Journal of Mathematical Systems, Estimation, and Control
Vol. 4, No. 2, 1994, pp. 181–203

Finite Dimensional Filters with Nonlinear Drift I: A class of filters including both Kalman-Bucy filters and Benes filters*

Stephen S.-T. Yau†

Abstract

Ever since the technique of the Kalman-Bucy filter was popularized, there has been an intense interest in finding new classes of finite dimensional recursive filters. In the late seventies, the concept of the estimation algebra of a filtering system was introduced. It was proven to be an invaluable tool in the study of nonlinear filtering problems. In 1981, Benes established exact finite dimensional filters for certain diffusions with nonlinear drift. In this paper, we established a simple algebraic necessary and sufficient condition for an estimation algebra of a general class of filtering systems to be finite dimensional. In particular we prove the Mitter conjecture for a large class of nonlinear filtering system. We also give partial solution to the Brockett problem on classification of finite dimensional estimation algebras. Using the Wei-Norman approach, we construct explicitly a general class of finite dimensional recursive filters.

Key words: nonlinear filters, Kalman-Bucy filters, Benes filters, Lie algebra, estimation algebra, Duncan-Mortensen-Zakai equation, Wei-Norman approach

AMS Subject Classifications: 17B30, 35K15, 60G35, 93E11

1 Introduction

The idea of using estimation algebras to construct finite dimensional nonlinear filters was first proposed in Brockett and Clark [2], Brockett [3]

*Received December 5, 1991; received in final form January 1, 1992.
†Research supported by Army Grant DAAL-03-89K-0123.

and Mitter [11]. The motivation came from the following Wei-Norman approach [15] of using Lie algebraic ideas to solve time variant linear differential equations. Consider the equation

$$\frac{d}{dt}X(t) = A(t)X(T) = \sum_{i=1}^{m} a_i(t)A_iX(t), \qquad X(0) = X_0$$

where X and A_i's are n by n matrices and a_i's are scalar-valued functions. Let $B_1, \cdots, B_n$ be a basis of the Lie algebra generated by $A_1, \cdots, A_m$. Then the Wei-Norman Theorem states that locally in t, $X(t)$ has a representation of the form

$$X(t) = \exp(b_i(t)B_1)\cdots\exp(b_n(t)B_n)X_0$$

where b_i's satisfy an ordinary differential equation of the form

$$\frac{db_i}{dt} = c_i(b_1, \cdots, b_n), \qquad b_i(0) = 0$$

for all i. The function c_i's in the above equation are determined by the structure constants of the Lie algebra generated by the A_i's.

The extension of Wei-Norman approach to the non-linear filtering problem is much more complicated. Instead of an ordinary differential equation, we have to solve the Duncan-Mortensen-Zakai (DMZ) equation, which is a stochastic partial differential equation. By working on the robust form of the DMZ equation we can reduce the complexity of the problem to that of solving a time variant partial differential equation. Working independently, Steinberg [13] applied the Wei-Norman approach to solve some partial differential equations which are roughly related to the linear filtering problem. Wong in [16] constructed some new finite dimensional estimation algebras and used the Wei-Norman approach to synthesize finite dimensional filters. However, the systems considered in [16] are very specific and the question whether the Wei-Norman approach works for a general system with finite dimensional estimation algebra remains open. The above introduction can be found in [14].

In [14], Tam, Wong and the present author have examined the properties of finite dimensional estimation algebras and the Wei-Norman approach in detail. There a class of filtering systems having the property that the drift-term, f, of the state evolution equation is a gradient vector field was considered. In [17], the concept of Ω is introduced, which is defined as the $n \times n$ matrix whose (i,j)-entry is $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}$. In view of Poincare lemma, f is a gradient vector field if and only if $\Omega = 0$. In this paper, we consider a more general class of filtering systems having the property that $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}$. are constants for all i, j, i.e. Ω is a skew symmetric constant matrix. In Section 2, we prove that Ω is a skew symmetric constant matrix if and only if $(f_1, \cdots, f_n) = (l_1, \cdots, l_n) + (\frac{\partial \psi}{\partial x_1}, \cdots, \frac{\partial \psi}{\partial x_n})$ where $l_1, \cdots, l_n$ are polynomials of degree one and ψ is a C^∞ function. If $\psi \equiv 0$ on $\mathbf{R}^n$, then we are in the situation of the Kalman-Bucy filtering system. If $(l_1, \cdots, l_n) \equiv 0$, then we have the Benes' filtering system [1] as special case.

Motivated by the results of Tam-Wong-Yau [14], we investigate the algebraic problem of characterizing and classifying finite dimensional estimation algebras. In this paper, we derive a simple necessary and sufficient condition for an estimation algebra of the above filtering system to be finite dimensional. As an important consequence of these algebraic results, we also prove that for a system of the above type with finite dimensional estimation algebra, the Wei-Norman approach always leads to finite dimensional filters. As a matter of fact, we are able to write down these finite dimensional filters explicitly.

It is a big problem in nonlinear filtering theory whether the observation terms h_i's in the filtering system (2.0) are necessarily affine if one has finite dimensional filter. In fact Professor Sanjoy Mitter conjectured that this is always the case. Previous known techniques do not fully explain why one has to assume h_i's be affine for Kalman-Bucy and Benes filters except for the filtering systems whose dynamics are governed by gradient vector fields which was studied by [14]. Using the Lie algebraic approach, we first established in Theorem 3 that for a quite general filter, h_i's are necessarily affine. In his talk at the International Congress of Mathematics in 1983, Professor Roger Brockett proposed to find a classification of finite dimensional estimation algebras. Our results in Theorem 5 and Theorem 6 give an important step toward a complete solution of his problem.

Although the key ideas behind the proof of the main theorems are similar to [14], the main results here are quite fundamental. This is because in the second paper (jointly with Wen-Lin Chiou) we have shown that from a Lie algebraic point of view the filters constructed here are the most general maximal rank dimensional filters when the dimension of the state space is two.

2 Basic Concepts

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

$$\begin{cases} dx(t) = f(x(t)) + g(x(t))dv(t) & x(0) = x_0 \\ dy(t) = h(x(t))dt + dw(t) & y(0) = 0 \end{cases} \tag{2.0}$$

in which x, v, y, and w are, respectively, $\mathbf{R}^n, \mathbf{R}^p, \mathbf{R}^m$, and $\mathbf{R}^m$ valued processes,
and v and w have components which are independent, standard Brownian processes. We further assume that $n = p, f, h$ are C^∞ smooth, and that g

is an orthogonal matrix. We will refer to $x(t)$ as the state of the system at time t 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 (see [7], for example) that $\rho(t,x)$ is given by normalizing a function, $\sigma(t,x)$, which satisfies the following Duncan-Mortensen-Zakai equation:

$$d\sigma(t,x) = L_0\sigma(t,x)dt + \sum_{i=1}^{m} L_i\sigma(t,x)dy_i(t), \quad \sigma(0,x) = \sigma_0 \tag{2.1}$$

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

Equation (2.1) is a stochastic partial differential equation. In real applications, we are interested in considering robust state estimators from observed sample paths with some properties of robustness. Davis in [6] studied this problem and proposed some robust algorithms. In our case, his basic idea reduces to defining a new unnormalized density

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

It is easy to show that $\xi(t,x)$ satisfies the following time varying partial differential equation

$$\begin{aligned}\frac{d\xi(t,x)}{dt} =& L_0\xi(t,x) + \sum_{i=1}^{m} y_i(t)[L_0, L_i]\xi(t,x)\\ &+ \frac{1}{2}\sum_{i,j=1}^{m} y_i(t)y_j(t)[[L_0, L_i], L_j]\xi(t,x)\\ \xi(0,x) =& \sigma_0\end{aligned} \tag{2.2}$$

where $[\cdot,\cdot]$ is the Lie bracket defined as :

Definition. If X and Y are differential operators, the Lie bracket of X and Y, $[X,Y]$, is defined by

$$[X,Y]\zeta = X(Y\zeta) - Y(X\zeta)$$

for any C^∞ function ζ.

