International Journal of Control
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tcon20

Complete classification of finite-dimensional estimation algebras of maximal rank

Stephen S.-T. Yau [a]

[a] Department of Mathematics, Statistics and Computer Science (M/C 249), University of Illinois, Chicago, 851 South Morgan Street, Chicago, IL, 60607-7045, USA E-mail:

Available online: 08 Nov 2010

To cite this article: Stephen S.-T. Yau (2003): Complete classification of finite-dimensional estimation algebras of maximal rank, International Journal of Control, 76:7, 657-677

To link to this article: http://dx.doi.org/10.1080/0020717031000098435

INT. J. CONTROL, 2003, VOL. 76, NO. 7, 657–677

Taylor & Francis
Taylor & Francis Group

Complete classification of finite-dimensional estimation algebras of maximal rank

STEPHEN S.-T. YAU

The idea of using estimation algebras to construct finite-dimensional non-linear filters was first proposed by Brockett and Clark, and Mitter independently. In his famous talk at the International Congress of Mathematics in 1983, Brockett proposed to classify all finite-dimensional estimation algebras. In this paper we explain why the theory of estimation algebras plays an important role in non-linear filtering. We show how to use the Wei–Norman approach to construct finite-dimensional filters from finite-dimensional estimation algebras. We survey some results in estimation algebras after 1984. We give a self-contained proof of complete classification of finite-dimensional estimation algebras of maximal rank in one place. The proof given here is simpler than those proofs scattered in several papers. This provides the readers with a complete coherent view of the important topic of the classification of finite-dimensional estimation algebras.

Dedicated to Roger Brockett on the occasion of his 65th birthday and to Sanjoy Mitter on the occasion of his 70th birthday.

1. Introduction

Filtering is concerned with making estimates of quantities associated with a stochastic process $\{x_t\}$ on the basis of information gleaned from a related process $\{y_t\}$. The process $\{x_t\}$ is called the signal or state process and $\{y_t\}$ is the observation process. The goal is the computation, for each t, of least square estimates of functions of the signal x_t given the observation history $\{y_s: 0 \leq s \leq t\}$, i.e. the computation of conditional expectations of the form $E[\phi(x_t)/y_s,\ 0 \leq s \leq t] = \widehat{\phi(x_t)}$ or perhaps even the computation of the entire conditional distributional of x_t, given the observation history. In many (engineering) applications the data come in sequentially and one does not really want a calculating procedure which needs all the data y_s, $0 \leq s \leq t$, every time t that it is desired to find $\widehat{\phi(x_t)}$; rather we would like to have a procedure which uses a statistics m_t which can be updated using only the new observations y_s, $t \leq s \leq t'$ to its value $m_{t'}$, i.e.

$$m_{t'} = a(m_t, t, t', \{y_s : t \leq s \leq t'\})$$

and from which the desired conditional expectation can be calculated directly, i.e.

$$\widehat{\phi(x_t)} = E[\phi(x_t)/y_s, 0 \leq s \leq t] = b(t, y_t, m_t)$$

Finally to actually implement the filter it would be nice if m_t were a finite dimensional quantity. All this leads to the (ideal) notion of a finite dimensional recursive filter. By definition such a filter is a system

$$\mathrm{d}\xi_t = \alpha(\xi_t)\mathrm{d}t + \sum_{i=1}^{p} \beta_i(\xi_t)\mathrm{d}y_{it}$$

driven by the observation y_{it}; y_{it} is the ith component of y_t $i = 1, \ldots, p$; together with an output map

$$\widehat{\phi(x_t)} = \gamma(\xi_t)$$

This was solved in the context of linear dynamics by Kalman and Bucy (1960, 1961) and the resulting 'Kalman filter' has of course enjoyed immense success in a wide variety of applications. Attempts were soon made to generalize the results to systems with non-linear dynamics. This is a substantially more difficult problem, being in general infinite-dimensional, but nevertheless equations describing the evolution of conditional distributions were obtained by several authors in the mid-sixties; for example, Bucy (1965), Duncan (1967), Kushner (1964), Mortensen (1966), Shiryaev (1967), Stratonovich (1968) and Wonham (1965). Wonham (1965) studied the important finite-state case and evaluated numerically performance of the optimal non-linear filter for one example and found the performance to be better than that of the simpler Wiener filter. Zakai (1969) obtained these equations in substantially simpler form using the so-called 'reference probability' method (see Wong (1971)).

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 1960s and early 1970s, the basic approach to non-linear filtering theory was via the 'innovation methods' originally proposed by Kailath (1968) and Frost and Kailath (1971) and subsequently rigorously developed by Fujisaki *et al.* (1972). As pointed out by Mitter (1979), the difficulty with this approach is that the innovation process is not, in general, explicitly computable (except in the well-known Kalman–Bucy case). In the late 1970s, Brockett and Clark (1980), Brockett (1981) and Mitter (1979) proposed the idea of using estimation

Received 1 January 2002. Revised 1 December 2002. Accepted 1 January 2003.

† Department of Mathematics, Statistics and Computer Science (M/C 249), University of Illinois at Chicago, 851 South Morgan Street, Chicago, IL 60607-7045, USA. e-mail: yau@uic.edu

International Journal of Control ISSN 0020–7179 print/ISSN 1366–5820 online © 2003 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/0020717031000098435

algebras to construct a finite-dimensional non-linear filter. This Lie algebra approach has several merits. First, it takes into account of geometrical aspects of the situation. Second, it explains convincingly why it is easy to find exact recursive filters for linear dynamical systems while it is very difficult to filter something like the cubic sensor described in the work of Hazewinkel *et al.* (1998 a). The third, and perhaps most important, merit of the Lie algebra approach is the following. As long as the estimation algebra is finite dimensional, not only can the finite dimensional recursive filter be constructed explicitly, but also the filter so constructed is universal in the sense of Chaleyat-Maurel and Michel (1984). Moreover, the number of sufficient statistics in the Lie algebra method, which requires computing the conditional probability density, is linear in n, where n is the dimension of the state space. This is a consequence of our classification result (see Corollary 2). Finally the Lie algebraic methods are useful for classifying equivalence of finite dimensional filters and for indicating when no finite dimensional filters exist. In those cases where no finite dimensional representations exist the available methods must be redirected to the construction of consistent and useful approximate filters (see Marcus (1984) for an example).

In his talk at the International Congress of Mathematics in 1983, Brockett proposed the problem of classifying finite-dimensional estimation algebras. Since then, the concept of estimation algebra has been proven to be invaluable tool in the study of non-linear filtering problems. Nevertheless, the structure and classification of finite-dimensional estimation algebras were studied in detail only in the early 1990s by Tam *et al.* (1990), Chiou and Yau (1994), Yau (1994), Chen and Yau (1996, 1997), Chen *et al.* (1996, 1997), Wu *et al.* (2002) and Yau and Hu (preprint). In Wong (1987), the concept of Ω was introduced, which is defined as the matrix whose (i,j) element is $\omega_{ij} = (\partial f_j/\partial x_i) - (\partial f_i/\partial x_j)$, where f is the drift term of the state evolution equation (1). The programme of classifying finite dimensional estimation algebras of maximal rank was begun in 1990 by Yau. There are four crucial steps here.

Step 1. In 1990, Yau first observed that Wong's Ω-matrix plays an important role. As the first crucial step, he classifies all finite dimensional estimation algebras of maximal rank if Wong's matrix has entries in constant coefficients. His result was announced in 1990 (Yau 1990) and the detail of the proof was published in 1994 (Yau 1994). Chiou and Yau (1991) formally introduced the concept of finite dimensional estimation algebra of maximal rank and gave classification when the state space dimension n is at most 2. Their results were published in 1994 (Chen and Yau 1996).

Step 2. The second crucial step was due to Chen and Yau in 1996 (Chen and Yau 1997). They developed quadratic structure theory for finite dimensional estimation algebra. They laid down all the ingredients which are needed to give classification of finite dimensional estimation algebras of maximal rank. In particular, they introduced the notion of quadratic rank k. In this way, the Wong's Ω-matrix is divided into three parts: (1) (ω_{ij}), $1 \le i, j \le k$; (2) (ω_{ij}), $k+1 \le i, j \le n$ and (3) (ω_{ij}), $1 \le i \le k$, $k+1 \le j \le n$, or $k+1 \le i \le n$, $1 \le j \le k$. Chen and Yau (1997) proved among many other things that part (1) (ω_{ij}), $1 \le i, j \le k$, is a matrix with constant coefficients.

Step 3. In their published paper, Chen *et al.* (1997) proved the weak Hessian matrix non-decomposition theorem for $n \le 4$. As a result, part (2), (ω_{ij}), $k+1 \le i, j \le n$, is a matrix with constant coefficients. In their paper, Wu *et al.* (2002) proved the weak Hessian matrix non-decomposition theorem for general n. Thus part (2), (ω_{ij}), $k+1 \le i, j \le n$ is also a matrix with constant coefficients for arbitrary n.

