu-Ya]). *Let $n = 2$. If E is finite-dimensional, then*

(1) *if h_i's are constants, $E = \{L_0\}$ or $E = \{L_0, 1\}$;*

(2) *otherwise, Ω-matrix has constant entries. h_i's must be affine in x_1 and x_2.*
E has dimension of either 4, 5, or 6.

Moreover, from the above discussion, it is easy to see that if E is finite-dimensional, it has only elements with order less than or equal to 2. Thus, the Levine conjecture holds for the finite-dimensional estimation algebras with state dimension 2.

4. Direct Method of Solving the DMZ Equation. The direct method was first introduced by Yau–Yau [Ya-Ya1][Ya-Ya2] to solve DMZ equations arising from a linear filtering model or an exact filtering model. This direct method was generalized by Yau–Hu [Ya-Hu1] [Ya-Hu2] to treat a more general filtering system.

Recall that in 1990, Yau [Ya1] (cf. [Ya2] for detailed version) first studied the filtering model (2.1) with the following conditions:

$$(c_1') \qquad \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = \text{constant } c_{ij} \text{ for all } 1 \leqslant i, j \leqslant n.$$

This was called the Yau filtering model in [Ch]. The Yau filtering model includes linear filtering model and exact filtering model. By Theorem 3.3, we know that (c_1') is equivalent to the following condition:

$$(4.1) \qquad\qquad (c_1) \qquad f_i(x) = l_i(x) + \frac{\partial F}{\partial x_i}(x), \quad 1 \leqslant i \leqslant n$$

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

A theorem in [Ya2] tells us that $h_1, \ldots, h_m$ are polynomials of degree at most one if the Yau filtering model has a finite-dimensional estimation algebra. So we list the following condition:

$$(4.2) \qquad\qquad (c_2) \qquad h_i(x) = \sum_{j=1}^n c_{ij}x_j + c_i, \quad 1 \leqslant i \leqslant m$$

where c_{ij} and c_i are constants.

Moreover, by a theorem in [Ya2], we know that $\eta(x)$ is a polynomial of degree at most two for Yau filtering models with finite-dimensional estimation algebra. Hence we assume the following condition:

$$(4.3) \qquad (c_3) \qquad \eta(x) = \sum_{i,j=1}^{n} \eta_{ij} x_i x_j + \sum_{i=1}^{n} \eta_i x_i + \eta_0$$

where η_{ij}, η_i, and η_0 are constants.

LEMMA 4.1 (Yau–Hu [Ya-Hu2]). *Equation* (2.4) *is equivalent to the following equation*

$$(4.4) \qquad \left. \begin{aligned} \frac{\partial u}{\partial t}(t,x) &= \frac{1}{2}\Delta u(t,x) + \sum_{i=1}^{n} \theta_i(t,x)\frac{\partial u}{\partial x_2}(t,x) + \theta(t,x)u(t,x) \\ u(0,x) &= \sigma_0(x) \end{aligned} \right\}$$

where

$$\theta_i(t,x) = \sum_{j=1}^{m} y_j(t)\frac{\partial h_j}{\partial x_i}(x) - f_i(x),$$

$$\theta(t,x) = \frac{1}{2}\left(\sum_{i=1}^{n} \theta_i{}^2(t,x) + \sum_{i=1}^{n} \frac{\partial \theta_i}{\partial x_i}(t,x) - \eta(x) \right).$$

THEOREM 4.1 (Yau–Hu [Ya-Hu2]). *Suppose* $u(t,x)$ *is a solution of* (4.4) *and*

$$(4.5) \qquad \tilde{u}(t,x) = e^{\Lambda(t,x)}u(t,x+b(t))$$

Then $\tilde{u}(t,x)$ *is the solution of the following Kolmogorov equation:*

$$(4.6) \qquad \left. \begin{aligned} \frac{\partial \tilde{u}}{\partial t}(t,x) &= \frac{1}{2}\Delta \tilde{u}(t,x) - \sum_{i=1}^{n} H_i(x)\frac{\partial \tilde{u}}{\partial x_i}(t,x) - P(x)\tilde{u}(t,x) \\ \tilde{u}(0,x) &= e^{\Lambda(0,x)}u(0,x+b(0)). \end{aligned} \right\}$$

if we can choose $H_i(x)$ *and* $P(x)$ *such that* $\Lambda(t,x)$ *satisfy the following systems:*

$$(4.7) \qquad b_i'(t) - \frac{\partial \Lambda}{\partial x_i}(t,x) + H(x) + \theta_i(t,x+b(t)) \equiv 0 \quad 1 \leqslant i \leqslant n$$

$$(4.8) \quad \frac{\partial \Lambda}{\partial t}(t,x) - \frac{1}{2}\sum_{i=1}^{n}(b_i'(t))^2 - \sum_{i=1}^{n}\theta_i(t,x+b(t))b_n'(t) - \frac{1}{2}\eta(x+b(t))$$

$$+ \frac{1}{2}\sum_{i=1}^{n} H_i{}^2(x) - \frac{1}{2}\sum_{i=1}^{n}\frac{\partial H_i}{\partial x_i}(x) + P(x) \equiv 0$$

Moreover, if $u, \dfrac{\partial u}{\partial x_1}, \ldots, \dfrac{\partial u}{\partial x_n}$ *are linearly independent, then* (4.7) *and* (4.8) *are also necessary conditions for* (4.6).

Next if we assume that condition (c_3) holds, then

$$(4.9) \qquad \eta(x + b(t)) = \eta(x) + \sum_{i=1}^{n} B_i(t)x_i + B(t)$$

where

$$B_i(t) = \sum_{j=1}^{n} (n_{ij} + \eta_{ji})b_j(t),$$

$$B(t) = \sum_{i,j=1}^{n} n_{ij}b_i(t)b_j(t) + \sum_{i=1}^{n} \eta_i b_i(t)$$

The following theorem says that in many situations if the solution of (4.4) can be represented in the form of (4.5) and (4.6), then condition (c_1) in (4.1) holds.

THEOREM 4.2 (Yau–Hu [Ya-Hu2]). *Assume conditions (c_2) and (c_3) hold. Suppose $u(t, x)$ is a solution of (4.4) and*

$$\tilde{u}(t, x) = e^{\Lambda(t,x)} u(t, x + b(t))$$

where $b'_i(t)$, $1 \leqslant i \leqslant n$, are linearly independent and $\tilde{u}(t, x)$ is the solution of the following Kolmogorov equation:

$$\begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t, x) = \dfrac{1}{2}\Delta\tilde{u}(t, x) - \sum_{i=1}^{n} H_i(x)\dfrac{\partial \tilde{u}}{\partial x_i}(t, x) - P(x)\tilde{u}(t, x) \\ \tilde{u}(0, x) = e^{\Lambda(0,x)} u(0, x + b(0)). \end{cases}$$

Furthermore, let

$$\Lambda(t, x) = c(t) + G(x) + \sum_{i=1}^{n} a_j(t)x_j - F(x + b(t)).$$

If we can choose $H(x)$, $G(x)$, and $P(x)$ such that

$$(4.10) \qquad \frac{1}{2}\sum_{i=1}^{n} H_i^2(x) - \frac{1}{2}\sum_{i=1}^{n} \frac{\partial H}{\partial x_i}(x) - \frac{1}{2}\eta(x) + P(x) \equiv 0,$$

then

$$f_i(x) = \frac{\partial F}{\partial x_i}(x) + l_i(x), \quad 1 \leqslant i \leqslant n,$$

where $l_i(x) = \sum_{j=1}^{n} d_{ij}x_j + d_i$, d_{ij} and d_i are constants, i.e., condition (c_1) holds.