The objective of constructing a robust finite-dimensional filter to equation (2.0) is equivalent to finding a smooth manifold, M, and complete C^∞ vector fields, μ_i, on M and C^∞ functions, ν, on $M \times \mathbf{R} \times \mathbf{R}^n$ and w_i's on $\mathbf{R}^m$, such that $\xi(t, x)$ can be represented in the form:

$$\frac{dz(t)}{dt} = \sum_{i=1}^{k} \mu_i(z(t)) w_i(y(t)), \qquad z(0) \in M \tag{2.3.a}$$

$$\xi(t, x) = \nu(z(t), t, x). \tag{2.3.b}$$

Following [4], we say that system (2.0) has a robust universal finite-dimensional filter if for all initial probability density, σ_0, there exists a z_0, such that (2.3.a) and (2.3.b) hold if $z(0) = z_0$, and μ_i, ν, w_i are independent of σ_0.

In §4, we will use the Wei-Norman approach to construct a finite dimensional filter for equation (2.0). Before we can achieve that, we need to introduce the concept of the estimation algebra of equation (2.0) and examine its algebraic structure.

Definition. Thr estimation algebra E of a filtering problem (2.0), is defined to be the Lie algebra generated by

$$\{L_0, L_1, \cdots, L_m\}, \quad \text{or} \quad E = \langle L_0, L_1, \cdots, L_m \rangle_{L.A.}$$

Theorem 1. $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ *are constants for all* i *and* j *if and only if*

$$(f_1, \cdots, f_n) = (l_1, \cdots, l_n) + \left(\frac{\partial \psi}{\partial x_1}, \cdots, \frac{\partial \psi}{\partial x_n}\right)$$

where $l_1, \cdots, l_n$ *are polynomials of degree one and* ψ *is a* C^∞ *function.*

Proof: Suppose first that

$$(f_1, \cdots, f_n) = (l_1, \cdots, l_n) + \left(\frac{\partial \psi}{\partial x_1}, \cdots, \frac{\partial \psi}{\partial x_n}\right)$$

where $l_i(x) = a_1 + b_{i1}x_1 + b_{i2}x_2 + \cdots + b_{in}x_n$, for $1 \leq i \leq n$. Here $a_i, b_{i1}, b_{i2}, \cdots, b_{in}$ are constants. Then

$$\begin{aligned} \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} &= \left(b_{ji} + \frac{\partial^2 \psi}{\partial x_i \partial x_j}\right) - \left(b_{ji} + \frac{\partial^2 \psi}{\partial x_j \partial x_i}\right) \\ &= b_{ji} - b_{ij} = \text{constant} \end{aligned}$$

for all $1 \leq i, j \leq n$.

Conversely suppose that $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ are constants for all $1 \leq i, j \leq n$. Observe that $c_{ij} = -c_{ji}$. Let $b_{ij} = -\frac{1}{2}c_{ij}$. Then we have

$$b_{ji} - b_{ij} = -\frac{1}{2}c_{ji} - \left(-\frac{1}{2}c_{ij}\right) = \frac{1}{2}c_{ij} + \frac{1}{2}c_{ij} = c_{ij} \tag{2.4}$$

for all $1 \leq i, j \leq n$. Let $l_i(x) = \sum_{j=1}^{n} b_{ij}x_j$ for $1 \leq i \leq n$. Clearly the exterior derivatives of the differential forms $f_1 dx_1 + f_2 dx_2 + \cdots + f_n dx_n$ and $l_1(x)dx_1 + l_2(x)dx_2 + \cdots + l_n(x)dx_n$ are given as follows:

$$\begin{aligned} d\left(\sum_{j=1}^{n} f_j\, dx_j\right) &= \sum_{i,j} \frac{\partial f_j}{\partial x_i} dx_i \wedge dx_j \\ &= \sum_{i<j} \left(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}\right) dx_i \wedge dx_j \\ &= \sum_{i<j} c_{ij}\, dx_i \wedge dx_j \end{aligned}$$

and

$$\begin{aligned} d\left(\sum_{j=1}^{n} l_j\, dx_j\right) &= \sum_{i,j} b_{ij}\, dx_i \wedge dx_j \\ &= \sum_{i<j} (b_{ji} - b_{ij}\, dx_i) \wedge dx_j. \end{aligned}$$

In view of (2.4), we have

$$d\left(\sum_{j=1}^{n} f_j\, dx_j\right) = d\left(\sum_{j=1}^{n} l_j\, dx_j\right),$$

i.e.,

$$d\left(\sum_{j=1}^{n} f_j\, dx_j - \sum_{j=1}^{n} l_j\, dx_j\right) = 0.$$

Since every d-closed differential form on $\mathbf{R}^n$ are d-exact, there exists a C^∞ ψ such that

$$\sum_{j=1}^{n} f_j\, dx_j - \sum_{j=1}^{n} l_j\, dx_j = d\psi = \sum_{j=1}^{n} \frac{\partial \psi}{\partial x_j} dx_j.$$

Here Theorem 1 follows immediately. QED

From now on, unless stated otherwise, we assume the estimation algebra of (2.0) satisfies the conditions stated in Theorem 1.

Define

$$D_i = \frac{\partial}{\partial x_i} - f_i$$

and

$$\eta = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{m} h_i^2. \tag{2.5}$$

Then,

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

We need the following basic results for later discussion.

Theorem 2. (Ocone) *Let E be a finite dimensional estimation algebra. If a function ξ is in E, then ξ is a polynomial of degree ≤ 2.*

Ocone's theorem ([10], see [5] for an extension) says that $h_1, \cdots, h_m$ in a finite dimensional estimation algebra are polynomials of degree ≤ 2.

Lemma 1. ([14]) *Let ξ be a C^∞ function on $\mathbf{R}^n$. Suppose $E_l(\xi)$ is a polynomial of degree at most k where $E_l = \sum_{i=1}^{l} x_i \frac{\partial}{\partial x_i}$. Then*

$$\xi = p_k(x_1, \cdots, x_n) + \xi(0, \cdots, 0, x_{l+1}, \cdots, x_n)$$

where p_k is a polynomial of degree k in $x_1, \cdots, x_n$.

Lemma 2. ([14]) *Let ξ be a C^∞ function on $\mathbf{R}^n$. Suppose that $E_l(\xi)+2\xi$ is a sum of polynomials of degree two and a C^∞ function on $\mathbf{R}^n$ which depends only on $x_{l+1}, \cdots, x_n$ variables. Then for any*

$$(a_{l+1}, \cdots, a_n) \in \mathbf{R}^{n-l}, \quad \xi(x_1, \cdots, x_l, x_{l+1}, \cdots, a_n)$$

is a polynomial of degree two in $x_1, \cdots, x_l$ variables.

Lemma 3. *Let xi be a C^∞ function on $\mathbf{R}^n$. Suppose that for any $a_{l+1}, \cdots, a_n) \in \mathbf{R}^{n-l}$, $E_l(\xi) - \xi$ is a polynomial of degree two in $x_1, \cdots, x_l$ variables. Then for any $(a_{l+1}, \cdots, a_n) \in \mathbf{R}^n$, ξ is also a polynomial of degree two in $x_1, \cdots, x_l$ variables.*

Proof: The proof is similar to the proof of Lemma 2.