Step 4. This final step was also done in 1997. Yau and Hu (preprint) used the full power of the quadratic structure theory developed by Chen and Yau (1997) to prove that the matrix (ω_{ij}), $1 \le i \le k$, $k+1 \le j \le n$ and the matrix (ω_{ij}), $k+1 \le i \le n$, $1 \le j \le k$ are with the constant coefficients.

The above four steps complete the classification of finite dimensional estimation algebras of maximal rank. Therefore Yau and his coworkers have proved the following theorem.

Theorem 1: *Suppose that the state space of the filtering system* (1) *is of dimension n. If E is the finite-dimensional estimation algebra with maximal rank, then* $f = \nabla\phi + (\alpha_i, \ldots, \alpha_n)$ *where* ϕ *is a smooth function and* α_i, $1 \le i \le n$, *are affine functions and E is a real vector space of dimension* $2n+2$ *with basis given by* $1, x_1, \ldots, x_n, D_1, \ldots, D_n$ *and* L_0 *where* D_i *and* L_0 *are defined in* (5) *and* (7).

Mitter conjectured a long time ago that all the functions in finite dimensional estimation algebras are polynomial of degree one. As an immediate consequence of the above theorem, we have the following corollary.

Corollary 1 (Mitter conjecture)**:** *Suppose that E is the finite-dimensional estimation algebra with maximal rank* *corresponding to the filter system* (1). *Then any function in E is a polynomial of degree one.*

The following corollary is an immediate consequence of the above theorem and Theorem 7 of Yau (1994) (cf. Theorem 14 below).

Corollary 2: *Suppose that the state space of the filtering system* (1) *is of dimension n. If E is the finite-dimensional estimation algebra with maximal rank, then the number of statistics in order to compute the conditional density by Lie algebraic methods is n.*

In § 2, we recall some basic concepts and notations. We prove two fundamental results: Ocone theorem (Theorem 2) and nonexistence solution of over-determined PDE (Theorem 3 and Corollary 3). We explain why one wants to work with robust DMZ equation (3) rather than stochastic partial differential equation (2). We also recall the gauge transformation of Mitter and Brockett's estimation equivalence group in non-linear filtering. In § 3, we survey some result developed after the beautiful survey article by Marcus (1984). We recall Wong's structure theorem of estimation algebra in case the drift term $f(x)$ is real analytic with some growth conditions as well as a new class of finite dimensional estimation algebra introduced by Wong. The concept of finite dimensional exact estimation algebra is introduced. The structure and classification of these algebras are discussed. We recall Cohen de Lara's structure theorem for those finite dimensional estimation algebras of maximal rank with very strong assumption on the structure of differential operators in the estimation algebras. We also recall the general construction of finite dimensional estimation algebra with non-maximal rank by Rasoulian and Yau. The most recent beautiful result by Chiou and Chiueh on classification of five-dimensional estimation algebras is discussed. In § 4, we survey some results obtained in Yau (1994). In particular, the classification result is proved under the assumption that Ω-matrix has constant coefficients. We describe in detail how to solve the time-varying parabolic partial differential equation by Wie–Norman theory. We characterize those drift $f(x)$ for which the Ω-matrix has constant coefficients. We use the Wei–Norman approach to construct a finite dimensional filter if the estimation algebra is finite dimensional. In § 5, we survey some results obtained in Chen and Yau (1996). In particular, quadratic structure theory is developed for finite dimensional estimation algebra. The linear structure of Ω-matrix is proved and the constant structure of the upper left corner of the Ω-matrix is also proved. The proof given here is different from those in Chen and Yau (1996). In § 6, we survey the result obtained in Wu *et al.* (2002). We prove the constant structure of the lower right corner of the Ω-matrix. In § 7, we survey some results obtained in Yau and Hu (preprint). We prove the constant structure of the lower left corner and the upper right corner of the Ω-matrix.

2. Some basic concepts, fundamental tools and equivalent filtering problems

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

$$\left.\begin{aligned} \mathrm{d}x(t) &= f(x(t))\,\mathrm{d}t = g(x(t))\,\mathrm{d}v(t), \quad & x(0) &= x_0 \\ \mathrm{d}y(t) &= h(x(t))\,\mathrm{d}t + \mathrm{d}w(t), & y(0) &= 0 \end{aligned}\right\} \tag{1}$$

Here x, v, y and w are respectively $\mathbb{R}^n$, $\mathbb{R}^p$, $\mathbb{R}^m$ and $\mathbb{R}^m$ valued processes, and v and w have components which are independent, standard Brownian processes. We assume that $n = p$; 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 $\{y(s) : 0 \le s \le t\}$. It is well known (see, e.g. Davis and Marcus 1981) that $\rho(t,x)$ is given by normalizing $\sigma(t,x)$, i.e. $\rho(t,x) = \sigma(t,x)/\int \sigma(t,x)\mathrm{d}x$, which satisfies the Duncan–Mortensen–Zakai (DMZ) equation

$$\left.\begin{aligned} \mathrm{d}\sigma(t,x) &= L_0\sigma(t,x)\,\mathrm{d}x + \sum_{i=1}^{m} L_i\sigma(t,x)\,\mathrm{d}y_i(t) \\ \sigma(0,x) &= \sigma_0 \end{aligned}\right\} \tag{2}$$

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) is a stochastic partial differential equation (with as 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 path is of measure zero. Thus there arises the question whether there exist a version of (2) which can be interpreted pathwise for all $y(t)$ and for which the solution of (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 (1980) studied this problem and proposed some robust algorithms. In our case, his basic idea reduces to defining a new unnormalized density

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

Davis reduced (2) to the following time-varying partial differential equation, which is called the robust DMZ equation

$$\left.\begin{aligned} \frac{\partial u}{\partial t}(t,x) &= L_0u(t,x) + \sum_{i=1}^{m} y_i(t)[L_0, L_i]u(t,x) \\ &\quad + \frac{1}{2}\sum_{i,j=-1}^{m} y_i(t)y_j(t)[[L_0, L_i], L_j]u(t,x) \\ u(0,x) &= \sigma_0(x) \end{aligned}\right\} \quad (3)$$

which is a time-varying partial differential equation. Here we have used the following notation.

Definition 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 $(x, y) \mapsto [x, y]$ (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) $[x, y] = 0$ for all $x \in \mathcal{F}$;
(iii) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \ (x, y, z \in \mathcal{F})$.

Definition 2: The estimation algebra E of a filtering system (1) is defined as the Lie algebra generated by $\{L_0, L_1, \ldots, L_m\}$ 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 $1 \le i \le n$, there exists a constant c_i such that $x_i + c_i$ is in E.

Definition 3: Wong's matrix of a filtering system (1) is a $n \times n$ matrix $\Omega = (\omega_{ij})$ defined by

$$\omega_{ij} = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_i}, \qquad \forall 1 \le i, j \le n \quad (4)$$

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

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

Define

$$D_i = \frac{\partial}{\partial x_i} - f_i \quad (5)$$

$$\eta = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i} \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{m} h_i^2 \quad (6)$$

Then

$$L_0 = \frac{1}{2}\left(\sum_{i=1}^{n} D_i^2 - \eta\right) \quad (7)$$

For the convenience of the readers, we list the following elementary lemmas without proof. The lemmas were proven in Chiou and Yau (1994) and Yau (1994).

Lemma 1:

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

(ii) $[gD_i, h] = g\dfrac{\partial h}{\partial x_i}$, *where* g, h *are any function defined on* $\mathbb{R}^n$

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

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

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

(vi) $$[D_i^2, D_j^2] = 4\omega_{ji}D_jD_i + 2\left(\frac{\partial \omega_{ji}}{\partial x_j}\right)D_i + 2\left(\frac{\partial \omega_{ji}}{\partial x_j}\right)D_i + 2\left(\frac{\partial \omega_{ji}}{\partial x_i}\right)D_j + \frac{\partial^2 \omega_{ji}}{\partial x_i \partial x_j} + 2\omega_{ji}^2$$

(vii) $$\begin{aligned}[D_k^2, hD_iD_j] &= 2\left(\frac{\partial h}{\partial x_k}\right)D_kD_iD_j + 2h\omega_{jk}D_iD_k \\ &\quad + 2h\omega_{ik}D_kD_j + \left(\frac{\partial^2 h}{\partial x_k^2}\right)D_iD_j \\ &\quad + 2h\left(\frac{\partial \omega_{jk}}{\partial x_i}\right)D_k + h\left(\frac{\partial \omega_{jk}}{\partial x_k}\right)D_i \\ &\quad + h\left(\frac{\partial \omega_{ik}}{\partial x_k}\right)D_j + h\left(\frac{\partial^2 \omega_{jk}}{\partial x_i \partial x_k}\right)\end{aligned}$$

(viii) $$\begin{aligned}[gD_iD_j, hD_k] &= g\left(\frac{\partial h}{\partial x_j}\right)D_iD_k + g\left(\frac{\partial h}{\partial x_i}\right)D_jD_k \\ &\quad + gh\omega_{kj}D_i + gh\omega_{ki}D_j \\ &\quad + g\left(\frac{\partial^2 h}{\partial x_i \partial x_j}\right)D_k + gh\left(\frac{\partial \omega_{kj}}{\partial x_i}\right) \\ &\quad - h\left(\frac{\partial g}{\partial x_k}\right)D_iD_j\end{aligned}$$