The following theorem plays an important role in our direct method.

THEOREM 4.3 (Yau-Hu [Ya-Hu2]). *Consider the filtering model (2.1) with conditions (c_1), (c_2), and (c_3). Then the solution $u(t, x)$ for the DMZ equation (2.4) or (4.4) is reduced to the solution of $\tilde{u}(t, x)$ for the Kolmogorov equation*

$$\begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t, x) = \dfrac{1}{2}\Delta\tilde{u}(t, x) - \sum_{i=1}^{n} H_i(x)\dfrac{\partial \tilde{u}}{\partial x_i}(t, x) - P(x)\tilde{u}(t, x) \\ \tilde{u}(0, x) = e^{G(x)-F(x)}\sigma_0(x) \end{cases}$$

where

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

and $a_i(t)$, $b_i(t)$, and $c(t)$ satisfy the following system of ODEs:

(4.11)
$$\begin{cases} b_i'(t) - a_i(t) - \sum_{j=1}^{n} d_{ij}b_j(t) + \sum_{j=1}^{n} c_{ji}y_j(t) = 0 \\ b_i(0) = 0 \qquad 1 \leqslant i \leqslant n \end{cases}$$

(4.12)
$$\begin{cases} a_i'(t) - \frac{1}{2}\sum_{j=1}^{n}(\eta_{ij} + \eta_{ji})b_j(t) + \sum_{j=1}^{n} d_{ji}b_j'(t) = 0 \\ a_i(0) = 0 \qquad 1 \leqslant i \leqslant n \end{cases}$$

(4.13)
$$\begin{cases} c_i'(t) = -\frac{1}{2}\sum_{i=1}^{n}(b_i'(t))^2 + \sum_{i=1}^{n} a_i(t)b_i'(t) - \sum_{i=1}^{n} d_i b_i'(t) \\ \qquad + \frac{1}{2}\sum_{i,j=1}^{n} \eta_{ij}b_i(t)b_j(t) + \frac{1}{2}\sum_{i=1}^{n} \eta_i b_i(t) \\ c(0) = 0, \end{cases}$$

if we can choose $H(x)$, $G(x)$, and $P(x)$ such that

(4.14)
$$\frac{1}{2}\sum_{i=1}^{n} H_i^{\,2}(x) - \frac{1}{2}\sum_{i=1}^{n}\frac{\partial H_i}{\partial x_i}(x) - \frac{1}{2}\eta(x) + P(x) \equiv 0$$

where $H_i(x) - \frac{\partial G}{\partial x_i}(x) = l_i(x)$.

We have several choices of $H(x)$, $G(x)$, and $P(x)$. We list some of them as follows.

COROLLARY 4.1. *Choose a C^∞ function $G(x)$ such that*

$$\Delta G(x) + |\nabla G|^2(x) + 2\sum_{i=1}^{n} l_i(x)\frac{\partial G}{\partial x_i}(x) \equiv \eta(x) - \sum_{i=1}^{n} l_i^{\,2}(x) - \sum_{i=1}^{n}\frac{\partial l_i}{\partial x_i}(x).$$

Let

$$H_i(x) = \frac{\partial G}{\partial x_i}(x) + l_i(x), \quad 1 \leqslant i \leqslant n,$$

and

$$P(x) = \sum_{i=1}^{n}\frac{\partial H_i}{\partial x_i}(x) = \sum_{i=1}^{n}\left(\frac{\partial^2 G}{\partial x_i^{\,2}}(x) + \frac{\partial l_i}{\partial x_i}(x)\right)$$

Then the corresponding Kolmogorov equation is

$$\begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t,x) = \dfrac{1}{2}\Delta \tilde{u}(t,x) - \sum_{i=1}^{n}\left(l_i(x) + \dfrac{\partial G}{\partial x_i}(x)\right)\dfrac{\partial \tilde{u}}{\partial x_i}(t,x) \\ \qquad\qquad\qquad - \sum_{i=1}^{n}\left(\dfrac{\partial l_i}{\partial x_i}(x) + \dfrac{\partial^2 G}{\partial x_i^{\,2}}(x)\right)\tilde{u}(t,x) \\ \tilde{u}(0,x) = e^{G(x)-F(x)}\sigma_0(x) \end{cases}$$

where

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

and $a_i(t)$, $b_i(t)$, and $c(t)$ satisfy ODEs (4.11), (4.12) and (4.13).

COROLLARY 4.2. *Choose*

$$G(x) \equiv 0$$

$$P(x) = \frac{1}{2}\eta(x) - \frac{1}{2}\sum_{i=1}^{n} l_i{}^2(x) - \frac{1}{2}\sum_{i=1}^{n} \frac{\partial l_i}{\partial x_i}(x)$$

$$H_i(x) = l_i(x), \quad 1 \leqslant i \leqslant n.$$

Then the corresponding Kolmogorov equation becomes

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

where

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

and $a_i(t)$, $b_i(t)$, and $c(t)$ satisfy ODEs (4.11), (4.12) and (4.13).

COROLLARY 4.3. *Choose a C^∞ function $G(x)$ such that $\frac{\partial G}{\partial x_i}(x) = -l_i(x)$ if $d_{ij} = d_{ji}$ for $1 \leqslant i, j \leqslant n$. Let $P(x) = \frac{1}{2}\eta(x)$ and $H_i(x) \equiv 0$, $1 \leqslant i \leqslant n$. Then the corresponding Kolmogorov equation becomes*

$$\begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t, x) = \dfrac{1}{2}\Delta \tilde{u}(t, x) - \dfrac{1}{2}\eta(x)\tilde{u}(t, x) \\ \tilde{u}(0, x) = e^{G(x) - F(x)}\sigma_0(x). \end{cases}$$

where

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

and $a_i(t)$, $b_i(t)$, and $c(t)$ satisfy ODEs (4.11), (4.12) and (4.13).

COROLLARY 4.4. *Choose*

$$G(x) = F(x)$$

$$P(x) = \frac{1}{2}\eta(x) - \frac{1}{2}\sum_{i=1}^{n} f_i{}^2(x) + \frac{1}{2}\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(x)$$

$$H_i(x) = f_i(x), \quad 1 \leqslant i \leqslant n.$$

Then the corresponding Kolmogorov equation becomes

$$\begin{cases} \dfrac{\partial \tilde{u}}{\partial t}(t,x) = \dfrac{1}{2}\Delta \tilde{u}(t,x) - \displaystyle\sum_{i=1}^{n} f_i(x) + \dfrac{\partial \tilde{u}}{\partial x_i}(t,x) \\[4mm] \qquad\qquad + \dfrac{1}{2}\Big(\displaystyle\sum_{i=1}^{n} f_i{}^2(x) - \sum_{i=1}^{n} \dfrac{\partial f_i}{\partial x_i}(x) - \eta(x) \Big)\tilde{u}(t,x) \\[4mm] \tilde{u}(0,x) = \sigma_0(x) \end{cases}$$

where

$$\tilde{u}(t,x) = \exp\Big[c(t) + \sum_{i=1}^{n} a_i(t)x_i + F(x) - F(x+b(t)) \Big] \cdot u(t,x+b(t))$$

and $a_i(t)$, $b_i(t)$, and $c(t)$ satisfy ODEs (4.11), (4.12) and (4.13).