3 Structure Theorems

We first begin with some elementary lemmas.

Lemma 4. *(i) $[XY, Z] = X[Y, Z] + [X, Z]Y$ where X, Y and Z are differential operators.*

(ii) $[gD_i, h] = g\frac{\partial h}{\partial x_i}$, *where* $D_i = \frac{\partial}{\partial x_i} - f_i$, g *and* h *are functions defined on* $\mathbf{R}^n$.

(iii) $[gD_i, hD_j] = -gh\omega_{ij} + g\frac{\partial h}{\partial x_i}D_j - h\frac{\partial g}{\partial x_j}D_i$, *where* $\omega_{ji} = [D_i, D_j] = \frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i}$.

(iv) $[gD_i^2, h] = 2g\frac{\partial h}{\partial x_i}D_i + g\frac{\partial^2}{\partial x_i^2}$.

(v) $[D_i^2, hD_j] = 2\frac{\partial h}{\partial x_i}D_iD_j - 2h\omega_{ij}D_i + \frac{\partial^2 h}{\partial x_i^2}D_j - h\frac{\partial \omega_{ij}}{\partial x_i}$.

Proof: (i) This follows immediately from the definition of the Lie bracket.

(ii) $[gD_i, h] = [g, h]D_i + g[D_i, h] = g\frac{\partial h}{\partial x_i}$

(iii)
$$\begin{aligned}[gD_i, hD_j] &= g[D_i, hD_j] + [g, hD_j]D_i \\ &= -g[hD_j, D_i] - [hD_j, g]D_i \\ &= -g(h[D_j, D_i] + [h, D_i]D_j) - (h[D_j, g])D_i \\ &= -gh\omega_{ij} + g\frac{\partial h}{\partial x_i} - h\frac{\partial g}{\partial x_j}D_i\end{aligned}$$

(iv)
$$\begin{aligned}[gD_i^2, h] &= g[D_i^2, h] + [g, h]D_i^2 \\ &= gD_i[D_i, h] + g[D_i, h]D_i \\ &= gD_i\left(\frac{\partial h}{\partial x_i}\right) + g\frac{\partial h}{\partial x_i}D_i \\ &= g\frac{\partial^2 h}{\partial x_i^2} + g\frac{\partial h}{\partial x_i}D_i + g\frac{\partial h}{\partial x_i}D_i \\ &= g\frac{\partial^2 h}{\partial x_i^2} + 2g\frac{\partial h}{\partial x_i}D_i\end{aligned}$$

(v)
$$\begin{aligned}[D_i^2, hD_j] &= -[hD_j, D_i^2] \\ &= -h[D_j, D_i^2] - [h, D_i^2]D_j \\ &= 2h\omega_{ji}D_i + h\frac{\partial \omega_{ji}}{\partial x_i} + \left(\frac{\partial^2 h}{\partial x_i^2} + 2\frac{\partial h}{\partial x_i}D_i\right)D_j.\end{aligned}$$

QED

The following theorem plays a fundamental role in the classification of estimation algebra.

Theorem 3. *Let* E *be a finite dimensional estimation algebra of (2.0) satisfying* $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ *where* c_{ij} *are constants for all* $1 \le i, j \le n$. *Then* $h_1, \cdots, h_m$ *are polynomials of degree at most one.*

Proof: By Theorem 2, each h_i is a polynomial of degree at most two. Suppose h_1 is of degree two, then by using the affine transformation $\tilde{x} = Ax + b$, where A is orthogonal, we may assume that h_1 is of the form

$$\sum_{i=1}^{l} c_i\tilde{x}_i^2 + \sum_{i=l+1} c_i\tilde{x}_i^2 + c_0$$

where $c_1, \cdots, c_l$ are nonzero real numbers, and $l \le n$. (If $l = n$, the second summation vanishes.) Define $\tilde{f}(\tilde{x}) = Af(x)$ and $\tilde{D}_i = \frac{\partial}{\partial \tilde{x}_i} -$

$\tilde{f}_i$. If $\tilde{\phi}(\tilde{x}) = \phi(x)$ and $(\tilde{l}_1(\tilde{x}), \cdots, \tilde{l}_n(\tilde{x}))^T = A(l_1(x), \cdots, l_n(x))^T$ where $(l_1(x), \cdots, l_n(x))^T$ is as in Theorem 1, then it is easy to see that

$$\tilde{f}(\tilde{x}) = \left(\frac{\partial \tilde{\phi}}{\partial \tilde{x}_1}(\tilde{x}), \cdots, \frac{\partial \tilde{\phi}}{\partial \tilde{x}_n}(\tilde{x})\right)^T + (\tilde{l}_1(\tilde{x}), \cdots, \tilde{l}_n(\tilde{x})).$$

Under the transformation, L_0 is mapped into:

$$\tilde{L}_0 = \frac{1}{2}\left(\sum_{i=1}^{n} \tilde{D}_i^2 - \tilde{\eta}(\tilde{x})\right)$$

where

$$\tilde{\eta}(\tilde{x}) = \sum_{i=1}^{n} \frac{\partial \tilde{f}_i}{\partial \tilde{x}_i}(\tilde{x}) + \tilde{f}(\tilde{x})^T \tilde{f}(\tilde{x}) + \sum_{i=1}^{n} \tilde{h}_i(\tilde{x})^2$$

and h is transformed into

$$\tilde{h}(\tilde{x}) = h(x).$$

E is isomorphic to the Lie algebra generated by $\tilde{L}_0$ and $\tilde{h}_i$. Note that the degree of h_i in x is the same as the degree of $\tilde{h}_i$ in $\tilde{x}$. Without causing any confusion, from now on, we drop the tilde notation.

Since h_1 is not of degree one, then $l \geq 1$. We shall produce a contradiction. Let $X_0 = h_1$, and define X_i for $i \geq 1$ recursively by $X_i = [[L_0, X_{i-1}], X_0]$. Since $L_0 = \frac{1}{2}\left(\sum_{i=1}^{n} D_i^2 - \eta\right)$, it is easy to see that

$$\begin{aligned}
X_1 &= 4\sum_{i=1}^{l} c_i^2 x_i^2 + \sum_{i=l+1}^{n} c_i^2 \\
X_2 &= 4^2 \sum_{i=1}^{l} c_i^3 x_i^2 \\
X_3 &= 4^3 \sum_{i=1}^{l} c_i^4 x_i^2 \\
&\vdots \\
X_j &= 4^j \sum_{i=1}^{l} x_i^2 c_i^{j+1}.
\end{aligned}$$

By the invertibility of the Vandermonde matrix, it follows after some relabeling, if necessary, that

$$p := \frac{1}{2}\sum_{i=1}^{l} x_i^2$$

is an element in E. Let Y_0 be the zero degree differential operator defined by multiplication by p. Define

$$\begin{aligned} Y_1 : &= [L_0, Y_o] \\ &= \sum_{i=1}^{l} x_i D_i + \frac{l}{2} \\ Y_2 : &= [L_0, Y_1] \\ &= \sum_{i=1}^{l} D_i^2 - \sum_{i=1}^{n} \sum_{j=1}^{l} x_j c_{ij} D_i + \frac{1}{2} \sum_{i=1}^{l} x_i \frac{\partial \eta}{\partial x_i}. \end{aligned}$$