Lemma 2:
(i) $[L_0, x_j + c_j] = D_j, \quad 1 \le j \le n$
(ii) $[D_i x_j + c_j] = \delta_{ij}, \quad 1 \le i,\ j \le n$
(iii) $[D_i, D_j] = \omega_{ji}, \quad 1 \le i,\ j \le n$
(iv) $Y_j := [L_0, D_j] = \sum_{i=1}^{n}\left(\omega_{ji} D_i + \frac{1}{2}\frac{\partial \omega_{ji}}{\partial x_i}\right) + \frac{1}{2}\frac{\partial \eta}{\partial x_j}$, $1 \le j \le n$
(v) $[Y_j, \omega_{kl}] = \sum_{i=1}^{n} \omega_{ji}\frac{\partial \omega_{kl}}{\partial x_i}, \quad 1 \le j,k,l \le n$
(vi) $[Y_j, D_k] = \sum_{i=1}^{n}\left(\omega_{ji}\omega_{ki} - \frac{\partial \omega_{ji}}{\partial x_k} D_i\right) - \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2 \omega_{ji}}{\partial x_k \partial x_i} - \frac{1}{2}\frac{\partial^2 \eta}{\partial x_k \partial x_j}$, $1 \le j,k \le n$

The following theorem due to Ocone (1980) is the first result which allows us to understand what kind of functions can appear in finite dimensional estimation algebra.

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 at most* (2).

Proof: Let $Ad_{L_0}(\xi) = [L_0, \xi]$ and $Ad^k_{L_0}\xi = [L_0, Ad^{k-1}_{L_0}(\xi)]$. Then it is easy to see that

$$Ad^k_{L_0}(\xi) = \sum_{i_l,\ldots,i_k=1}^{n} \frac{\partial^k \xi}{\partial x_{i_l}\ldots\partial x_{i_k}} D_{i_l}\ldots D_{i_k} + (k-1)\text{th order differential operator}$$

Since $Ad^k_{L_0}(\xi)$ is in E for all k, the finite dimensionality of E implies that $\partial^k \xi/\partial x_{i_l}\ldots\partial x_{i_k} = 0$, for $1 \le i_l, \ldots, i_k \le n$, if k is large enough. It follows that ξ is a polynomial.

Observe that $\xi \in E$ implies

$$\sum_{i=1}^{n}\left(\frac{\partial \xi}{\partial x_i}\right)^2 = [Ad_{L_0}(\xi), \xi], \in E$$

The facts that ξ is a polynomial and E is finite dimensional imply ξ is a polynomial of degree at most 2. □

We shall now prove a very useful theorem in PDE which can be found in Yau (1994).

Theorem 3: *Let $F(x_1, \ldots, x_n)$ be a C^∞ function on $\mathbb{R}^n$. Suppose that there exists a path $c\colon \mathbb{R} \longrightarrow \mathbb{R}^n$ 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 $B_\delta(c(t)) = \{x \in \mathbb{R}^n : \|x - c(t)\| < \delta\}$. Then there are no C^∞ functions $f_1, f_2, \ldots, f_n$ on $\mathbb{R}^n$ satisfying the equation*

$$\sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 = F \tag{8}$$

Proof: Let $\psi \in C_0^\infty$ be any C^∞ function with compact support. Multiplying (8) with ψ^2 and integrating the equation of $\mathbb{R}^n$, we get

$$\int_{\mathbb{R}^n}(\operatorname{div} f)\psi^2 + \int_{\mathbb{R}^n}\psi^2(f\cdot f) = \int_{\mathbb{R}^n} F\psi^2$$

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

$$\begin{aligned}\int_{\mathbb{R}^n} F\psi^2 &= -\int_{\mathbb{R}^n} 2\psi\nabla\psi\cdot f = \int \psi^2(f\cdot f)\\ &\ge -\int_{\mathbb{R}^n}|\nabla\psi|^2 - \int_{\mathbb{R}^n}\psi^2(f\cdot f) + \int\psi^2(f\cdot f)\\ &= -\int_{\mathbb{R}^n}|\nabla\psi|^2\end{aligned}$$

Therefore we get

$$\int_{\mathbb{R}^n} F\psi^2 + \int_{\mathbb{R}^n}|\nabla\psi|^2 \ge 0 \tag{9}$$

for all $\psi \in C_0^\infty$. Take any non-zero C^∞ function θ with compact support in the ball $B_\delta(0)$ of radius δ. Define ψ to be θ followed by a translation by $c(t)$. Observe that $\int_{\mathbb{R}^n}|\nabla\psi|^2$ is independent of the translation selected. On the other hand, $\int_{\mathbb{R}^n} F\psi^2 \to -\infty$ as $t \to \infty$ by our assumptions. This leads to a contradiction to (9). □

Corollary 3: *Let $F(x_1, \ldots, x_n)$ be a polynomial $\mathbb{R}^n$. Suppose that degree of F is odd. Then there are no C^∞ functions $f_1, \ldots, f_n$ on $\mathbb{R}^n$ satisfying the equation*

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

The estimation algebra can be useful in recognizing equivalent filtering problems in the sense that E is invariant under certain transformations of a filtering problem. First, note that if we perform a 'change of scale' on the unnormalized conditioned density function, multiplying it by a non-negative function $\psi(x)$ taking $\sigma \to \tilde\sigma = \psi(x)\sigma$, the DMZ equation becomes

$$\mathrm{d}\tilde\sigma(t,x) = \psi(x)L_0\psi^{-1}(x)\tilde\sigma(t,x)\,\mathrm{d}x + \sum_{i=1}^{n} L_i\tilde\sigma(t,x)\,\mathrm{d}y_i(t)$$

This transformation takes $L_0 \mapsto \psi L_0 \psi^{-1}$ and $h_i \mapsto \psi h_i \psi^{-1} = h_i$, $1 \le i \le m$ and the corresponding Lie algebras are isomorphic. Specifically we have the following theorem.

Theorem 4: *If $\psi\colon \mathbb{R}^n \to \mathbb{R}$ is smooth and positive, then the Lie algebra E generated by $L_0, h_1, \ldots, h_m$ and the Lie algebra E generated by $\psi L_0 \psi^{-1}, h_1, \ldots, h_m$ are isomorphic with an isomorphism $\phi\colon A \to \psi A \psi^{-1}$ for all $A \in E$.*

The proof of Theorem 4 can be found for example in Marcus (1984). The transformation in Theorem 4 is called gauge transformation by Mitter (1978).

A related phenomenon occurs when one performs a smooth non-singular change of variables $z = \alpha(x)$ with inverse $x = \beta(z)$. Then Brockett (1979) proved the following theorem.

Theorem 5: *If the estimation problem* (1), (2) *is transformed by a smooth non-singular change of coordinates* $z_t = \alpha(x_t)$, *so that* $\{z_t\}$ *has generator* L_{0z}, *then the mapping*

$$\phi: L_0 \to L_{0z}, \quad \phi: h_i \mapsto h_i \circ \beta, \quad 1 \le i \le m$$

extends to an isomorphism of the Lie algebras $\{L_0, h_1, \ldots, h_m\}_{L.A.}$ *and* $\{L_{0z}, h_1 \circ \beta, \ldots, h_m \circ h\beta\}_{L.A.}$.

Since the set of all transformations consisting of successive applications of the two types of transformations described in Theorems 4 and 5 forms a group under composition, Brockett (1979) has called this the *estimation equivalence group* and he has termed two estimation problems equivalent if their estimation algebras can be transformed into one another by elements of this group. This group is also called the (stochastic) invariance group by Hijab (1980).

3. Structures of finite-dimensional estimation algebras

The concept of the estimation algebra has played a very important role in the recent studies of non-linear filtering systems. The beautiful survey article by Marcus (1984) has provided a detail account of many developments that involve the estimation algebra. In this section, we shall survey some estimation algebra related results developed after Marcus (1984). Wong (1987 a) proved several theorems concerning the structure of finite dimensional estimation algebras. Among other things, these results together with his other results in Wong (1987 b) shed new light on the classification problem of finite dimensional estimation algebras. The structure theorem of Wong (1987 a) can be stated as follows.

Theorem 6: *Assume that h and f in* (1) *are real analytic functions on* $\mathbb{R}^n$, *and f satisfies the growth condition for any i, all the first, second, and third order partial derivatives of* f_i *are bounded functions:*

(1) *If the degree of h in x is greater than* 1, *then the estimation of* (1) *is infinite dimensional.*
(2) *If the estimation algebra of* (1) *is finite dimensional, then it has no differential operator of degree higher than two. It has a basis consisting of one second degree differential operator,* L_0, *first degree operator(s) of the form* $\sum_{i=1}^n \alpha_i D_i + \sum_{i=1}^n \beta_i(\partial\eta/\partial x_i)$ *where* α_i, β_i *are constants, and zero degree differential operator(s) affine in x.*
(3) *All finite dimensional estimation algebras* (1) *are solvable.*

The growth condition in Theorem 6 guarantees that (1) has a well-defined solution for all time. It also implies that for all i, $f_i = O(|x|)$ at infinity. (We say $a(x) = O(b(x))$ at infinity if there exist constants M and N such that $|a(x)| \le M|b(x)|$ for $|x| \ge N$).