5. New Algorithms in Real Time Solution of the Nonlinear Filtering Problem. The purpose of this section is to describe our new algorithms which can solve all engineering problems in real time. In section 5.1, we give some detail why the solution arising from our algorithm converges to the true solution of the DMZ equation for the filtering model (2.1). In section 5.2, we describe how our algorithm works for the general filtering model (1.1).

5.1. Yau–Yau algorithm of solving nonlinear filtering problems.
The direct method described in the previous section works very nicely for the Yau filtering system and it is very easy to implement. However, in all the direct methods in [Ya-Ya1] [Ya-Ya2] [Ya-Hu1][Ya-Hu2], they need to assume that all the observation terms $h_i(x)$, $1 \leqslant i \leqslant m$, are degree one polynomials. On the other hand, in many practical examples, e.g., tracking problem, the observation terms may be nonlinear. In 2000, we [Ya-Ya3] proposed a novel algorithm to solve the DMZ equation in real time and memoryless way. Under the assumptions that the drift terms $f_i(x)$, $1 \leqslant i \leqslant n$, and their first and second derivatives, and the observation terms $h_i(x)$, $1 \leqslant i \leqslant m$, and their first derivatives, have linear growth, we showed that the solution obtained from our algorithms converges to the true solution of the DMZ equation. Recently we [Ya-Ya4] show that under very mild conditions which essentially say that the growth of $|h|$ is greater than the growth of $|f|$, the DMZ equation admits a unique nonnegative solution u which can be approximated by solutions u_R of DMZ equation on the ball B_R with $u_R|_{\partial B_R} = 0$. The rate of convergence can be efficiently estimated in L^1-norm. The solutions u_R can in turn be approximated efficiently by an algorithm depending only on solving the time independent Kolmogorov equation on B_R. Our algorithm can solve practically all engineering problems including the cubic sensor problem in real time and memoryless manner.

It is shown in [Ya-Ya3] that the robust DMZ equation can be written in the

following form

$$
(5.1) \quad
\begin{cases}
\dfrac{\partial u}{\partial t}(t,x) = \dfrac{1}{2}\Delta u(t,x) + (-f(x) + \nabla K(t,x)) \cdot \nabla u(t,x) \\
\qquad\qquad + \left(-\operatorname{div} f(x) - \dfrac{1}{2}|h(x)|^2 + \dfrac{1}{2}\Delta K(t,x) \right. \\
\qquad\qquad\qquad \left. - f(x)\cdot\nabla K(t,x) + \dfrac{1}{2}|\nabla K(t,x)|^2 \right) u(t,x) \\
u(0,x) = \sigma_0(x).
\end{cases}
$$

where $K = \sum_{j=1}^{m} y_i(t)h_j(x)$, $f = (f_1,\ldots,f_n)$ and $h = (h_1,\ldots,h_m)$.

To simplify our presentation, we introduce the following condition.

Condition (C_1):

$$
-\frac{1}{2}|h|^2 - \frac{1}{2}\Delta K - f\cdot\nabla K + \frac{1}{2}|\nabla K|^2 + |f - \nabla K| \leq c_1 \; \forall \; (t,x) \in [0,T]\times\mathbb{R}^n,
$$

where c_1 is a constant possibly depending on T.

Our main theorems are as follows:

THEOREM 5.1 ([Ya-Ya4]). *Consider the filtering model* (2.1). *For any $T > 0$, let u be a solution of the robust DMZ equation* (5.1) *in $[0,T]\times\mathbb{R}^n$. Assume Condition* (C_1) *is satisfied.*

Then

$$
(5.2) \qquad \sup_{0\leq t\leq T}\int_{\mathbb{R}^n} e^{\sqrt{1+|x|^2}}u(t,x) \leq e^{(c_1+\frac{n+1}{2})T}\int_{\mathbb{R}^n} e^{\sqrt{1+|x|^2}}u(0,x).
$$

In particular,

$$
(5.3) \qquad \sup_{0\leq t\leq T}\int_{|x|\geq R} u(t,x) \leq e^{-\sqrt{1+R^2}}e^{(c_1+\frac{n+1}{2})T}\int_{\mathbb{R}^n} e^{\sqrt{1+|x|^2}}u(0,x).
$$

Theorem 5.1 above says that one can choose a ball large enough to capture almost all the density. In fact we have a precise estimate of density lying outside this ball by (5.3).

THEOREM 5.2 ([Ya-Ya4]). *Consider the filtering model* (2.1). *For any $T > 0$, let u be a solution of the robust DMZ equation* (5.1) *in $[0,T]\times\mathbb{R}^n$. Assume*

(1) *Condition* (C_1) *is satisfied.*

(2) $-\frac{1}{2}|h|^2 - \frac{1}{2}\Delta K - f(x)\cdot\nabla K(t,x) + \frac{1}{2}|\nabla K|^2 + 12 + 2n + 4|f - \nabla K| \leq c_2$ *for all $(t,x) \in [0,T]\times\mathbb{R}^n$, where c_2 is a constant possibly depending on T.*

(3) $e^{-\sqrt{1+|x|^2}}[12 + 2n + 4|f - \nabla K|] \leq c_3$ *for all $(t,x) \in [0,T]\times\mathbb{R}^n$.*

Let $R \geq 1$ and u_R be the solution of the following DMZ equation on the ball B_R

$$
(5.4) \quad
\begin{cases}
\dfrac{\partial u_R}{\partial t} = \dfrac{1}{2}\Delta u_R + (-f + \nabla K)\cdot\nabla u_R \\
\qquad\quad + \left(-\operatorname{div} f - \dfrac{1}{2}|h|^2 + \dfrac{1}{2}\Delta K - f\cdot\nabla K + \dfrac{1}{2}|\nabla K|^2 \right) u_R \\
u_R(t,x) = 0 \qquad \text{for } (t,x) \in [0,T]\times\partial B_R \\
u_R(0,x) = \sigma_0(x).
\end{cases}
$$

Let $v = u - u_R$. Then $v \geq 0$ for all $(t, x) \in [0, T] \times B_R$ and

$$(5.5) \qquad \int_{B_R} \phi v(T, x) \leq \frac{e^{c_2 T} - 1}{c_2} c_3 e^{-R} e^{(c_1 + \frac{n+1}{2})T} \int_{\mathbb{R}^n} e^{\sqrt{1 + |x|^2}} u(0, x)$$

where $\phi(x) = e^{\frac{|x|^4}{R^3} - \frac{2|x|^2}{R}} - e^{-R}$. In particular

$$(5.6) \quad \int_{B_{\frac{R}{2}}} v(T, x)$$

$$\leq \frac{2(e^{c_2 T} - 1)}{c_2} c_3 e^{-\frac{9}{16} R} e^{(c_1 + \frac{n+1}{2})T} \int_{\mathbb{R}^n} e^{\sqrt{1 + |x|^2}} u(0, x).$$

Theorem 5.2 above says that we can approximate u by u_R. The approximation is good if R is large enough. In fact we have a precise error estimate of this approximation by (5.6).