Let $\alpha_i = \sum_{j=1}^{l} x_j c_{ij}$ and $E_l = \sum_{i=1}^{l} x_i \frac{\partial}{\partial x_i}$. Then

$$\begin{aligned} Y_2 &= \sum_{i=1}^{l} D_i^2 - \sum_{i=1}^{n} \alpha_i D_i + \frac{1}{2} E_l(\eta) \\ Y_3 &= [Y_2, Y_1] \\ &= 2 \sum_{i=1}^{l} D_i^2 - 2 \sum_{i=1}^{l} \alpha_i D_i + \sum_{i=l+1}^{n} \alpha_i D_i + \sum_{i=1}^{n} \alpha_i^2 - \frac{1}{2} E_l^2(\eta) \\ Y_4 &= [Y_3, Y_1] \\ &= 4 \sum_{i=1}^{l} D_i^2 - 4 \sum_{i=1}^{l} \alpha_i D_i - \sum_{i=l+1}^{n} \alpha_i D_i + \frac{1}{2} E_l^3(\eta) - 3 \sum_{i=l+1}^{n} \alpha_i^2. \end{aligned}$$

Then

$$\begin{aligned} Y_3 - 2Y_2 &= 3 \sum_{i=l+1}^{n} \alpha_i D_i + \sum_{i=1}^{n} \alpha_i^2 - \frac{1}{2} E_l^2(\eta) - E_l(\eta) \\ Y_4 - 2Y_3 &= -3 \sum_{i=l+1}^{n} \alpha_i D_i - 5 \sum_{i=l+1}^{n} \alpha_i^2 + \frac{1}{2} E_l^3(\eta) + E_l^2(\eta) - 2 \sum_{i=1}^{l} \alpha_i^2 \end{aligned}$$

and

$$(Y_4 - 2Y_3) + (Y_3 - 2Y_2) = -4 \sum_{i=l+1}^{n} \alpha_i^2 + \frac{1}{2} E_l [E_l^2(\eta) + E_l(\eta) - 2\eta] - \sum_{i=1}^{l} \alpha_i^2.$$

By Theorem 2, $(Y_4 - 2Y_3) + (Y_3 - 2Y_2)$ is a polynomial of degree at most two. In view of Lemma 1, $E_l^2(\eta) + E_l(\eta) - 2\eta$ is a sum of polynomials of degree two and a C^∞ function which depends on $x_{l+1}, \cdots, x_n$ variables. Since

$E_l^2(\eta)+E_l(\eta)-2\eta = (E_l+2)(E_l-1)\eta$, by Lemma 2, for any $(a_{l+1}, \cdots, a_n) \in \mathbf{R}^{n-1}$, $E_l(\eta)-\eta$ is a polynomial of degree two in $x_1, \cdots, x_l$. It follows from Lemma 3 that for any $(a_{l+1}, \cdots, a_n) \in \mathbf{R}^{n-l}$, η is a polynomial of degree two in $x_1, \cdots, x_l$. Recall that

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = -\sum_{i=1}^{m} h_i^2 + \eta. \tag{3.0}$$

Let $\psi \in C_0^\infty$ be any C^∞ function with compact support. Multiply equation (3.0) with ψ^2 and integrate the equation over $\mathbf{R}^n$.

$$\int_{\mathbf{R}^n} \left(\eta - \sum_{i=1}^{m} h_i^2\right)\psi^2 = \int_{\mathbf{R}^n} \psi^2 \mathrm{div} f + \int_{\mathbf{R}^n} \psi^2 (f \cdot f)$$

where $f = (f_1, \cdots, f_n)$ and div $f = \sum_{i=1}^m \frac{\partial f_i}{\partial x_i}$. In view of divergence theorem, we have

$$\begin{aligned} \int_{\mathbf{R}^n} \left(\eta - \sum_{i=1}^{m} h_i^2\right)\psi^2 &= -\int_{\mathbf{R}^n} 2\psi \nabla\psi \cdot f + \int_{\mathbf{R}^n} \psi^2 (f \cdot f) \\ &= -2\int_{\mathbf{R}^n} \nabla\psi \cdot \psi f + \int_{\mathbf{R}^n} \psi^2 (f \cdot f). \end{aligned} \tag{3.1}$$

By the Schwartz inequality,

$$\begin{aligned} 2\int_{\mathbf{R}^n} \nabla\psi \cdot \psi f &\le 2\left[\int_{\mathbf{R}^n} |\nabla\psi|^2\right]^{1/2} \left[\int_{\mathbf{R}^n} \psi^2 (f \cdot f)\right]^{1/2} \\ &\le \int_{\mathbf{R}^n} |\nabla\psi|^2 + \int_{\mathbf{R}^n} \psi^2 (f \cdot f). \end{aligned} \tag{3.2}$$

Putting (3.2) into (3.1), we get

$$\int_{\mathbf{R}^n} |\nabla\psi|^2 - \int_{\mathbf{R}^n} \left(\sum_{i=1}^{m} h_i^2 - \eta\right)\psi^2 \ge 0 \tag{3.3}$$

which is true for all $\psi \in C_0^\infty$. Take any nonzero C^∞ function θ with compact support. Define ψ to be θ followed by a translation in $x_1, \cdots, x_l$ variables direction. Observe that $\int_{\mathbf{R}^n} |\nabla\psi|^2$ is independent of the translation selected. On the other hand, since η is quadratic in $x_1, \cdots, x_l, \sum_{i=1}^m h_i^2 - \eta$ becomes very positive when one of the $x_1, \cdots, x_l$ tends to infinity while the other variables remain fixed. We can choose translation in directions $x_1, \cdots, x_l$ in such a way that

$$\int_{\mathbf{R}^n} \left(\sum_{i=1}^{m} h_i^2 - \eta \right) \psi^2$$

is arbitrarily large while $\int_{\mathbf{R}^n} |\nabla \psi|^2$ is bounded. This of course contradicts the inequality (3.3). QED

The argument above actually proves the following theorem.

Theorem 4. *Let $F(x_1, \cdots, x_n)$ be a C^∞ function on $\mathbf{R}^n$. Suppose that there exists a path $c : \mathbf{R} \longrightarrow \mathbf{R}^n$ and $\delta > 0$ such that $\lim_{t\to\infty} ||c(t)|| = \infty$ and $\lim_{t\to\infty} \sum_{B_\delta(c(t))} F = -\infty$, where $B_\delta(c(t)) = \{x \in \mathbf{R}^n : ||x - c(t)|| < \delta\}$. Then there are no C^∞ functions $f_1, f_2, \cdots, f_n$ on $\mathbf{R}^n$ satisfying the equation*

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F.$$

Corollary. *Let $F(x_1, \cdots, x_n)$ be a polynomial on $\mathbf{R}^n$. Suppose that there exists a polynomial path $c : \mathbf{R} \longrightarrow \mathbf{R}^n$ such that $\lim_{t\to\infty} ||c(t)|| = \infty$ and $\lim_{t\to\infty} F \circ c(t) = -\infty$. Then there are no C^∞ functions $f_1, \cdots, f_n$ on $\mathbf{R}^n$ satisfying the equation*

$$\sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F.$$

Proof: The proof of this corollary is the same as the corresponding corollary in [14]. QED

Definition. Let E be an estimation algebra of (2.0) satisfying $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ where c_{ij} are constants for all $1 \leq i, j \leq n$. If E is finite dimensional, then the matrix

$$H = [\nabla h_1, \nabla h_2, \cdots, \nabla h_m] \tag{3.4}$$

is a constant matrix in view of Theorem 3. H is called the observation matrix of (2.0).

The following result provides a simple characterization of when the dimension of an estimation algebra is finite.