Wong (1987 b) introduced a new class of solvable finite dimensional estimation algebras. Using either the Wei and Norman (1964) method or the function-space integral approach of Benés (1981), one can derive from these results new finite dimensional non-linear filters. In our case, Wong's (1987 b) result can be stated as follows.

Theorem 7: *Let* $h_i = H_i^{\mathrm{T}} x$ *where* $H_i^{\mathrm{T}} = (H_{i1}, \ldots, H_{in})$ *is a constant vector,* $1 \le i \le n$. *Let* Ω *be the skew-symmetric matrix defined in Definition 3 and* $J_\eta = (\partial^2\eta/\partial x_i \partial x_j)$ *denote the Hessian of* η. *Define* $\nabla\eta = (\partial\eta/\partial x_1, \ldots, \partial\eta/\partial x)^{\mathrm{T}}$ *and* $D = (D_1, \ldots, D_n)^{\mathrm{T}}$. *Let U denote the associative algebra of n by n matrix-valued function of x over* $\mathbb{R}$ *generated by* $\{\Omega, J_\eta, I\}$, *where I stands for the identity matrix. If* $H_i^{\mathrm{T}}\Gamma$ *is a vector of constant functions for any i and any* Γ *in U, then the dimension of the estimation algebra of* (1) *is bounded above by* $2n + m + 2$.

Tam *et al.* (1990) introduced the concept of an exact estimation algebra, i.e. estimation algebra with $f = \nabla\phi$ for some smooth function ϕ defined on $\mathbb{R}^n$. A simple algebraic necessary and sufficient condition was proved for an exact estimation algebra to be finite-dimensional. They also provided a detailed examination of the relationship between finite-dimensional exact estimation algebras and finite-dimensional non-linear filters. More specifically they proved the following structure theorems.

Theorem 8: *Let E be a finite-dimensional exact estimation algebra. Then:*

(1) $h_1, \ldots, h_m$ *are polynomials of degree at most one.*
(2) *E has a basis consisting of one second-degree differential operator* L_0, *first-degree differential operator(s) with constant coefficients for the* $\partial/\partial x_i$ *terms, and zero-degree differential operator(s) affine in x. Moreover, if X and Y are in E with degree less than or equal to one, then* $[X, Y]$ *is a constant.*
(3) *E is a solvable Lie algebra.*

Theorem 9: *Suppose E is an exact estimation algebra. Then E is finite-dimensional if and only if* $\nabla h_i^{\mathrm{T}} J_\eta^j$ *is a constant for* $1 \le i \le m$ *and all* $j = 0, 1, \ldots,$ *where* J_η *is the Hessian matrix of* η.

Given the importance of the estimation algebra, a natural question arises as to whether we can classify all finite-dimensional exact estimation algebras up to

Lie algebraic isomorphism. Theorems 8 and 9 provide a starting point for solving this problem. Dong *et al.* (1991) provided a more explicit structure theorem for an important subclass of finite-dimensional exact estimation algebras as follows.

Theorem 10: *Suppose E is a finite-dimensional exact estimation algebras of maximal rank. Then it is a real vector space of dimension* $2n+2$ *with basis given by 1,* $x_1, x_2, \ldots, x_n, D_1, \ldots, D_n$ *and* L_0. *Moreover,* η *is a polynomial of degree at most two and the quadratic part of* $\eta - \sum_{i=1}^m h_i^2$ *is positive semidefinite.*

A next question that arises naturally is whether we can classify all filtering systems with finite-dimensional exact estimation algebras up to state-space diffeomorphism. This is apparently a very difficult problem and requires a careful study of partial differential equations of type (8) with $f_i = \partial\phi/\partial x_i$. The connection between these types of equations and the non-linear filtering problem was first noted by Benés (1981). The properties of these equations, however, are not well-known. In Dong *et al.* (1991), the authors provided some answers in regard to the existence and uniqueness of the solutions of these types of equations.

Cohen de Lara (1997) proved a structure theorem under a severe assumption of estimation algebra as follows.

Theorem 11: *Suppose E is a finite-dimensional estimation algebra of the form* $\mathbb{R}L_0 \oplus F$, *where F is a finite-dimensional Lie algebra consisting of linear partial differential operators of order less than or equal to one. If E is of maximal rank, then*

(1) $h_1, \ldots, h_p$ *are polynomials of degree less than or equal to one,*

(2) *there exists a skew-symmetric matrix K and a smooth function* ϕ *such that*

a. *the drift f may be written as* $f(x) = \nabla\phi(x) + Kx$

b. *the function* $\nabla\phi + \|\nabla\phi + Kx\|^2$ *is quadratic.*

Rasoulian and Yau (1997) studied finite-dimensional estimation algebras of non-maximal rank. They gave general construction of finite-dimensional estimation algebras of non-maximal rank. Suppose that E is the finite-dimensional estimation algebra of (1). Consider the enlarged filter system

$$\left.\begin{aligned} \mathrm{d}\tilde{x}(t) &= \tilde{f}(\tilde{x}(t))\,\mathrm{d}t + \tilde{g}(\tilde{x}(t))\,\mathrm{d}\tilde{v}(t), \quad & \tilde{x}(0) = x_0 \\ \mathrm{d}y(t) &= h(\tilde{x}(t))\,\mathrm{d}t + \mathrm{d}w(t), & y(0) = 0 \end{aligned}\right\} \quad (10)$$

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 w have components which are independent, standard Brownian processes. Let $\tilde{E}$ be the estimation algebra associated to (10). Rasoulian and Yau showed that $\tilde{E}$ is isomorphic to E. Note that although E is of maximal rank with respect to (1), $\tilde{E}$ is of non-maximal rank with respect to (10) in general. They suspected that all finite dimensional estimation algebras of non-maximal rank are essentially arising in this way. In Yau and Rasoulian, they classified all estimation algebras of dimension at most four. In a recent preprint of Chiou and Chiueh (preprint), the authors have done spectacular works on five-dimensional estimation algebra. Specifically, they have proved the following theorem.

Theorem 12: *The five-dimensional estimation algebra is isomorphic to a Lie algebra having a basis given by* $\{1, x_1, D_1, Y_1, L_0\}$ *where*

$$D_1 = \frac{\partial}{\partial x_1} - f_1, \quad Y_1 = [L_0, D_1] = \sum_{i=1}^n \omega_{i1} D_i + \frac{1}{2}\frac{\partial\eta}{\partial x_1},$$

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

Moreover $\omega_{1j} =$ constant ($\neq 0$, *for some* $j = 2, \ldots, n$), $\eta = \alpha x_1^2 + \beta(x_2, \ldots, x_n)x_1 + \gamma(x_2, \ldots, x_n)$, *where* $\beta(x_1, \ldots, x_n)$ *and* $\gamma_2(x_2, \ldots, x_n)$ *are* C^∞ *functions. In particular,* $f_1, \ldots, f_n$ *have to satisfy the equations*

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

$$\frac{1}{2}\frac{\partial\beta}{\partial x_i} = c_1\omega_{1i} + \sum_{j=1}^n \omega_{1j}\omega_{ij}, \quad i = 2, \ldots, n$$

$$\sum_{j=1}^n \omega_{1j}\frac{\partial\beta}{\partial x_j} = c_2$$

$$\sum_{j=1}^n \omega_{1j}\frac{\partial\gamma}{\partial x_j} = c_3\beta(x_2, \ldots, x_n) + c_4$$

where α_1, c_1, c_2, c_3 *and* c_4 *are constants, and* $\alpha \geq 1$.

4. Estimation algebras of maximal rank with Ω-matrix in constant coefficients and Wei–Norman approach to construct finite dimensional filters

The application of the Lie algebra method to non-linear filtering problems has led to a number of new results concerning finite dimensional filters and to a deeper understanding of the structure of non-linear filtering problems in general. In this section we shall show how to construct finite dimensional filter by Lie algebra method via Wei–Norman approach.

We begin with the following general lemma observed in Yau (1994)

Lemma 3: *Let E be a finite dimensional estimation algebra with maximal rank. Then $E \supseteq \langle 1, x_1, \dots, x_n, D_1, \dots, D_n, L_0 \rangle$ and $\omega_{ij} \in E$ is a polynomial of degree 2 for all $1 \le i, j \le n$.*

Proof: This is an immediate consequence of Lemma 2 and Theorem 2. □

We now prove the following theorem (Yau 1994) which plays a fundamental role in the classification of finite-dimensional estimation algebras of maximal rank.

Theorem 13: *Let E be a finite-dimensional estimation algebra of* (1) *such that $\omega_{ij} = (\partial f_j/\partial x_i) - (\partial f_i/\partial x_j) =$ constant c_{ij}. If E is of maximal rank, then E is a real vector space of dimension $2n+2$ with basis given by $1, x_1, \dots, x_n$, $D_1, \dots, D_n$ and L_0 and η defined in* (6) *is a polynomial of degree* 2.