THEOREM 5.3 ([Ya-Ya4]). *Let Ω be a bounded domain in $\mathbb{R}^n$. Let $F \colon [0, T] \times \Omega \to \mathbb{R}^n$ be a family of vector fields C^∞ in x and Holder continuous in t with exponent α and $J \colon [0, T] \times \Omega \to \mathbb{R}$ be a C^∞ function in x and Holder continuous in t with exponent α such that the following properties are satisfied*

$$(5.7) \qquad |\operatorname{div} F(t, x)| + 2|J(t, x)| + |F(t, x)| \leq c \text{ for } (t, x) \in [0, T] \times \Omega$$

$$(5.8) \quad |F(t, x) - F(\bar{t}, x)| + |\operatorname{div} F(t, x) - \operatorname{div} F(\bar{t}, x)|$$

$$+ |J(t, x) - J(\bar{t}, x)| \leq c_1 |t - \bar{t}|^\alpha$$

$$\text{for } (t, x), (\bar{t}, x) \in [0, T] \times \Omega.$$

Let $u(t, x)$ be the solution on $[0, T] \times \Omega$ of the equation

$$(5.9) \qquad \begin{cases} \dfrac{\partial u}{\partial t}(t, x) = \dfrac{1}{2} \Delta u(t, x) + F(t, x) \cdot \nabla u(t, x) + J(t, x) u(t, x) \\ u(0, x) = \sigma_0(x) \\ u(t, x)\big|_{\partial \Omega} = 0. \end{cases}$$

For any $0 \leq \tau \leq T$, let $\mathcal{P}_k = \{0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_k = \tau\}$ be a partition of $[0, \tau]$ where $\tau_i = \frac{i\tau}{k}$. Let $u_i(t, x)$ be the solution on $[\tau_{i-1}, \tau_i] \times \Omega$ of the following equation

$$(5.10) \qquad \begin{cases} \dfrac{\partial u_i}{\partial t}(t, x) = \dfrac{1}{2} \Delta u_i(t, x) + F(\tau_{i-1}, x) \cdot \nabla u_i(t, x) \\ \qquad\qquad + J(\tau_{i-1}, x) u_i(t, x) \\ u_i(\tau_{i-1}, x) = u_{i-1}(\tau_{i-1}, x) \\ u_i(t, x)\big|_{\partial \Omega} = 0. \end{cases}$$

Here we use the convention $u_0(t, x) = \sigma(x)$. Then the solution $u(t, x)$ of (5.9) can be computed by means of the solution $u_i(t, x)$ of (5.10). More specifically, $u(\tau, x) = \lim_{k \to \infty} u_k(\tau_k, x)$ in L^1-sense on Ω and the following estimate holds

$$(5.11) \qquad \int_\Omega |u - u_k|(\tau_k, x) \leq \frac{2c_2}{\alpha + 1} \frac{T^{\alpha+1} e^{cT}}{k^\alpha}$$

where

$$(5.12) \qquad c_2 = c_1 e^{cT} + c_1 \sqrt{\mathrm{Vol}\,(\Omega)} e^{c^2 T} \sqrt{2c^2 T \int_\Omega u^2(0, x) + \int_\Omega |\nabla u(0, x)|^2}.$$

The right hand side of (5.11) goes to zero as $k \to \infty$.

In case (5.9) and (5.10) are DMZ equations, i.e., $F(t, x) = -f(x) + \nabla K$ and $J(t, x) = -\mathrm{div}\, f - \frac{1}{2}|h|^2 + \frac{1}{2}\Delta K - f \cdot \nabla K + \frac{1}{2}|\nabla K|^2$, where $K = \sum_{j=1}^m y_j(t)h_j(x)$, by Proposition 5.1 below (which is similar to Proposition 3.1 of [Ya-Ya3]), $u_i(\tau_i, x)$ can be computed by $\widetilde{u}_i(\tau_i, x)$ where $\widetilde{u}_i(t, x)$ for $\tau_{i-1} \leq t \leq \tau_i$ satisfies the following Kolmogorov equation

$$(5.13) \qquad \begin{cases} \dfrac{\partial \widetilde{u}_i}{\partial t}(t, x) = \dfrac{1}{2}\Delta \widetilde{u}_i(t, x) - \displaystyle\sum_{j=1}^n f_j(x)\dfrac{\partial \widetilde{u}_i}{\partial x_j}(t, x) \\ \qquad\qquad - \Big(\mathrm{div} f(x) + \dfrac{1}{2}\displaystyle\sum_{j=1}^m h_j^2(x)\Big)\widetilde{u}_i(t, x) \\ \widetilde{u}_i(\tau_{i-1}, x) = \exp\Big(\displaystyle\sum_{j=1}^m (y_j(\tau_{i-1}) - y_j(\tau_{i-2}))h_j(x)\Big)\widetilde{u}_{i-1}(\tau_{i-1}, x). \end{cases}$$

In fact

$$(5.14) \qquad u_i(\tau_i, x) = \exp\Big(-\sum_{j=1}^m y_j(\tau_{i-1})h_j(x)\Big)\widetilde{u}_i(\tau_i, x).$$

Therefore theoretically to solve the DMZ equation in a real time manner, we only need to compute the following Kolmogorov equations off-line

$$\begin{cases} \dfrac{\partial \widetilde{u}}{\partial t}(t, x) = \dfrac{1}{2}\Delta \widetilde{u}(t, x) - \displaystyle\sum_{j=1}^n f_j(x)\dfrac{\partial \widetilde{u}}{\partial x_j}(t, x) \\ \qquad\qquad - \Big(\mathrm{div}\, f(x) + \dfrac{1}{2}\displaystyle\sum_{j=1}^m h_j^2(x)\Big)\widetilde{u}(t, x) \\ \widetilde{u}(0, x) = \phi_i(x) \end{cases}$$

where $\{\phi_i(x)\}$ is an orthonormal base in $L^2(\mathbb{R}^n)$. The only real time computation here is to express arbitrary initial condition $\phi(x)$ as the linear combination of $\phi_i(x)$. But this can be done by means of parallel computation.

The following Proposition 5.1 plays a fundamental role in our real time solution to the robust DMZ equation (5.1) in memoryless manner.