Theorem 5. *Let E be an estimation algebra of (2.0) satisfying $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ where c_{ij} are constants for all $1 \leq i, j \leq n$.*

(i) If η is a polynomial of degree at most two, then E is finite dimensional and has a basis consisting of $E_0 = L_0$, differential operators $E_1, \cdots, E_p$ (for some p) of the form

$$\sum_{j=1}^{n} \alpha_{ij} D_j + \beta_i, \qquad 1 \le i \le p$$

where α_{ij}'s are constants and β_i's are affine in x, and zero degree differential operators $E_{p+1}, \cdots, E_q, 1$ (for some $q > p$) where E_i's are affine in x for $p+1 \le i \le q$. Moreover the quadratic part of $\eta - \sum_{i=1}^{m} h_i^2$ is positive semi-definite.

(ii) Conversely, if E is finite dimensional, then $h_1, \cdots, h_m$ are affine in x, i.e. the observation matrix is a constant matrix. Furthermore if the observation matrix has rank n (in particular $m \ge n$), then η is a polynomial of degree at most two.

Proof: (ii) If E is finite dimensional, then $h_1, \cdots, h_m$ are affine in x by Theorem 3. Since the observation H is a constant matrix with rank n, there are constants c_i's such that $x_i + c_i$ is in E for $i = 1, \cdot, n$. In view of Lemma 4, we have

$$[L_0, x_i + c_i] = \frac{1}{2}\left[\sum_{j=1}^{n} D_j^2, x_i + c_i\right] = D_i \tag{3.5}$$

$$[D_i, x_j + c_j] = \delta_{ij} \tag{3.6}$$

$$\begin{aligned} [L_0, D_i] &= \frac{1}{2}\left[\sum_{j=1}^{n} D_j^2, D_i\right] + \frac{1}{2}[D_i, \eta] \\ &= \sum_{j=1}^{n} c_{ij} D_j + \frac{1}{2}\frac{\partial \eta}{\partial x_i} \end{aligned} \tag{3.7}$$

(3.5) and (3.7) imply that $\frac{\partial \eta}{\partial x_i}$ is in E for all $1 \le i \le n$. If η is a quadratic polynomial, then in view of (3.5), (3.6) and (3.7), we see easily that E is a finite dimensional real vector space spanned by $1, x_1, \cdot, x_n, D_1, \cdots, D_n$ and L_0.

To see that η is a quadratic polynomial, we first observe that by Theorem 2, $\frac{\partial \eta}{\partial x_i}$, for all $1 \le i \le n$, are polynomials of degree at most two. It follows from lemma 1 that η is a polynomial of degree at most three. If $\frac{\partial \eta}{\partial x_i}$ is a polynomial of degree at most one for all $1 \le i \le n$, then clearly η is a polynomial of degree at most two and we are done. Assume that $\frac{\partial \eta}{\partial x_i}$

is a polynomial of degree exactly two for some i; we shall prove that this is impossible. Without loss of generality, we may assume that the degree of $\frac{\partial\eta}{\partial x_1}$ is exactly two. So,

$$\eta = x_1 q + r$$

where q is a polynomial with degree two, r is independent of x_1. Depending on the degree of q in x_1, we have three possible cases.

(i) **Degree 2 Case.** Clearly $\eta - \sum_{i=1}^m h_i^2$ can be arbitrarily negative on some polynomial path as the path tends to infinity.

(ii) **Degree 1 Case.** It follows that $\eta = \sum_{i=2}^n \alpha_i x_i x_1^2 + \beta x_1 + r$ where α_i's are constants, at least one of them nonzero, β and r are independent of x_1. Clearly, $\eta - \sum_{i=1}^m h_i^2$ can be arbitrarily negative on some polynomial path as the path tends to infinity.

(iii) **Degree 0 Case.** Since q is independent of x_1, $\eta = sx_1 + t$, where s and t are independent of x_1. s is a polynomial of degree exactly two while t is a polynomial of degree at most three. If $\sum_{i=1}^m h_i^2$ is independent of x_1, then $\eta - \sum_{i=1}^m h_i^2$ can be arbitrarily negative. If $\sum_{i=1}^m h_i^2$ is dependent on x_1, it must be of degree 2 in x_1. Again $\eta - \sum_{i=1}^m h_i^2$ can be arbitrarily negative on some polynomial path as the path tends to infinity.

In all three cases, there is a contradiction to the corollary of Theorem 4. We have proved that $\frac{\partial\eta}{\partial x_i}$ is a polynomial of degree at most one, for all $1 \le i \le n$. Hence η is a polynomial of degree at most two. This completes the proof of statement (ii) in Theorem 5.

To prove state (i) in Theorem 5, let F be the real vector space spanned by $\sum_{i=1}^n \alpha_i D_i + \beta$ where α_i's are constants and β is affine in x. Clearly the dimension of F is $1 + 2n$. If $\sum_{i=1}^n \alpha_i D_i + \beta$ is an element in F, then

$$\begin{aligned}
&\left[L_0, \sum_{i=1}^n \alpha_i D_i + \beta\right] \\
&= \frac{1}{2}\left[\sum_{j=1}^n D_j^2, \sum_{i=1}^n \alpha_i D_i\right] + \frac{1}{2}\left[\sum_{i=1}^n \alpha_i D_i, \eta\right] + \frac{1}{2}\left[\sum_{j=1}^n D_j^2, \beta\right] \\
&= \frac{1}{2}\sum_{i,j=1}^n \alpha_i [D_j^2, D_i] + \frac{1}{2}\sum_{i=1}^n \alpha_i \frac{\partial\eta}{\partial x_i} + \frac{1}{2}\left[\sum_{i=1}^n D_j^2, \sum_{i=1}^n \beta_i x_i + \beta_0\right] \\
&= \sum_{i,j=1}^n \alpha_i c_{ij} D_j + \frac{1}{2}\sum_{i=1}^n \alpha_i \frac{\partial\eta}{\partial x_i} + \sum_{i=1}^n \beta_i D_i \in F \qquad (3.8)
\end{aligned}$$

where $\beta = \sum_{i=1}^n \beta_i x_i + \beta_0, \beta_0, \beta_1, \cdots, \beta_n$ are constants. If $\sum_{i=1}^n \alpha_i' D_i + \beta'$ is another element in F, then

$$\left[\sum_{i=1}^n \alpha_i D_i + \beta, \sum_{j=1}^n \alpha_j' D_j + \beta'\right]$$

$$= \sum_{i,j=1}^{n} \alpha_i \alpha_j' [D_i, D_j] + \sum_{i=1}^{n} \alpha_i [D_i, \beta'] - \sum_{j=1}^{n} \alpha_j' [D_j, \beta]$$
$$= \sum_{i,j=1}^{n} \alpha_i \alpha_i' c_{ji} + \sum_{i=1}^{n} \alpha_i \frac{\partial \beta'}{\partial x_i} - \sum_{j=1}^{n} \alpha_j' \frac{\partial \beta}{\partial x_j}$$
$$= \text{constant}. \tag{3.9}$$

By (3.8) and (3.9), we know that the Lie algebra $\bar{F}$ generated by L_0 and F is exactly the vector space spanned by L_0 and F. In particular $\bar{F}$ has dimension $2n+2$. Since $\langle h_1, \because, h_m \rangle$ is contained in F by Theorem 4 and the assumption that η is a polynomial of degree 2, we have $E \subseteq \bar{F}$. So E is finite dimensional with basis as claimed in the statement (i) provided we can show that 1 is in E. For this purpose we only need to show that a nonzero constant is in E. Let $h_i = \sum_{k=1}^{n} h_{ik} x_k + e_i$ where h_{ik} and e_i are constants. Choose any h_i such that $h_{ik} \neq 0$ for some k. Then

$$[[L_0, h_i], h_j] = \left[\sum_{i=1}^{n} h_{ik} D_k, \sum_{l=1}^{n} h_{il} x_l + e_l \right] = \sum_{k=1}^{n} (h_{ik})^2$$

is a nonzero constant in E. QED

The following theorem follows from the proof of Theorem 5.