Proof: Since E is of maximal rank, there are constants c_is such that $x_i + c_i$ is in E for $i = 1, \dots, n$. In view of Lemma 2, the following elements are in E

$$[L_0, x_i + c_i] = D_i \in E \tag{11}$$

$$[D_i, x_i + c_i] = \delta_{ij} \in E \tag{12}$$

$$[L_0, D_i] = \sum_{i=1}^{n} c_{ij} D_j + \frac{1}{2}\frac{\partial \eta}{\partial x_i} \in E \tag{13}$$

Equations (11) and (13) imply that $\partial\eta/\partial x_i$ is in E for all $1 \le i \le n$. If η is a quadratic polynomial, then in view of (11), (12) and (13), we see easily that E is a finite dimensional real vector space spanned by $1, x_1, \dots, x_n$, $D_1, \dots, D_n$ and L_0. Therefore to finish the proof of this theorem, we only need to prove that η is a polynomial of degree at most 2.

To see that η is a quadratic polynomial, we first observe that by Theorem 2, $\partial\eta/\partial x_i$, for all $1 \le i \le n$, are polynomials of degree at most two because $\partial\eta/\partial x_i \in E$ by (13). It follows that η is a polynomial at most three. If the homogeneous degree 3 part of η is non-zero, then clearly there exists a straight line $c(t)$ passing through the origin such that $\lim_{t\to\infty} \eta(c(t)) = -\infty$. In particular

$$\lim_{t\to\infty} \left(\eta - \sum_{i=1}^{m} h_i^2 \right)(c(t)) = -\infty$$

Recall that

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

In view of Corollary 3, we get a contradiction. Therefore the homogeneous degree 3 part of η must be zero. □

Constructing a robust finite-dimensional filter to (1) is equivalent to finding a smooth manifold M, complete C^∞ vector fields μ_i on M, C^∞ function ν on $M \times \mathbb{R}^n$, and ω_is on $\mathbb{R}^m$ such that $u(t, x)$ in (3) can be represented in the form

$$\frac{dz}{dt}(t) = \sum_{i=1}^{k} \mu_i(z(t))\omega_i(y(t)), \qquad z(0) \in M \tag{14}$$

$$u(t, x) = \nu(z(t), t, x) \tag{15}$$

Following Chaleyat-Maurel and Michel (1984), we say that system (1) has a robust universal finite-dimensional filter if, for each initial probability density σ_0, there exists a z_0 such that (14) and (15) hold if $z(0) = z_0$ and μ_i, ω_i are independent of σ_0.

The method of Wei and Norman (1964) of using Lie algebraic ideas to solve time-varying linear differential equations is roughly as follows. Consider the equation

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

where X and A_is are $n \times n$ matrices and the a_is are scalar-valued functions. Let $B_1, \dots, B_l$ be a basis of the Lie algebra generated by $A_1, \dots, A_m$. Then the Wei–Norman theorem states that, locally in t, $X(t)$ has a representation of the form

$$X(t) = e^{b_1(t)B_1} \dots e^{b_l(t)B_l} X_0$$

where the b_is satisfy an ordinary differential equation of the form

$$\frac{db_i}{dt} = c_i(b_1, \dots, b_l), \qquad b_i(0) = 0, \quad 1 \le i \le l$$

The functions c_i, $1 \le i \le n$ in the above equation are determined by the structure constants of the Lie algebra (generated by the A_is) relative to the basis $\{B_1, \dots, B_l\}$.

The extension of Wei and Norman's approach to the non-linear filtering problem is much more complicated. Instead of an ordinary differential equation, we have to solve the robust DMZ equation, which is a time-varying differential equation.

Suppose that the Wei–Norman theory is applied to solve partial differential equations of the form

$$\frac{\partial u}{\partial t} = a_1A_1u + \dots + a_mA_mu \tag{16}$$

where the A_i, $1 \le u \le m$, are linear partial differential operators in $x_1, \dots, x_n$, and the a_i, $1 \le i \le m$, are given functions of time t. The idea is to solve (16) in terms of solutions of the simpler equations

$$\frac{\partial u}{\partial t} = A_iu, \qquad 1 \le i \le m \tag{17}$$

which we write as

$$u(t,x) = \mathrm{e}^{A_i t}\psi(x) \qquad \psi(x) = u(0,x) \tag{18}$$

We shall assume that the Lie algebra generated by the operators $A_1, \ldots, A_m$ in (16) is finite dimensional. By setting, if necessary, some of the $a_i(t)$ equal to zero, and by combining other $a_j(t)$ in case of linear dependence among the operators on the r.h.s. of (16), without loss of generality, we can assume that we are dealing with equation (16) with the additional property that

$$[A_i, A_j] = \sum_k \gamma_{ij}^k \cdot A_k, \ldots, i,j = 1, \ldots, m \tag{19}$$

for suitable real constants γ_{ij}^k, $1 \le i,j,k \le m$.

The central idea of Wei–Norman theory is now to try for a solution of the form

$$u(t) = \mathrm{e}^{g_1(t)A_1}\ \mathrm{e}^{g_2(t)A_2} \ldots \mathrm{e}^{g_m(t)A_m}\psi \tag{20}$$

where the g_i, $1 \le i \le m$, are still to be determined functions of time. The next step is to insert (20) into (16), to obtain

$$\dot{u} = \dot{g}_1 A_1\ \mathrm{e}^{g_1 A_1} \ldots \mathrm{e}^{g_m A_m}\psi + \mathrm{e}^{g_1 A_1}\dot{g}_2 A_2 \mathrm{e}^{g_2 A_2} \ldots \mathrm{e}^{g_m A_m}\psi + \cdots$$
$$+ \mathrm{e}^{g_1 A} \ldots \mathrm{e}^{g_{m-1}A_{m-1}}\dot{g}_m A_m\ \mathrm{e}^{g_m A_m}\psi \tag{21}$$

Now for $i = 2, \ldots, n$ insert a term

$$\mathrm{e}^{-g_{i-1}A_{i-1}} \ldots \mathrm{e}^{-g_1 A_1}\ \mathrm{e}^{g_1 A_1} \ldots \mathrm{e}^{g_{i-1}A_{i-1}}$$

just behind $\dot{g}_i A_i$ in the ith term of (21). Then use the adjoint representation formula

$$\mathrm{e}^A B \mathrm{e}^{-A} = B + [A,B] + \frac{1}{2!}[A,[A,B]] + \frac{1}{3!}[A,[A,B]] + \cdots \tag{22}$$

and (19) repeatedly, and use the linear independence of the $A_1, \ldots, A_m$ to obtain a system of ordinary differential equations for the $g_1, \ldots, g_m$ (with initial conditions $g_1(0) = 0 = g_2(0) = \cdots = g_m(0)$). These system of ODEs are always solvable for small time. However they may not be solvable for all time, meaning that finite escape time phenomena may occur.

Fortunately, Theorem 13 above will allow us to prove the following theorem which shows in particular how to construct finite dimensional filters from finite-dimensional estimation algebras. Since the estimation algebra is solvable, the corresponding system of ODEs are solvable for all $t \ge 0$. The detail can be found in Yau (1994).

Theorem 14: *Let E be an estimation algebra of* (1) *satisfying $(\partial f_j/\partial x_i) - (\partial f_i/\partial x_i) = c_{ij}$, where the c_{ij}s are constants for all $1 \le i, j \le n$. Suppose that E is a finite dimensional estimation algebra of maximal rank. Then E has a basis of the form* 1, $x_1, \ldots, x_n$, $D_1, \ldots, D_n$, *and* L_0 *and*

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

is a degree two polynomial

$$\sum_{i,j=1}^n a_{ij}x_i x_j + \sum_{i=1}^n b_i x_i + d$$

The robust DMZ equation (3) *has a solution for all $t \ge 0$ of the form*

$$u(t,x) = \mathrm{e}^{T(t)}\mathrm{e}^{r_n(t)x_n} \ldots \mathrm{e}^{r_1(t)x_1}\mathrm{e}^{s_n(t)D_n} \ldots \mathrm{e}^{s_1(t)D_1}\mathrm{e}^{tL_0}\sigma_0$$

where $T(t), r_1(t), \ldots, r_n(t)$, $s_1(t), \ldots, s_n(t)$ satisfies the ordinary differential equations

$$\frac{\mathrm{d}s_i}{\mathrm{d}t}(t) = r_i(t) + \sum_{i=1}^n s_j(t)c_{ji} + \sum_{k=1}^n h_{ki}y_k(t), \quad 1 \le i \le n \tag{23}$$

$$\frac{\mathrm{d}r_j}{\mathrm{d}t}(t) = \frac{1}{2}\sum_{i=1}^n s_i(t)(a_{ij} + a_{ji}), \quad 1 \le j \le n \tag{24}$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} = -\frac{1}{2}\sum_{i=1}^n r_i^2(t) - \frac{1}{2}\sum_{i=1}^n s_i^2(t)\left(\sum_{j=1}^n c_{ij}^2 - a_{ii}\right) + \sum_{i=1}^n r_i(t) \tag{25}$$
$$-\sum_{j=2}^n \sum_{i=1}^j s_j(t)c_{ij} + \sum_{1 \le i < k \le n} s_i(t)s_k(t)$$
$$\times \left[\sum_{j=1}^n c_{ij}c_{jk} + \frac{1}{2}(a_{ik} + a_{ki})\right]$$
$$+\frac{1}{2}\sum_{i=1}^n s_i(t)b_i + \frac{1}{2}\sum_{i,j=1}^m y_i(t)y_j(t)\sum_{k=1}^n h_{ik}h_{jk}$$
$$-\sum_{i,j=1}^n s_i(t)r_j(t)c_{ij}$$

where $h_k(x) = \sum_{j=1}^n h_{kj}x_j + e_k$, $1 \le k \le m$, h_{kj} and e_k are constants. In particular, a universal finite-dimensional filter exists.