PROPOSITION 5.1. $\widetilde{u}(t, x)$ satisfies the following Kolmogorov equation

$$\frac{\partial \widetilde{u}}{\partial t}(t, x) = \frac{1}{2}\Delta\widetilde{u}(t, x) - f(x) \cdot \nabla\widetilde{u}(t, x)$$

(5.15)
$$- \left(\operatorname{div} f(x) + \frac{1}{2}\sum_{i=1}^{m} h_i^2(x) \right)\widetilde{u}(t, x)$$

for $\tau_{\ell-1} \leq t \leq \tau_\ell$ if and only if

(5.16)
$$u(t, x) = e^{-\sum_{i=1}^{m} y_i(\tau_{\ell-1})h_i(x)}\widetilde{u}(t, x)$$

satisfies the robust DMZ equation with observation being frozen at $y(\tau_{\ell-1})$

$$\frac{\partial u}{\partial t}(t, x) = \frac{1}{2}\Delta u(t, x) + (-f(x) + \nabla K(\tau_{\ell-1}, x)) \cdot \nabla u(t, x)$$

(5.17)
$$+ \left(-\operatorname{div} f(x) - \frac{1}{2}|h(x)|^2 + \frac{1}{2}\Delta K(\tau_{\ell-1}, x)\right.$$

$$\left. -f(x) \cdot \nabla K(\tau_{\ell-1}, x) + \frac{1}{2}|\nabla K(\tau_{\ell-1}, x)|^2 \right) u(t, x).$$

We remark that (5.17) is obtained from the robust DMZ equation by freezing the observation term $y(t)$ to $y(\tau_{\ell-1})$. We shall show that the solution of (5.17) approximates the solution of the robust DMZ equation very well in L^1-sense.

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

(5.18)
$$\begin{cases} \dfrac{\partial u_i}{\partial t}(t, x) = \dfrac{1}{2}\Delta u_i(t, x) + (-f(x) + \nabla K(\tau_{i-1}, x)) \cdot \nabla u_i(t, x) \\ \qquad\quad + \left(-\operatorname{div} f(x) - \dfrac{1}{2}|h(x)|^2 \right. \\ \qquad\qquad\quad + \dfrac{1}{2}\Delta K(\tau_{i-1}, x) \\ \qquad\qquad\qquad - f(x) \cdot \nabla K(\tau_{i-1}, x) \\ \qquad\qquad\qquad\quad \left. + \dfrac{1}{2}|\nabla K(\tau_{i-1}, x)|^2 \right) u_i(t, x) \\ u_i(\tau_{i-1}, x) = u_{i-1}(\tau_{i-1}, x). \end{cases}$$

By Theorem 5.3 we have that $u(\tau, x) = \lim_{k \to \infty} u_k(\tau_k, x)$ in L^1-sense. By Proposition 5.1, $u_1(\tau_1, x)$ can be computed by $\widetilde{u}_1(\tau_1, x)$ where $\widetilde{u}_1(t, x)$ for $0 \leq t \leq \tau_1$ satisfies equation (5.15) with initial condition

(5.19)
$$\widetilde{u}_1(0, x) = \sigma_0(x).$$

In fact

(5.20)
$$u_1(\tau_1, x) = \widetilde{u}_1(\tau_1, x).$$

In general Proposition 5.1 tells us that for $i \geq 2$, $u_i(\tau_i, x)$ can be computed by $\widetilde{u}_i(\tau_i, x)$, where $\widetilde{u}_i(t, x)$ for $\tau_{i-1} \leq t \leq \tau_i$ satisfies equation (5.15) with initial condition

$$(5.21) \qquad \widetilde{u}_i(\tau_{i-1}, x) = \exp\Big[\sum_{j=1}^{m}(y_j(\tau_{i-1}) - y_j(\tau_{i-2}))h_j(x)\Big]\widetilde{u}_{i-1}(\tau_{i-1}, x)$$

where the last initial condition comes from

$$\widetilde{u}_i(\tau_{i-1}, x)$$

$$= u_i(\tau_{i-1}, x)\exp\Big(\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)\Big)$$

$$= u_{i-1}(\tau_{i-1}, x)\exp\Big(\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)\Big)$$

$$= \exp\Big(-\sum_{j=1}^{m} y_j(\tau_{i-2})h_j(x)\Big)\widetilde{u}_{i-1}(\tau_{i-1}, x)\exp\Big(\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)\Big)$$

$$= \exp\Big[\sum_{j=1}^{m}(y_j(\tau_{i-1}) - y_j(\tau_{i-2}))h_j(x)\Big]\widetilde{u}_{i-1}(\tau_{i-1}, x).$$

In fact,

$$(5.22) \qquad u_i(\tau_i, x) = \exp\Big(-\sum_{j=1}^{m} y_j(\tau_{i-1})h_j(x)\Big)\widetilde{u}_i(\tau_i, x)$$

THEOREM 5.4 ([Ya-Ya4]). *Let u_R be the solution of (5.4) the DMZ equation on B_R. Assume that*

(1) *$f(x)$ and $h(x)$ have at most polynomial growth.*
(2) *For any $0 \leq t \leq T$, there exist positive integer m and positive constants c' and c'' independent of R such that the following two inequalities hold on $\mathbb{R}^n$.*
   (a) $\frac{m^2}{2}|x|^{2m-2} - \frac{m}{2}(m+n-2)|x|^{m-2} - m|x|^{m-2}x \cdot (f - \nabla K) - \frac{\Delta K}{2} - \frac{1}{2}|h|^2 - f \cdot \nabla K + \frac{1}{2}|\nabla K|^2 \geq -c'$
   (b) $\big|\frac{m^2|x|^{2m-2}}{2} - \frac{m(m+n-2)}{2}|x|^{m-2} - m|x|^{m-2}(f - \nabla K) \cdot x\big|$
        $\leq \frac{m(m+1)}{2}|x|^{2m-2} + c''$
(3) $-\frac{1}{2}|h|^2 - \frac{1}{2}\Delta K - \sum_{j=1}^{n} f_j\frac{\partial K}{\partial x_j} + \frac{1}{2}|\nabla K|^2 \leq c_1$ *for all $(t, x) \in [0, T] \times \mathbb{R}^n$ where c_1 is a constant possibly depending on T.*

Then for any $R_0 < R$,

$$\int_{B_{R_0}} \big(e^{-|x|^m} - e^{-R_0^m}\big)u_R(T, x)$$

$$\geq e^{-c'T}\int_{B_{R_0}} (e^{-|x|^m} - e^{-R_0^m})\sigma_0(x)$$

$$+ \frac{e^{-R_0^m}}{c'}\Big(\frac{m(m+1)}{2}R_0^{2m-2} + c''\Big)\Big(1 - e^{c'T}\Big)\int_{B_R}\sigma_0(x).$$

In particular, the solution u of the robust DMZ equation on $\mathbb{R}^n$ has the following estimate

$$\int_{\mathbb{R}^n} e^{-|x|^m} u(T, x) \geq e^{-c'T} \int_{\mathbb{R}^n} e^{-|x|^m} \sigma_0(x).$$

In practical nonlinear filtering computation, it is important to know how much density remains within the given ball. Theorem 5.4 provides such a lower estimate. In particular, the solution u of the DMZ equation in $\mathbb{R}^n$ obtained by taking $\lim\limits_{R \to \infty} u_R$, where u_R is the solution of the DMZ equation in the ball B_R, is a nontrivial solution.