Theorem 6. *Let E be an estimation algebra of (2.0) satisfying $\frac{\partial f_i}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ where c_{ij} are constants for all $1 \leq i, j \leq n$. Suppose $m \geq n$ and the observation matrix is a constant matrix with full rank. If E is finite dimensional, then it is of dimension $2n+2$ with basis given by $1, x_1, \cdots, x_n, D_1, \cdots, D_n$ and L_0.*

4 Finite Dimensional Filters

In this section we will use the structural results of previous sections to derive finite-dimensional filters by the Wei-Norman approach.

Definition. Suppose X is a differential operator, ρ_0 is the domain of X, r is a continuous function, and $R(t) = \int_0^t r(s)ds$. We denote by $e^{R(t)X}\rho_0$ the solution at time T of the following equation

$$\frac{d\rho(t,x)}{dt} = r(t)X\rho(t,x), \qquad \rho(0,x) = \rho_0(x)$$

if it is well defined.

For $1 \leq i \leq n$, $e^{tD_i}\rho_0(x)$ can be expressed in the form:

$$e^{tD_i}\rho_0(x) = \rho_0(x_1, \cdots, x_i + t, \cdots, x_n) e^{-\int_0^t f_i(x_1, \cdots, x_i + t - s, \cdots, x_n) ds}.$$

Hence, we can extend easily the definition of $e^{tD_i}\rho_0(x)$ to $e^{tD_i}\rho_0(t,x)$, where ρ_0 is also a function of t.

Proposition 1. *Suppose* $\eta = \sum_{i,j=1}^n a_{ij}x_ix_j + \sum_{i=1}^n b_ix_i + d$ *and* $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ *where* a_{ij}, c_{ij}, b_i *and* d *are constants for all* $1 \le i,j \le n$. *If* ρ *is a* C^∞ *function in* x *and* t, *then the following Baker-Campbell-Hausdorff type relations hold:*

(1) For $1 \le i \le n$

$$e^{s_i(t)D_i}L_0\rho = \left[L_0 - s_i(t)\sum_{j=1}^n c_{ij}D_j - \frac{s_i(t)}{2}\left(\sum_{j=1}^n (a_{ij}+a_{ji})x_j + b_i\right) + \frac{s_i^2(t)}{2}\left(\sum_{j=1}^n c_{ij}^2 - a_{ii}\right)\right]e^{s_i(t)D_i}\rho$$

(2) For $1 \le i \le n$, $1 \le k \le n$

$$e^{s_k(t)D_k}M_i\rho = \left[M_i - s_i(t)s_k(t)\left(\sum_{j=1}^n c_{ij}c_{jk} + \frac{1}{2}(a_{ik}+a_{ki})\right)\right]e^{s_k(t)D_k}\rho$$

where

$$M_i = -s_i(t)\sum_{j=1}^n c_{ij}D_j - \frac{s_i(t)}{2}\left(\sum_{j=1}^n (a_{ij}+a_{ji})x_j + b_i\right) + \frac{s_i^2(t)}{2}\left(\sum_{j=1}^n c_{ij}^2 - a_{ii}\right)$$

(3) For $1 \le i \le n$

$$e^{r_i(t)x_i}L_0\rho = \left[L_0 - r_i(t)D_i + \frac{r_i^2(t)}{2}\right]e^{r_i(t)x_i}\rho$$

(4) For $1 \le i \le n$

$$e^{r_i(t)x_i}N\rho = \left[N + r_i(t)\sum_{k=1}^n s_k(t)c_{ki}\right]e^{r_i(t)x_i}\rho$$

where

$$\begin{aligned} N = &- \sum_{i,j=1}^n s_i(t)c_{ij}D_j - \frac{1}{2}\sum_{i=1}^n s_i(t)((a_{ij}+a_{ji})x_j + b_i) \\ &- \sum_{i\le i<j\le n} s_i(t)s_k(t)\left(\sum_{j=1}^n c_{ij}c_{jk} + \frac{1}{2}(a_{ik}+a_{ki})\right) \\ &+ \frac{1}{2}\sum_{i=1}^n s_i^2(t)\left(\sum_{j=1}^n c_{ij}^2 - a_{ii}\right) \end{aligned}$$

(5) For $1 \le i \le n$, $1 \le k \le n$

$$e^{r_k(t)x_k} D_i \rho = [D_i - r_k(t)\delta_{ik}] e^{r_k(t)x_k} \rho$$

(6) For $1 \le i \le n$, $1 \le k \le n$

$$e^{s_k(t)D_k} D_i \rho = [D_i + s_k(t)c_{ik}] e^{s_k(t)D_k} \rho$$

Proof: The following computations are legitimate by the argument given in the Proposition 1 of [14].

(1) $e^{s_i(t)D_i} L_0 \rho$

$$\begin{aligned}
&= \Big\{ L_0 + s_i(t)[D_i, L_0] + \frac{s_i^2(t)}{2}[D_i,[D_i,L_0]] \\
&\quad + \frac{s_i^3(t)}{3!}[D_i,[D_i,[D_i,L_0]]] + \cdots \Big\} e^{s_i(t)D_i} \rho \\
&= \Big\{ L_0 + s_i(t)\Big(-\sum_{j=1}^{n} c_{ij} D_j - \frac{1}{2}\sum_{j=1}^{n}(a_{ij}+a_{ji})x_j - \frac{1}{2} b_i \Big) \\
&\quad + \frac{s_i^2(t)}{2}\Big(\sum_{j=1}^{n} c_{ij}^2 - a_{ii} \Big) \Big\} e^{s_i(t)D_i} \rho \\
&= \Big[L_0 - s_i(t)\sum_{j=1}^{n} c_{ij} D_j - \frac{s_i(t)}{2}\Big(\sum_{j=1}^{n}(a_{ij}+a_{ji})x_j + b_i \Big) \\
&\quad + \frac{s_i^2(t)}{2}\Big(\sum_{j=1}^{n} c_{ij}^2 - a_{ii} \Big) \Big] e^{s_i(t)D_i} \rho
\end{aligned}$$

(2) $e^{s_k(t)D_k} M_i \rho$

$$\begin{aligned}
&= \Big\{ M_i + s_k(t)[D_k, M_i] + \frac{s_k^2(t)}{2}[D_k,[D_k,M_i]] \\
&\quad + \frac{s_i^3(t)}{3!}[D_k,[D_k,[D_k,M_i]]] + \cdots \Big\} e^{s_k(t)D_k} \rho \\
&= \Big\{ M_i + s_k(t)(-s_i(t)\sum_{j=1}^{n} c_{ij}c_{jk} - \frac{s_i(t)}{2}(a_{ik}+a_{ki})) \Big\} e^{s_k(t)D_k} \rho \\
&= \Big\{ M_i - s_i(t)s_k(t)\Big(\sum_{j=1}^{n} c_{ij}c_{jk} + \frac{1}{2}(a_{ik}+a_{ki}) \Big) \Big\} e^{s_k(t)D_k} \rho
\end{aligned}$$

The proofs of (3)-(7) are similar. QED

Theorem 7. *Let E be an estimation algebra of (2.0) satisfying $\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij}$ where c_{ij} are constants for all $1 \le i, j \le n$. Suppose E is finite dimensional, then $h_1, \cdots, h_m$ are affine. Suppose further that $m \ge n$ and the observation matrix has full rank, then $\eta = \sum_{i,j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n b_i x_i + d$ where a_{ij}, b_i and d are constants for all $1 \le i, j \le n$ and the robust Duncan-Mortensen-Zakai equation (2.2) has a solution for all $t \ge 0$ of the form:*

$$\xi(t,x) = e^{T(t)} e^{r_n(t)x_n} \cdots e^{r_1(t)x_1} e^{s_n(t)x_n D_n} \cdots e^{s_1(t)D_1} e^{tL_0} \sigma_0 \tag{4.1}$$

where $T(t), r_1(t), \cdots, r_n(t), s_1(t), \cdots, s_n(t)$ satisfy the following ordinary differential equations (4.2), (4.3) and (4.4)

For $1 \le i \le n$

$$\frac{ds_i(t)}{dt} = r_i(t) + \sum_{j=1}^n s_j(t) c_{ji} + \sum_{k=1}^m h_{ki} y_k(t) \tag{4.2}$$

where $h_k(x) = \sum_{j=1}^n h_{kj} x_j + e_k$, for $1 \le k \le m$; h_{kj} and e_k are constants.