The following theorem in Yau (1994) gives a characterization when the drift term $f(x)$ satisfies the conditions $(\partial f_j/\partial x_i) - (\partial f_i/\partial x_j) = c_{ij}$, where c_{ij} are constants for all $1 \le i,j \le n$.

Theorem 15: *$(\partial f_j/\partial x_i) - (\partial f_i/\partial x_j) = c_{ij}$ are constants for all i and j if and only if*

$$(f_1, \ldots, f_n) = (l_1, \ldots, l_n) + \left(\frac{\partial\psi}{\partial x_i}, \ldots, \frac{\partial\psi}{\partial x_n}\right) \tag{26}$$

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

Proof: It is clear that if (26) is satisfied, then $(\partial f_j/\partial x_i) - (\partial f_i/\partial x_j) = c_{ij}$ are constants for all i and j.

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

$$b_{ji} - b_{ij} = c_{ij}, \qquad 1 \le i,\ j \le n \tag{27}$$

Let $l_i(x) = \sum_{j=1}^{n} b_{ij}x_j$ for $1 \le i \le n$

$$d\left(\sum_{j=1}^{n} f_j\,\mathrm{d}x_j\right) = \sum_{i<j}\left(\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}\right)\mathrm{d}x_i \wedge \mathrm{d}x_j$$

$$= \sum_{i<j} c_{ij}\,\mathrm{d}x_i \wedge \mathrm{d}x_j \tag{28}$$

$$d\left(\sum_{j=1}^{m} l_j\,\mathrm{d}x_j\right) = \sum_{i<j}(b_{ji} - b_{ij})\,\mathrm{d}x_i \wedge \mathrm{d}x_j \tag{29}$$

In view of (27), (28) and (29) we have

$$d\left(\sum_{j=1}^{n} f_i\,\mathrm{d}x_j - \sum_{j=1}^{n} l_j\,\mathrm{d}x_j\right) = 0$$

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

$$\sum_{j=1}^{n} f_j\,\mathrm{d}x_j - \sum_{j=1}^{n} l_j\,\mathrm{d}x_j = \mathrm{d}\psi = \sum_{j=1}^{n}\frac{\partial \psi}{\partial x_j}\,\mathrm{d}x_j \qquad \square$$

5. Structures of quadratic forms and linear structure of Ω-matrix

We shall recall the theory of quadratic forms in estimation algebras developed by Chan and Yau (1996). We first introduce the notion of quadratic rank k for any estimation algebra. This concept plays a fundamental role in the theory of classification of finite dimensional estimation algebras. We show that any quadratic polynomial in the estimation algebra depends on the variables only up to quadratic rank k (cf. Lemma 4).

We show that there is a natural decomposition $\{1, 2, \ldots, k\}$ into disjoint union of S_i, where S_i is described in (39) below. For each S_i, we associate a basic quadratic polynomial p_i (cf. (41) below) in the estimation algebra. We show some important properties of quadratic polynomials in the estimation algebras in terms of this decomposition (cf. Lemmas 5–7). These properties of quadratic polynomials are used to prove the constant structure of the $k \times k$ left upper corner of the Ω matrix (cf. Lemma 10, Theorem 20 and Theorem 21). The proofs given are easier than those in Chen and Yau (1996). Quadratic polynomial properties were also used to prove the constant structure of the $k \times (n-k)$ right upper corner of the Ω matrix (cf. § 7). In § 5, we also develop a new simple proof of linear structure of Ω matrix than those given in Chen and Yau (1996). The proof given here depends on some special properties of partial Euler operators developed in Theorems 16–18.

Let Q be the space of quadratic forms in n variables, that is, real vector space spanned by x_ix_j, with $1 \le i \le j \le n$. Let $X = (x_1, x_2, \ldots, x_n)^{\mathrm{T}}$ and let $M_n(\mathbb{R})$ be the group of $n \times n$ matrices.

Definition 4: For any quadratic form $p \in Q$, there exists a symmatric matrix A such that $p(x) = X^{\mathrm{T}}AX$. The rank of the quadratic form p is denoted by $r(p)$ and is defined to be the rank of the matrix A. A *fundamental quadratic form* of the estimation algebra E is an element $p_0 \in E \cap Q$ with the greatest positive rank, that is, $r(p_0) \ge r(p)$ for any $p \in E \cap Q$. The maximal rank of quadratic forms in the estimation algebra E is defined to be $k = r(p_0)$ and is called the *quadratic rank* of E.

After an orthogonal transformation on x, p_0 can be written as

$$p_0 = c_1x_1^2 + c_2x_2^2 + \cdots + c_kx_k^2, \quad c_i \ne 0, \quad 0 \le k \le n \tag{30}$$

From $p_0(x)$, we can construct a sequence of quadratic forms in $E \cap Q$ as

$$q_0(x) = p_0(x) \tag{31}$$

$$q_j(x) = [[L_0, q_{j-1}], q_0] = \sum_{i=1}^{k} 4^j c_i^{j+1} x_1^2 \tag{32}$$

In view of the invertibility of the Vandermonde matrix, we can assume that

$$p_0(x) = x_1^2 + x_2^2 + \cdots + x_k^2 \in E$$

Lemma 4: *If p is a quadratic form in the estimation algebra E, then p is independent of x_j for $j > k$, where $k = r(p_0)$. In other words, $\partial p/\partial x_j = 0$ for $k+1 \le j \le n$.*

Proof: Suppose on the contrary that $\partial p/\partial x_i \ne 0$ for some $j > k$. Let A be a symmetric matrix such that $p = X^{\mathrm{T}}AX$. A can be written as

$$A = \begin{pmatrix} A_1 & A_2 \\ A_2^{\mathrm{T}} & A_4 \end{pmatrix} \tag{33}$$

where A_1 is a $k \times k$ symmetric matrix and A_4 is an $(n-k) \times (n-k)$ symmetric matrix. There is a $k \times k$ orthogonal matrix S_1 and an $(n-k) \times (n-k)$ orthogonal matrix S_2 such that $S_1^{\mathrm{T}}A_1S_1$ and $S_2^{\mathrm{T}}A_4S_2$ are diagonal matrices. So we can assume that A_1 and A_4 are diagonal matrices. $\partial p/\partial x_j \ne 0$ for some $j > k$ implies $A_2 \ne 0$ or $A_4 \ne 0$. Since

$$r(\lambda p_0 + \sigma p) = \operatorname{rank}\begin{pmatrix} \lambda I + \sigma A_1 & \sigma A_2 \\ \sigma A_2^{\mathrm{T}} & \sigma A_4 \end{pmatrix} \tag{34}$$

if we choose λ large enough, it is easy to see that

$$r(\lambda p_0 + \sigma p) > k \tag{35}$$

This contradicts the greatest positive rank assumption of p_0. □

Let $p_1 \in E \cap Q$ be an element with least positive rank, that is $0 < r(p_1) \le r(q)$ for any non-zero $q \in E \cap Q$. After an orthogonal transform that fixes $x_{k+1}, \ldots, x_n$ variables (i.e. an orthogonal transform on $x_1, x_2, \ldots, x_k$) and the Vandermonde matrix procedure as above, we can assume

$$p_1 = \sum_{i=1}^{k_1} x_1^2 \in E, \qquad 1 \le k_1 \le k \tag{36}$$

Note that the orthogonal transform on $x_1, \ldots, x_k$ leaves p_0 invariant. In summary, we deduce that $p_0 = \sum_{i=1}^{k} x_i^2$ has the greatest positive rank and $p_1 = \sum_{i=1}^{k_1} x_i^2$ has the least positive rank. Define

$$S_1 = \{1, 2, \ldots, k_1\} \subseteq S = \{1, 2, \ldots, k\} \tag{37}$$

and $Q_1 =$ real vector space spanned by $\{x_i x_j : k_1 + 1 \le i \le j \le k\} \subseteq Q$.