5.2. Extended Yau–Yau Algorithm for General Nonlinear Filtering Model.

The purpose of this subsection is to describe the extended Yau–Yau algorithm for general nonlinear filtering model (1.1) developed by Yan and Yau [Yan-Ya]. For the system (1.1), the unnormalized density $\sigma(x, t)$ of x_t conditioned on the observation history $Y_t = \{y_s, 0 \leqslant s \leqslant t\}$ is given by the following DMZ equation

$$(5.23) \qquad \begin{cases} d\sigma(x,t) = L\sigma(x,t)dt + \sigma(x,t)h^T(x,t)R^{-1}(t)dy_t \\ \sigma(x,0) = \sigma_0(x) \end{cases}$$

in which $\sigma_0(x)$ is the probability density of the initial point x_0, and

$$(5.24) \qquad L(*) \equiv \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2}{\partial x_i \partial x_j}\left[(GQG^T)_{ij}*\right] - \sum_{i=1}^{n}\frac{\partial(f_i*)}{\partial x_i}$$

The DMZ equation (5.23) is a stochastic partial differential equation due to the term with dy_t, there is no easy way to derive a recursive algorithm for solving this equation. In real applications, we are interested in constructing robust state estimators from observed sample paths with some property of robustness. Using Rozovsky's transformation [Ro]

$$(5.25) \qquad \sigma(x,t) = \exp\left[h^T(x,t)R^{-1}(t)y_t\right]\rho(x,t),$$

we obtain the following robust DMZ equation for $\rho(x, t)$ which involves y_t only in the coefficients:

$$(5.26) \qquad \begin{cases} \dfrac{\partial\rho}{\partial t} + \dfrac{\partial}{\partial t}\left(h^T R^{-1}\right)^T y_t \rho \\ \qquad = \exp\left(-h^T R^{-1}y_t\right)\left[L - \dfrac{1}{2}h^T R^{-1}h\right]\left[\exp\left(h^T R^{-1}y_t\right)\rho\right] \\ \rho(x,0) = \sigma_0(x) \end{cases}$$

PROPOSITION 5.2 ([Yan-Ya]). *The robust DMZ equation (5.26) is equivalent to*

the following partial differential equation,

$$
(5.27)\quad
\begin{cases}
\dfrac{\partial \rho}{\partial t} + \dfrac{\partial}{\partial t}\left(h^T R^{-1}\right)^T y_t \rho \\[2mm]
\quad = \dfrac{1}{2}\sum_{i,j=1}^{n}\left(GQG^T\right)\dfrac{\partial^2 \rho}{\partial x_i \partial x_j} \\[2mm]
\quad + \sum_{i=1}^{n}\left[\sum_{j=1}^{n}\dfrac{\partial}{\partial x_j}\left(GQG^T\right)_{ij}\right. \\[2mm]
\qquad\qquad \left. + \sum_{j=1}^{n}\left(GQG^T\right)_{ij}\left(\dfrac{\partial h}{\partial x_j}R^{-1}y_t\right) - f_i\right]\dfrac{\partial \rho}{\partial x_i} \\[2mm]
\quad + \left\{\dfrac{1}{2}\sum_{i,j=1}^{n}\dfrac{\partial^2}{\partial x_i \partial x_j}\left(GQG^T\right)_{ij}\right. \\[2mm]
\qquad + \sum_{i,j=1}^{n}\dfrac{\partial}{\partial x_i}\left(GQG^T\right)_{ij}\left(\dfrac{\partial h}{\partial x_j}R^{-1}y_t\right) \\[2mm]
\qquad + \dfrac{1}{2}\sum_{i,j=1}^{n}\left(GQG^T\right)_{ij}\left[\left(\dfrac{\partial^2 h}{\partial x_i \partial x_j}R^{-1}y_t\right)\right. \\[2mm]
\qquad\qquad\qquad \left. + \left(\dfrac{\partial h}{\partial x_i}R^{-1}y_t\right)\left(\dfrac{\partial h}{\partial x_j}R^{-1}y_t\right)\right] \\[2mm]
\qquad \left. - \sum_{i=1}^{n}\dfrac{\partial f_i}{\partial x_i} - \sum_{i=1}^{n}f_i\left(\dfrac{\partial h}{\partial x_i}R^{-1}y_t\right) - \dfrac{1}{2}\left(h^T R^{-1}h\right)\right\}\rho \\[2mm]
\rho(x,0) = \sigma_0(x)
\end{cases}
$$

where for h being an $m \times 1$ column vector, we define

$$
\frac{\partial h}{\partial t} \equiv \left[\frac{\partial h_1}{\partial t}\ \frac{\partial h_2}{\partial t}\ \cdots\ \frac{\partial h_m}{\partial t}\right],\quad
\frac{\partial h}{\partial x_i} \equiv \left[\frac{\partial h_1}{\partial x_i}\ \frac{\partial h_2}{\partial x_i}\ \cdots\ \frac{\partial h_m}{\partial x_i}\right]
$$

$$
\frac{\partial^2 h}{\partial x_i \partial x_j} \equiv \left[\frac{\partial^2 h_1}{\partial x_i \partial x_j}\ \frac{\partial^2 h_2}{\partial x_i \partial x_j}\ \cdots\ \frac{\partial^2 h_m}{\partial x_i \partial x_j}\right]
$$

The fundamental problem of nonlinear filtering theory is how to solve the robust DMZ equation (5.26) or (5.27) in real time and in memoryless manner. In this section we shall describe the extended Yau–Yau algorithm which achieves this goal for any filtering system with arbitrary initial distribution.

Let $P_i = \{0 = \tau_0 < \tau_1 < \tau_2 \cdots < \tau_i = T\}$ be a partition of $[0,T]$. Let $\rho_i(x,t)$ be a solution of the following partial differential equation for $\tau_{i-1} \leqslant t \leqslant \tau_i$,

$$
(5.28)\quad
\begin{cases}
\dfrac{\partial \rho_i}{\partial t} + \dfrac{\partial}{\partial t}\left(h^T R^{-1}\right)^T y_{\tau_{i-1}}\rho_i \\[2mm]
\quad = \exp\left(-h^T R^{-1}y_{\tau_{i-1}}\right)\left(L - \dfrac{1}{2}h^T R^{-1}h\right)\left[\exp\left(h^T R^{-1}y_{\tau_{i-1}}\right)\rho_i\right] \\[2mm]
\rho_i(x,\tau_{i-1}) = \rho_{i-1}(x,\tau_{i-1})
\end{cases}
$$

Define the norm of the partition P_k by $|P_k| = \sup_{1 \leqslant i \leqslant k}\{|\tau_i - \tau_{i-1}|\}$. As $|P_k| \to 0$ we have

$$
(5.29)\qquad\qquad\qquad \rho(x,t) = \rho_i(x,t)
$$

where $\rho(x, t)$ is the solution of the robust DMZ equation (5.26) or (5.27). As shown in [Ya-Ya4] (See section 4 above), the solutionof (5.28) approximates the solution of robust DMZ equation (5.26) or (5.27) very well in both pointwise sense and L^2 sense for the model (2.1). For the most general (1.1), the proof of (5.29) can be done in a similar way, but it is substantially more difficult.