For $1 \le j \le n$

$$\frac{dr_j(t)}{dt} = \frac{1}{2} \sum_{i=1}^n s_i(t)(a_{ij} + a_{ji}) \tag{4.3}$$

and

$$\begin{aligned}
\frac{dT(t)}{dt} = & -\frac{1}{2} \sum_{i=1}^n r_i^2(t) - \frac{1}{2} \sum_{i=1}^n s_i^2(t) \left(\sum_{j=1}^n c_{ij}^2 - a_{ii} \right) \\
& + \sum_{1 \le i < k \le n} s_i(t) s_k(t) \left(\sum_{j=1}^n c_{ij} c_{jk} + \frac{1}{2}(a_{ik} + a_{ki}) \right) \\
& + \sum_{i=1}^n r_i(t) - \sum_{j=2}^n \sum_{i=1}^j s_j(t) c_{ij} + \frac{1}{2} \sum_{i=1}^n s_i(t) b_i \\
& + \frac{1}{2} \sum_{i,j=1}^m y_i(t) y_j(t) \left(\sum_{k=1}^n h_{ik} h_{jk} \right) - \sum_{i,j=1}^n s_i(t) s_j(t) c_{ij}.
\end{aligned} \tag{4.4}$$

It follows then that a universal finite dimensional filter exists for (2.0).

Proof: Since L_0 is uniformly elliptic, for any $t > 0$, $e^{tL_0}\sigma_0$ is C^∞. By differentiating $\xi(t,x)$ we have

$$\frac{\partial \xi(t,x)}{\partial t} = e^{T(t)} e^{r_n(t)x_n} \cdots e^{r_1(t)x_1} e^{s_n(t)D_n} \cdots e^{s_1(t)D_1} L_0 e^{tL_0} \sigma_0$$

$$
\begin{aligned}
&+\frac{ds_1(t)}{dt}e^{T(t)}e^{r_n(t)x_n}\dots e^{r_1(t)x_1}e^{s_n(t)D_n}\dots e^{s_2(t)D_2}D_1e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
&+\cdots\\
&+\frac{ds_n(t)}{dt}e^{T(t)}e^{r_n(t)x_n}\dots e^{r_1(t)x_1}D_ne^{s_n(t)D_n}\dots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
&+\frac{dr_1(t)}{dt}e^{T(t)}e^{r_n(t)x_n}\dots e^{r_2(t)x_2}x_1e^{r_1(t)x_1}e^{s_n(t)D_n}\dots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
&+\cdots\\
&+\frac{dr_n(t)}{dt}e^{T(t)}x_ne^{r_n(t)x_n}\dots e^{r_1(t)x_1}e^{s_n(t)D_n}\dots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
&+\frac{dT(t)}{dt}e^{T(t)}e^{r_n(t)x_n}\dots e^{r_1(t)x_1}e^{s_n(t)D_n}\dots e^{s_1(t)D_1}e^{tL_0}\sigma_0. \qquad (4.5)
\end{aligned}
$$

By applying Proposition 1, we have

$$
\begin{aligned}
&e^{T(t)}e^{r_n(t)x_n}\dots e^{r_1(t)x_1}e^{s_n(t)D_n}\dots e^{s_1(t)D_1}L_0e^{tL_0}\sigma_0\\
=&e^{T(t)}e^{r_n(t)x_n}\dots e^{r_1(t)x_1}\Bigg[L_0-\sum_{i,j=1}^{n}s_i(t)c_{ij}D_j\\
&-\frac{1}{2}\sum_{i=1}^{n}s_i(t)\left(\sum_{j=1}^{n}(a_{ij}+a_{ji})x_j+b_i\right)\\
&+\frac{1}{2}\sum_{i=1}^{n}s_i^2(t)\left(\sum_{j=1}^{n}c_{ij}^2-a_{ii}\right)\\
&-\sum_{1\le i<k\le n}s_i(t)s_k(t)\left(\sum_{j=1}^{n}c_{ij}c_{jk}+\frac{1}{2}(a_{ik}+a_{ki})\right)\Bigg]\\
&\quad e^{s_n(t)D_n}\dots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
=&\Bigg[L_0-\sum_{i=1}^{n}r_i(t)D_i+\frac{1}{2}\sum_{i=1}^{n}r_i^2(t)-\sum_{i,j=1}^{n}s_i(t)c_{ij}D_j\\
&-\frac{1}{2}\sum_{i=1}^{n}s_i(t)\left(\sum_{j=1}^{n}(a_{ij}+a_{ji})x_j+b_i\right)\\
&+\frac{1}{2}\sum_{i=1}^{n}s_i^2(t)\left(\sum_{j=1}^{n}c_{ij}^2-a_{ii}\right)\\
&-\sum_{1\le i<k\le n}s_i(t)s_k(t)\left(\sum_{j=1}^{n}c_{ij}c_{jk}+\frac{1}{2}(a_{ik}+a_{ki})\right)\\
&+\sum_{j,k=1}^{n}s_k(t)r_j(t)c_{kj}\Bigg]\xi(t,x) \qquad (4.6)
\end{aligned}
$$

$$
\begin{aligned}
&e^{T(t)}e^{r_n(t)x_n}\cdots e^{r_1(t)x_1}e^{s_n(t)D_n}\cdots e^{s_{i+1}(t)D_{i+1}}D_i e^{s_i(t)D_i}\cdots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
=&e^{T(t)}e^{r_n(t)x_n}\cdots e^{r_1(t)x_1}e^{s_n(t)D_n}\cdots e^{s_{i+2}(t)D_{i+2}}[D_i+s_{i+1}(t)c_{i,i+1}]\\
&e^{s_{i+1}(t)D_{i+1}}\cdots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
=&e^{T(t)}e^{r_n(t)x_n}\cdots e^{r_1(t)x_1}[D_i+s_n(t)c_{i,n}+s_{n-1}(t)c_{i,n-1}+\cdots\\
&+s_{i+1}(t)c_{i,i+1}]e^{s_n(t)D_n}\cdots e^{s_1(t)D_1}e^{tL_0}\sigma_0\\
=&(D_i-r_i(t)+s_n(t)c_{i,n}+s_{n-1}(t)c_{i,n-1}+\cdots+s_{i+1}(t)c_{i,i+1})\xi(t,x) \quad (4.7)
\end{aligned}
$$