If $k_1 < k$, then $Q_1 \cap E$ is a non-trivial space, since $p - p_0 \in E \cup Q$. In a similar procedure as above, there exists

$$p_2 = \sum_{i=k_1+1}^{k_2} x_i^2 \in E \cap Q_1 \tag{38}$$

with the least positive rank in $E \cap Q_1$. By induction, we construct a series of S_i, Q_i and p_i such that

$$S_i = \{k_{i-1} + 1, \ldots, k_i\}, \quad k_0 = 0, \quad k_i \le k \tag{39}$$

and

$Q_i =$ real vector space spanned by

$$\{x_l x_j : k_i + 1 \le l \le j \le k\} \tag{40}$$

$$p_i = \sum_{j=k_{i-1}+1}^{k_i} x_j^2 = \sum_{j \in S_i} x_j^2, \quad i > 0 \tag{41}$$

and p_i has the least positive rank in $E \cap Q_{i-1}$ for $i > 0$.

Lemma 5: *If $p \in E \cap Q$, then*

$$p(0, \ldots, 0 x_{k_{i-1}+1}, \ldots, x_{k_i}, 0, \ldots, 0) = \lambda p_i \quad \textit{for } i > 0$$

Proof: In view of Lemma 1 and the fact that $[L_0, p_i] \in E$, $[L_0, p_0 - p_i] \in E$, we have

$$\sum_{j \in S_i} x_j D_j \in E, \qquad \sum_{j \in S - S_i} x_j D_j \in E \tag{42}$$

Hence

$$\left[\sum_{j \in S_i} x_j D_j, p\right] - \left[\sum_{j \in S - S_i} x_j D_j, \left[\sum_{j \in S_i} x_j D_j, p\right]\right] = 2p(0, \ldots, 0, x_{k_{i-1}+1}, \ldots, x_{k_i}, 0, \ldots, 0) \in E$$

Because p_i has the least positive rank for polynomials in $x_{k_{i-1}+1}, \ldots, x_{k_i}$, there is a λ such that

$$p(0, \ldots, 0, x_{k_{i-1}+1}, \ldots, x_{k_i}, 0, \ldots, 0) = \lambda p_i$$ □

Similarly, we also have the following lemma.

Lemma 6: *if $p \in E \cap Q$, then*

$$p(x_1, \ldots, x_{k_{i-1}}, 0, \ldots, 0, x_{k_i+1}, \ldots, x_k) \in E \quad \textit{for } i > 0$$

Proof: The lemma follows immediately from the formula

$$\begin{aligned} p(x_1, \ldots, x_{k_{i-1}}, 0, \ldots, 0, x_{k_i+1}, \ldots, x_k) &= p - \left[\sum_{j \in S - S_i} x_j D_j, \left[\sum_{j \in S_i} x_j D_j, p\right]\right] \\ &\quad - p(0, \ldots, 0, x_{k_{i-1}+1}, \ldots, x_{k_i}, 0, \ldots, 0) \end{aligned}$$ □

Lemma 7: *Let $p = \sum_{i \in S_{l_1}} \sum_{j \in S_{l_2}} 2a_{ij} x_i x_j \in E$, where $a_{ij} \in \mathbb{R}$ and $l_1 < l_2$. Let $X_i = (x_{k_{i-1}+1}, \ldots, x_{k_i})^{\mathrm{T}}$ be a $(k_1 - k_{i-1})$-vector. Under this notation, p can be written as*

$$p = (X_{l_1}^{\mathrm{T}}, X_{l_2}^{\mathrm{T}}) \begin{pmatrix} 0 & A \\ A^{\mathrm{T}} & 0 \end{pmatrix} \begin{pmatrix} X_{l_1} \\ X_{l_2} \end{pmatrix} \tag{43}$$

Then $|S_{l_1}| = |S_{l_2}|$ and $A = bT$, where b is a constant and T is an orthogonal matrix

Proof: $[L_0, p] = 2 \sum_{i \in S_{l_1}} \sum_{j \in S_{l_2}} a_{ij}(x_i D_j + x_j D_i) \in E$. Hence

$$\begin{aligned} [[L_0, p], p] &= 4 \sum_{i,m \in S_{l_1}} \sum_{j,l \in S_{l_2}} a_{ij} a_{ml} [x_i D_j + x_j D_i, x_m x_l] \\ &= 4 \sum_{i,m \in S_{l_1}} \sum_{j,l \in S_{l_2}} a_{ij} a_{ml} \\ &\quad \times (x_i x_l S_{im} + x_i x_m \delta_{jl} + x_j x_l \delta_{im} + x_j x_m \delta_{il}) \\ &= 4 \sum_{i \in S_{l_1}} \sum_{j,l \in S_{l_2}} a_{ij} a_{jl} x_i x_l \\ &\quad + 4 \sum_{i,m \in S_{l_1}} \sum_{j \in S_{l_2}} a_{ij} a_{mj} x_i x_m \\ &\quad + 4 \sum_{i \in S_{l_1}} \sum_{j,l \in S_{l_2}} a_{ij} a_{il} x_j x_l \\ &\quad + 4 \sum_{i,m \in S_{l_1}} \sum_{j \in S_{l_2}} a_{ij} a_{mi} x_j x_m \end{aligned}$$

Since $[[L_0, p], p] \in E$, from Lemma 5, we have

$$\sum_{i,m \in S_{l_1}} \left(\sum_{j \in S_{l_2}} a_{ij} a_{mj}\right) x_i x_m = \lambda_1 p_{l_1} \tag{44}$$

$$\sum_{j,l \in S_{l_2}} \left(\sum_{i \in S_{l_1}} a_{ij} a_{il}\right) x_j x_l = \lambda_2 p_{l_2} \tag{45}$$

Equations (44) and (45) show that the rows of A are mutually orthogonal and so are the columns. Since for any matrix the row rank is equal to column rank, we have $|S_{l_1}| = |S_{l_2}|$. As the column vectors have the same Euclidean length, it follows that A is a constant multiple of an orthogonal matrix. $\square$

If E is a finite dimensional estimation algebra with maximal rank, then Lemma 3 says that $\omega_{ij} \in E$ is a polynomial of degree at most 2 for all $1 \le i, j \le n$. Let $\omega_{ij}^{(2)}$, $\omega_{ij}^{(1)}$ be the homogeneous part of degree 2, and 1 of ω_{ij} respectively. Then we have the following lemma.

Lemma 8: *Suppose that E is a finite dimensional estimation algebra of maximal rank. Then*

(i) $\omega_{ij}^{(2)}$ *depends only on* $x_1, \dots, x_k$ *for* $i \le k$ *or* $j \le k$

(ii) $\omega_{ij}^{(2)} = 0$ *for* $k+1 \le i,\ j \le n$

(iii) $\dfrac{\partial \omega_{ij}^{(2)}}{\partial x_l} + \dfrac{\partial \omega_{jl}^{(2)}}{\partial x_i} + \dfrac{\partial \omega_{li}^{(2)}}{\partial x_j} = 0$ *for* $1 \le i, j, l \le n$

(iv) $\dfrac{\partial \omega_{ij}^{(1)}}{\partial x_l} + \dfrac{\partial \omega_{jl}^{(1)}}{\partial x_i} + \dfrac{\partial \omega_{li}^{(1)}}{\partial x_j} = 0$ *for* $1 \le i, j, l \le n$

Proof: Since E is finite dimensional of maximal rank and $\omega_{ij} \in E$, it follows that $\omega_{ij}^{(2)} \in E$. Hence $\omega_{ij}^{(2)}$ depends only on $x_1, \dots, x_k$ by Lemma 4. The cyclic conditions of part (iii) and part (iv) of this Lemma follow from the corresponding cyclic conditions

$$\frac{\partial \omega_{ij}}{\partial \omega_l} + \frac{\partial \omega_{jl}}{\partial x_i} + \frac{\partial \omega_{li}}{\partial x_j} = 0 \tag{46}$$

Let $k+1 \le i,\ j \le n$, and $1 \le l \le k$. Then (iii) gives $\partial \omega_{ij}^{(2)} / \partial x_l = 0$. It follows that $\omega_{ij}^{(2)} = 0$ for $k+1 \le i$, $j \le n$. $\square$

The following three theorems are due to Yau and Rasoulian (1999)

Theorem 16: *Let* $E_k = \sum_{j=1}^k x_j(\partial/\partial x_j)$ *be a Euler operator in* $x_1, \dots, x_k$ *variables. Suppose that m is an integer and* ξ *is a* C^∞ *function on* $\mathbb{R}^n$ *such that* $E_k(\xi) + m\xi$ *is a polynomial of degree r, r a positive integer, in* $x_1, \dots, x_k$ *variables with coefficients in* C^∞ *functions of* $x_{k+1}, \dots, x_n$ *variables. If* $r + m \ge 0$, *then* ξ *is a polynomial of degree r in* $x_1, \dots, x_k$ *variables with coefficients in* C^∞ *functions of* $x_{k+1}, \dots, x_n$. *If* $r + m < 0$, *then* ξ *is a polynomial of degree at most* $-m$ *in* $x_1, \dots, x_k$ *variables with coefficients in* C^∞ *functions of* $x_{k+1}, \dots, x_n$.