PROPOSITION 5.3. *For* $\tau_{i-1} \leqslant t \leqslant \tau_i$, $\rho_i(x, t)$ *satisfies* (5.28) *if and only if*

$$(5.30) \qquad u_i(x, t) = \exp\left[h^T(x, t)R^{-1}(t)y_{\tau_{i-1}}\right]\rho_i(x, t)$$

satisfies the following Kolmogorov-type equation

$$(5.31) \qquad \frac{\partial u_i}{\partial t}(x, t) = \left(L - \frac{1}{2}h^T R^{-1}h\right)u_i(x, t)$$

Observe (5.25), (5.29) and (5.30) for time interval $\tau_{i-1} \leqslant t \leqslant \tau_i$, at $t = \tau_{i-1}$, the left hand of time interval, we have

$$(5.32) \qquad \begin{cases} \sigma(x, \tau_{i-1}) = \exp\left[h^T(x, \tau_{i-1})R^{-1}(\tau_{i-1})y_{\tau_{i-1}}\right]\rho(x, \tau_{i-1}) \\ \rho(x, \tau_{i-1}) = \rho_i(x, \tau_{i-1}) \\ u_i(x, \tau_{i-1}) = \exp\left[h^T(x, \tau_{i-1})R^{-1}(\tau_{i-1})y_{\tau_{i-1}}\right]\rho(x, \tau_{i-1}) \end{cases}$$

$$(5.33) \qquad \Rightarrow u_i(x, \tau_{i-1}) = \sigma(x, \tau_{i-1})$$

at $t = \tau_i$, the right hand of the time interval, we have

$$(5.34) \qquad \begin{cases} \sigma(x, \tau_i) = \exp\left[h^T(x, \tau_i)R^{-1}(\tau_i)y_{\tau_i}\right]\rho(x, \tau_i) \\ \rho(x, \tau_i) = \rho_i(x, \tau_i) \\ u_i(x, \tau_i) = \exp\left[h^T(x, \tau_i)R^{-1}(\tau_i)y_{\tau_{i-1}}\right]\rho_i(x, \tau_i) \end{cases}$$

$$(5.35) \qquad \Rightarrow \sigma(x, \tau_i) = \exp\left[h^T(x, \tau_i)R^{-1}(\tau_i)\left(y_{\tau_i} - y_{\tau_{i-1}}\right)\right]u_i(x, \tau_i)$$

Our main idea is given as follows. First, to solve the DMZ equation (5.23) is equivalent to solve the robust DMZ equation (5.26) or (5.27). Second, the solutions of the robust DMZ equation (5.26) or (5.27) can be obtained approximately by solving (5.28) on each time interval $\tau_{i-1} \leqslant t \leqslant \tau_i$. Finally, to solve (5.28) is equivalent to the Kolmogorov-type equation (5.31) through *Proposition* 5.3.

On each time interval $\tau_{i-1} \leqslant t \leqslant \tau_i$, (5.33) can be treated as the initial condition of (5.31) $u_i(x, \tau_i)$ can be treated as the prediction of unnormalized density when the observation $y_{\tau_{i-1}}$ is given, but y_{τ_i} is not available. (5.35) on the one hand, is the update unnormalized density when the observation y_{τ_i} is available; on the other hand, it is the initial condition for next time step.

Proof. [**Conclusion:**] The Yan–Yau algorithm can be described as follows: for

$0 \leqslant t \leqslant \tau_1$, $u_1(x,t)$ can be calculated through

(5.36)
$$\begin{cases} \dfrac{\partial u_1}{dt}(x,t) = \left(L - \dfrac{1}{2} h^T R^{-1} h \right) u_1(x,t) \\ u_1(x,0) = \sigma_0(x) \end{cases}$$

(5.37)   and $\sigma(x,\tau_1) = \exp\left[h^T(x,\tau_1) R^{-1}(\tau_1) y_{\tau_1} \right] u_1(x,\tau_1)$

For $i \geqslant 2$ and $\tau_{i-1} \leqslant t \leqslant \tau_i$, $u_i(x,t)$ can be calculated through

(5.38)
$$\begin{cases} \dfrac{\partial u_i}{dt}(x,t) = \left(L - \dfrac{1}{2} h^T R^{-1} h \right) u_i(x,t) \\ u_i(x,\tau_{i-1}) = \sigma(x,\tau_{i-1}) \end{cases}$$

and $\sigma(x,\tau_i)$ is obtained by (5.35).

Observe that in our algorithm at time step i, we only need the observation at time τ_{i-1} and τ_i to calculate $\sigma(x,\tau_i)$. We do not need any other previous observation data. Observe also that the partial differential equation (5.31) is independent of observation y_t. It can be computed off-line. $\qquad\square$

REFERENCES

[Br] R. W. BROCKETT, *Nonlinear systems and nonlinear estimation theory*, The Mathematics of Filtering and Identification and Applications (M. Hazewinkel and J. C. Willems, eds.), Reidel, Dordrecht, 1981.

[Br-Cl] R. W. BROCKETT AND J. M. C. CLARK, *The geometry of the conditional density functions*, Analysis and Optimization of Stochastic Systems (O. L. R. Jacobs et al., eds.), Academic Press, New York, 1980, pp. 299–309.

[Bu-Pa] R. S. BUCY AND J. PAGES, *A prior error bounds for the cubic sensor problems*, IEEE Transactions on Automatic Control, 24(1979), pp. 948–953.

[Ch-Mi] M. CHALEYAT-MAUREL AND D MICHEL, *Des résultats de non existence de filtre de dimension finie*, Stochastics, 13(1984), pp. 83–102.

[Ch] J. CHEN, *On uniquity of Yau filters*, Proceedings of the American Control Conference (Baltimore, MD, USA), 1994, pp. 252–254.

[Ch-Ya1] J. CHEN AND S. S.-T. YAU, *Finite-dimensional filters with nonlinear drift VI: Linear structure of* Ω, Mathematics of Control, Signals and Systems, 9(1996), pp. 370–385.

[Ch-Ya2] ———, *Finite-dimensional filters with nonlinear drift VII: Mitter conjecture and structure of* η, SIAM Journal of Control and Optimization, 35(1997), pp. 1116–1131.

[C-Y-L1] J. CHEN, S. S.-T. YAU, AND C. W. LEUNG, *Finite-dimensional filters with nonlinear drift IV: Classification of finite-dimensional estimation algebras of maximal rank with state space dimension 3*, SIAM Journal of Control and Optimization, 34(1996), pp. 179–198.

[C-Y-L2] ———, *Finite-dimensional filters with nonlinear drift VII: Classification of finite-dimensional estimation algebras of maximal rank with state space dimension 4*, SIAM Journal of Control and Optimization, 35(1997), pp. 1132–1141.

[C-C-Y] W.-L. CHIOU, W.-R. CHIUEH, AND S. S.-T. YAU, *Structure theorem for five-dimensional estimation algebras*, Systems and Control Letters, 55(2006), pp. 275–281.

[Chi-Ya] W. L. CHIOU AND S. S.-T. YAU, *Finite-dimensional filters with nonlinear drift II: Brock-ett's problem on classification of finite-dimensional estimation algebras*, SIAM Journal of Control and Optimization, 32(1994), pp. 297–310.

[Da] M. H. A. DAVIS, *On a multiplicative functional transformation arising in nonlinear filtering theory*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 54(1980), pp. 125–139.

[Da-Ma] M. H. A. DAVIS AND S. I. MARCUS, *An introduction to nonlinear filter*, The Mathematics of Filtering and Identification and Applications (M. Hazewinkel and J. S. Willems, eds.), Reidel, Dordrecht, 1981, pp. 53–75.

[D-T-W-Y] R. T. DONG, L. F. TAM, W. S. WONG, AND S. S.S.-T. YAU, *Structure and classification theorems of finite-dimensional exact estimation algebras*, SIAM Journal of Control and Optimization, 29(1991), pp. 866–877.