Putting (4.6) and (4.7) in (4.5), we have

$$
\begin{aligned}
&\frac{\partial\xi(t,x)}{\partial t}\\
&= \Bigg[L_0-\sum_{i=1}^n r_i(t)D_i+\frac{1}{2}\sum_{i=1}^n r_i^2(t)-\sum_{i,j=1}^n s_i(t)c_{ij}D_j\\
&\quad -\frac{1}{2}\sum_{i=1}^n s_i(t)\left(\sum_{j=1}^n(a_{ij}+a_{ji})x_j+b_i\right)\\
&\quad +\frac{1}{2}\sum_{i=1}^n s_i^2(t)\left(\sum_{j=1}^n c_{ij}^2-a_{ii}\right)\\
&\quad -\sum_{1\le i<k\le n} s_i(t)s_k(t)\left(\sum_{j=1}^n c_{ij}c_{jk}+\frac{1}{2}(a_{ik}+a_{ki})\right)\\
&\quad +\sum_{j,k=1}^n s_k(t)r_j(t)c_{kj}\Bigg]\xi(t,x)\\
&\quad +\frac{ds_1(t)}{dt}[D_1-r_1(t)+s_n(t)c_{1n}+s_{n-1}(t)c_{1,n-1}\\
&\quad +\cdots+s_1(t)c_{12}]\xi(t,x)\\
&\quad +\frac{ds_2(t)}{dt}[D_2-r_2(t)+s_n(t)c_{2n}+s_{n-1}(t)c_{2,n-1}\\
&\quad +\cdots+s_3(t)c_{23}]\xi(t,x)\\
&\quad +\cdots+\frac{ds_n(t)}{dt}[D_n-r_n(t)]\xi(t,x)\\
&\quad +\frac{dr_1(t)}{dt}x_1\xi(t,x)+\frac{dr_2(t)}{dt}x_2\xi(t,x)\\
&\quad +\cdots+\frac{dr_n(t)}{dt}x_n\xi(t,x)+\frac{dT}{dt}\xi(t,x)\\
&= L_0\xi(t,x)+\sum_{i=1}^n\left[-r_i-\sum_{j=1}^n s_j(t)c_{ji}+\frac{ds_i(t)}{dt}\right]D_i\xi(t,x)
\end{aligned}
$$

$$+\sum_{j=1}^{n}\left[\frac{dr_j(t)}{dt}-\frac{1}{2}\sum_{j=1}^{n}(a_{ij}+a_{ji})s_(t)\right]x_j\xi(t,x)$$

$$-\left[-\frac{1}{2}\sum_{i=1}^{n}r_i^2(t)+\sum_{i=1}^{n}r_i(t)-\frac{1}{2}\sum_{i=1}^{n}s_i^2(t)\Big(\sum_{j=1}^{n}c_{ij}^2-a_{ii}\Big)\right.$$

$$+\sum_{1\le i<k\le n}s_i(t)s_k(t)\Big(\sum_{j=1}^{n}c_{ij}c_{jk}+\frac{1}{2}(a_{ik}+a_{ki})\Big)$$

$$-\sum_{j,k=1}^{n}s_k(t)r_j(t)c_{kj}$$

$$\left.+\frac{1}{2}\sum_{i=1}^{n}s_i(t)-\sum_{j=2}^{n}\sum_{i=1}^{j}s_j(t)c_{ij}-\frac{dT(t)}{dt}\right]\xi(t,x) \tag{4.8}$$

On the other hand, recall that L_i is the zero degree differential operator of multiplication by $h_i=\sum_{j=1}^{n}h_{ij}x_j+e_i$ where h_{ij} and e_i are constants. Then the right hand side of (2.2) is of the form

$$\begin{aligned}\frac{\partial\xi(t,x)}{\partial t}=&L_0\xi(t,x)+\sum_{i=1}^{m}y_i(t)[L_0,L_i]\xi(t,x) \qquad (4.9)\\ &+\frac{1}{2}\sum_{i,j=1}^{m}y_i(t)y_j(t)[[L_0,L_i],L_j]\xi(t,x)\\ =&L_0\xi(t,x)+\sum_{i=1}^{m}y_i(t)\sum_{j=1}^{n}h_{ij}D_j\xi(t,x)\\ &+\frac{1}{2}\sum_{i=1}^{m}y_i^2(t)\left[\sum_{j=1}^{n}h_{ij}D_j,\sum_{k=1}^{n}h_{ik}x_k+e_i\right]\xi(t,x)\\ =&L_0\xi(t,x)+\sum_{i=1}^{n}\Big(\sum_{j=1}^{m}h_{ji}y_j(t)\Big)D_i\xi(t,x)\\ &+\frac{1}{2}\sum_{i,j=1}^{m}y_i(t)y_j(t)\Big(\sum_{k=1}^{n}h_{ik}h_{jk}\Big)\xi(t,x).\end{aligned}$$

By comparing (4.8) and (4.9) it is clear that $\xi(t,x)$ is a solution to (2.2), if (4.2), (4.3) and (4.4) are satisfied. It is also clear that (4.2), (4.3) and (4.4) have solutions for all t.

Finally, to see that these results lead to a finite dimensional filter for (2.0), notice that if we let r_i's, s_i's and T play the role of the z_i's in (2.3a), then (4.2), (4.3) and (4.4) are of the form (2.3a). It is also easy to check that (4.1) is one of the form (2.3b). QED

Acknowledgement

We would like to thank Dr. Wing Shing Wong for introducing the subject to us and for his helpful comments.

References

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

[2] R.W. Brockett. Nonlinear systems and nonlinear estimation theory, in *The Mathematics of Filtering and Identification and Applications* (M. Hazewinkel and J.S. Williams, eds.). Dordrecht: Reidel, 1981.

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

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

[5] P.CC. Collingwood. Some remarks on estimation algebras, *Systems Control Lett.* **7** (1986), 217-224.

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

[7] M.H.A. Davis and S.I. Marcus. An introduction to nonlinear filtering, in *The Mathematics of Filtering and Identification and Applications* (M. Hazewinkel and J.S. Williams, eds.). Dordrecht: Reidel, 1981.

[8] R.T. Dong, L.F. Tam, W.S. Wong and S. S.-T. Yau. Structure and classification theorems of finite dimensional exact estimation algebras *SIAM J. Control and Optimization*, to appear.

[9] U.G. Haussman and E. Paradoux. A conditionally almost linear filtering problem with non-Gaussian initial condition, *Stochastics* **23(2)** (1988), 244-275.

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

[11] D.L. Ocone. Finite dimensional estimation algebras in nonlinear filtering, in *The Mathematics of Filtering and Identification and Applications* (M. Hazewinkel and J.S. Willems, eds.). Dordrecht: Reidel, 1981.

[12] T. Shukhman. Explicit filters for linear systems and certain nonlinear systems with stochastic initial conditions (preprint).

[13] S. Steinberg. Applications of the Lie algebraic formulas of Baker, Campbell, Hausdorff and Zassenhaus to the calculation of explicit solutions of partial differential equations, *J. Differential Equations* **26** (1979), 404-434.

[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 and Optimization* **28(1)** (1990), 173-185.

[15] J. Wei and E. Norman. On global representation of the solutions of linear differential equations as a product of exponentials, *Proc. Amer. Math. Soc.* **15** (1964), 327-334.

[16] W.S. Wong. New classes of finite dimensional nonlinear filters, *Systems Control Lett.* **3** (1983), 155-164.

[17] W.S. Wong. On a new class of finite dimensional estimation algebras, *Systems Control Lett.* **9** (1987), 79-83.

[18] W.S. Wong. Theorems on the structure of finite dimensional estimation algebras, *Systems Control Lett.* **9** (1987), 117-124.

Department of Mathematics, University of Illinois at Chicago, Box 4348, Chicago, Illinois 60680

Communicated by Clyde F. Martin