[Du] T. E. DUNCAN, *Probability densities for diffusion processes with applications to nonlinear filtering theory*, Ph.D. thesis, Stanford, 1967.

[H-M-S] M. HAZEWINKEL, S. I. MARCUS, AND H. J. SUSSMANN, *Nonexistence of finite dimensional filters for conditional statistics of the cubic sensor problem*, Systems and Control Letters, 3(1983), pp. 331–340.

[Hu-Ya] G. Q. HU AND S. S.-T. YAU, *Finite-dimensional filters with nonlinear drift XV: New direct method for construction of universal finite-dimensional filter*, IEEE Transactions on Aerospace and Electronic Systems, 38:1(2002), pp. 50–57.

[H-Y-C] G. Q. HU, S. S.-T. YAU, AND W.-L. CHIOU, *Finite dimensional filters with nonlinear drift XIII: classification of finite-dimensional estimation algebras of maximal rank with state space dimension five*, Asian Journal of Mathematics, 4(2000), pp. 915–932.

[Ka] R. E. KALMAN, *A new approach to linear filtering and prediction problem*, Transactions on ASME, Series D, Journal of Basic Engineering, 82(1960), pp. 35–45.

[Ka-Bu] R. E. KALMAN AND R. S. BUCY, *New results in linear filtering and prediction theory*, Transactions on ASME, Series D, Journal of Basic Engineering, 83(1961), pp. 95–108.

[Ma] S. I. MARCUS, *Algebraic and geometric methods in nonlinear filtering*, SIAM J. Control Optim., 22(1984), pp. 817–844.

[Mi] S. K. MITTER, *On the analogy between mathematical problems of nonlinear filtering and quantum physics*, Ricerche Automatica, 10(1979), pp. 163–216.

[Mo] R. E. MORTENSEN, *Optimal control of continuous time stochastic systems*, Ph.D. thesis, University of California, Berkeley, CA, USA, 1996.

[Oc] D. OCONE, *Topics in nonlinear filtering theory*, Ph.D. thesis, Massachusetts Institute of Technology, MA, USA, 1980.

[Ra-Ya] A. RASOULIAN AND S. S.-T. YAU, *Finite-dimensional filters with nonlinear drift IX: Construction of finite-dimensional estimation algebra of non-maximal rank*, Systems and Control Letters, 30(1997), pp. 109–118.

[Ro] B. L. ROZOVSKY, *Stochastic partial differential equations arising in nonlinear filtering problems*, Usp. Mat. Nauk., 27(1972), pp. 213–214.

[T-W-Y] L. F. TAM, W. S. WONG, AND S. S.-T. YAU, *On a necessary and sufficient condition for finite dimensionality of estimation algebras*, SIAM Journal of Control and Optimization, 28(1990), pp. 173–185.

[We-No] J. WEI AND E. NORMAN, *On the global representation of the solutions of linear differential equations as a product of exponentials*, Proceedings of American Mathematical Science, 15(1964), pp. 327–334.

[Wo1] W. S. WONG, *Theorems on the structure of finite dimensional estimation algebras*, Systems and Control Letters, 9(1987), pp. 117–124.

[Wo2] _____, *On a new class of finite-dimensional estimation algebras*, Systems and Control Letters, 9(1987), pp. 79–83.

[Wo-Ya] W. S. WONG AND S. S.-T. YAU, *The estimation algebra of nonlinear filtering systems*, Mathematical Control Theory, Special Volume Dedicated to 60th Birthday of Brockett (J. Bailliaeul and J. C. Willems, eds.), Springer Verlag, (1998), 33–65.

[Wu-Ya] X. WU AND S. S.-T. YAU, *Classification of estimation algebras with state dimension 2*, SIAM Journal of Control and Optimization, 45:3(2006), pp. 1039–1073.

[W-Y-H] X. WU, S. S.-T. YAU, AND G. Q. HU, *Finite dimensional filters with nonlinear drift XII: linear and constant structure of Ω, stochastic theory and control*, Proceedings of the Workshop held in Lawrence, Kansas (B. Pasik-Duncan, ed.), Lecture Notes in Control and Information Sciences, no. 280, Springer Verlag, 2002, pp. 507–518.

[Yan-Ya] C. L. YAN AND S. S-T. YAU, *Real time solution of nonlinear filtering problem without memory III: Extended Yau's Algorithm*, submitted for publication.

[Ya1] S. S.-T. YAU, *Recent results on nonlinear filtering: new class of finite dimensional filters*, Proceedings of the 29th Conference on Decision and Control (Honolulu, Hawaii, USA), 1990, pp. 231–233.

[Ya2] ———, *Finite dimensional filters with nonlinear drift I: A class of filters including both Kalman–Bucy filters and Benes filters*, Journal of Mathematical Systems, Estimation and Control, 4(1994), pp. 181–203.

[Ya3] ———, *Brockett's problem on nonlinear filtering theory*, Lectures on Systems, Control, and Information, Studies in Advanced Mathematics, vol. 17, AMS/IP, 2000, pp. 177–212.

[Ya4] ———, *Complete classification of finite-dimensional estimation algebras of maximal rank*, International Journal of Control, 76:7(2003), pp. 657–677.

[Ya-Hu1] S. S.-T. YAU AND G.-Q. HU, *Direct method without Riccati equation for Kalman–Bucy filtering system with arbitrary initial conditions*, Proceedings of the 13th World Congress IFAC (San Francisco, CA), vol. H, June 30–July 5 1996, pp. 469–474.

[Ya-Hu2] ———, *Finite-dimensional filters with nonlinear drift X: Explicit solution of DMZ equation*, IEEE Transactions on Automatic Control, 46:1(2001), pp. 142–148.

[Ya-Hu3] ———, *Classification of finite-dimensional estimation algebras of maximal rank with arbitrary state-space dimension and Mitter conjecture*, International Journal of Control, 78:10(2005), pp. 689–705.

[Ya-Ra] S. S.-T. YAU AND A. RASOULION, *Classification of four-dimensional estimation algebras*, IEEE Transactions on Automatic Control, 44(1999), pp. 2312–2318.

[Y-W-W] S. S.-T. YAU, X. WU, AND W. S. WONG, *Hessian Matrix Non-decomposition Theorem*, Mathematical Research Letters, 6(1999), pp. 1–11.

[Ya-Ya1] S. S.-T. YAU AND S. T. YAU, *New direct method for Kalman–Bucy filtering system with arbitrary initial condition*, Proceedings of the 33rd conference on Decision and Control, Lake Buena Vista, FL, December 14–16 1994, pp. 1221–1225.

[Ya-Ya2] ———, *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 Transactions on Aerospace and Electronic Systems, 33(1997), pp. 1277–1294.

[Ya-Ya3] ———, *Real time solution of nonlinear filtering problem without memory I*, Mathematical Research Letters, 7(2000), pp. 671–693.

[Ya-Ya4] ———, *Real time solution of nonlinear filtering problem without memory II*, SIAM Journal of Control and Optimization, 47:1(2008), pp. 163–195.

[Za] M. ZAKAI, *On the optimal filtering of diffusion processes*, Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 11(1969), pp. 230–243.