Proof: First let $r + m \ge 0$, that is, $r + m + 1 > 0$. Also let $D = (\partial/\partial x_1)^{\alpha_1} \cdots (\partial/\partial x_k)^{\alpha_k}$, $\alpha_1 + \cdots + \alpha_k = r + 1$ be a differential operator of order $r+1$. Since $E_k(\xi) + m\xi$ is a polynomial of degree r in $x_1, x_2, \dots, x_k$ variables with coefficients in C^∞-functions of $x_{k+1}, \dots, x_n$ variables, we have $D[E_k(\xi) + m\xi] = 0$. On the other hand, in view of

$$\frac{\partial}{\partial x_i} E_k = E_k \frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_i} \quad \text{for } 1 \le i \le k$$

it is easy to see by induction that

$$\begin{aligned} D[E_k(\xi) + m\xi] &= \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_{k+1}}\right)^{\alpha_{k-1}} \left(\frac{\partial}{\partial x_k}\right)^{\alpha_k} \\ &\quad \times [E_k(\xi) + m\xi] \\ &= \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_{l-1}}\right)^{\alpha_{k-1}} \\ &\quad \times \left[E_k\left(\frac{\partial}{\partial x_k}\right)^{\alpha_k} \xi + (\alpha_k + m)\left(\frac{\partial}{\partial x_k}\right)\xi\right] \\ &= E_k(D\xi) + (\alpha_1 + \cdots + \alpha_k + m) D\xi \end{aligned}$$

So we have $E_k(D\xi) + (r + 1 + m) D\xi = 0$. Observe that

$$\begin{aligned} E_k[x_1^{r+1+m} D\xi] &= (r + 1 + m) x_1^{r+1+m} D\xi + x_1^{r+1+m} E_k(D\xi) \\ &= x_1^{r+1+m} [E_k(D\xi) + (r + 1 + m) D\xi] = 0 \end{aligned}$$

Denote $\phi = x_1^{r+1+m} D\xi$. Because $r + 1 + m > 0$, we have

$$\begin{aligned} &\phi(x_1, \dots, x_k, x_{k+1}, \dots, x_n) - \phi(\epsilon x_1, \dots, \epsilon x_k, x_{k+1}, \dots, x_n) \\ &= \int_\epsilon^1 \frac{\mathrm{d}\phi}{\mathrm{d}t}(t x_1, \dots, t x_k, x_{k+1}, \dots, x_n)\, \mathrm{d}t \\ &= \int_\epsilon^1 \left[x_1 \frac{\partial \phi}{\partial x_1}(t x_1, \dots, t x_k, x_{k+1}, \dots, x_n) + \cdots \right. \\ &\qquad \left. + x_k \frac{\partial \phi}{\partial x_k}(t x_1, \dots, t x_k, x_{k+1}, \dots, x_n)\right] \mathrm{d}t \\ &= \int_\epsilon^1 \frac{1}{t}(E_k \phi)(t x_1, \dots, t x_k, x_{k+1}, \dots, x_n)\, \mathrm{d}t = \int_\epsilon^1 \frac{0}{t} \mathrm{d}t = 0 \end{aligned}$$

for $\epsilon > 0$. Now let $\epsilon \to 0$. Then we get $\phi(x_1, \dots, x_k, x_{k+1}, \dots, x_n) = 0$. This implies that

$$D\xi = \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial x_k}\right)^{\alpha_k} \xi = 0$$

for all $\alpha_1 + \cdots + \alpha_k = r + 1$ and $\alpha_1 \ge 0, \dots, \alpha_k \ge 0$. In other words ξ is a polynomial of degree at most r in $x_1, \dots, x_k$ variables with coefficients in C^∞-functions of $x_{k+1}, \dots, x_n$ variables. Now by two methods we can prove that ξ is a polynomial of degree r. One method is by induction on r and using the same method as above; the other method is by assumption that

$$\xi = \sum_{0 \le i_1 + \cdots + i_k \le s} a_{i_1 \dots i_k}(x_{k+1}, \dots, x_n) x_1^{i_1} \dots x_k^{i_k}, \quad s \le r$$

is a polynomial of degree s, and then using the definition of $E_k(\xi)+m\xi$ and the hypothesis that the last one is a degree r polynomial. We provide the proof using the second method. Let ξ by a polynomial of degree s

$$\begin{aligned}
E_k(\xi)+m\xi &= E_k\left(\sum_{0\le|i|\le s} a_i(x_{k+1},\dots,x_n)x_1^{i_1}\dots x^{i_k}\right)+m \\
&\quad\times\sum_{0\le|i|\le s} a_i(x_{k+1},\dots,x_n)x_1^{i_1}\dots x_k^{i_k} \\
&= \sum_{0<|i|\le s} |i|a_i(x_{k+1},\dots,x_n)x_1^{i_1}\dots x_k^{i_k}+m \\
&\quad\times\sum_{0\le|i|\le s} a_i(x_{k+1},\dots,x_n)x_1^{i_1}\dots x_k^{i_k} \\
&= \sum_{0<|i|\le s} (|i|+m)a_i(x_{k+1},\dots,x_n)x_1^{i_1}\dots x_k^{i_k} \\
&\quad+ma_0(x_{k+1},\dots,x_n) \\
&= \sum_{0\le|i|\le r} b_i(x_{k+1},\dots,x_n)x_1^{i_1}\dots x_k^{i_k}
\end{aligned}$$

where $i=(i_1,\dots,i_k)$ and $|i|=i_1+\dots+i_k$ and $b_i(x_{k+1},\dots,x_n)$ is C^∞. By looking at the coefficients on both sides we see that $s=r$ and $(|i|+m)a_i=b_i$ for all i, $0<|i|\le r$. That is, ξ is a polynomial of degree r in $x_1,\dots,x_k$ variables with coefficients being C^∞ functions in $x_{k+1},\dots,x_n$.

Now let $r+m<0$. In this case m is a negative integer. Let $m=-m'$, $m'>0$. Then $E_k(\xi)+m\xi=E_k(\xi)-m'\xi=P_r$ where P_r is a polynomial of degree r in $x_1,\dots,x_k$ variables with coefficients in C^∞ functions of $x_{k+1},\dots,x_n$. We have

$$\frac{\partial}{\partial x_{i_1}}[E_k(\xi)-m'\xi]=\frac{\partial}{\partial x_{i_1}}P_r=P_{r-1}\qquad 1\le i_1\le k$$

$$\Rightarrow\quad E_k\left(\frac{\partial\xi}{\partial x_{i_1}}\right)-(m'-1)\frac{\partial\xi}{\partial x_{i_1}}=P_{r-1}$$

where P_{r-1} is a polynomial of degree $r-1$. Using the same technique, we get

$$\frac{\partial}{\partial x_{i_2}}\left[E_k\left(\frac{\partial\xi}{\partial x_{i_2}}\right)-(m'-1)\frac{\partial\xi}{\partial x_{i_1}}\right]=\frac{\partial}{\partial x_{i_2}}P_{r-1}=P_{r-2},\quad 1\le i_2\le k$$

$$\Rightarrow\quad E_k\left(\frac{\partial^2\xi}{\partial x_{i_1}\partial x_{i_2}}\right)-(m'-2)\frac{\partial^2\xi}{\partial x_{i_1}\partial x_{i_2}}=P_{r-2}$$

where P_{r-2} is a polynomial of degree $r-2$. After $m'-1$ times, we have

$$E_k\left(\frac{\partial^{m'-1}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'-1}}}\right)-\frac{\partial^{m'-1}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'-1}}}=P_{r-(m'-1)},\qquad 1\le i_{m'-1}\le k$$

where $P_{r-(m'-1)}$ is a polynomial of degree 0 in $x_1,\dots,x_k$ variables, i.e. a C^∞-function in $x_{k+1},\dots,x_n$.

Once more, we have

$$\frac{\partial}{\partial x_{i_{m'}}}\left[E_k\left(\frac{\partial^{m'-1}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'-1}}}\right)-\frac{\partial^{m'-1}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'-1}}}\right]=0\quad 1\le i_{m'}\le k$$

$$\Rightarrow\quad E_k\left(\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}\right)=0$$

Now let $\epsilon>0$. By the same technique we have

$$\begin{aligned}
&\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}(x_1,\dots,x_k,x_{k+1},\dots,x_n) \\
&\quad-\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}(\epsilon x_1,\dots,\epsilon x_k,x_{k+1},\dots,x_m) \\
&\quad=\int_\epsilon^1\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}(tx_1,\dots,tx_k,x_{k+1},\dots,x_n)\right]\mathrm{d}t \\
&\quad=\int_\epsilon^1\frac{1}{t}E_k\left[\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}(tx_1,\dots,tx_k,x_{k+1},\dots,x_n\right]\mathrm{d}t \\
&\quad=\int_\epsilon^1\frac{0}{t}\,\mathrm{d}t=0
\end{aligned}$$

Let $\epsilon\to 0$. Then

$$\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}(x_1,\dots,x_k,x_{k+1},\dots,x_n)=\frac{\partial^{m'}\xi}{\partial x_{i_1}\dots\partial x_{i_{m'}}}(0,\dots,0,x_{k+1},\dots,x_n)$$

The right-hand side is a function of $x_{k+1},\dots,x_n$. This means that $\partial^{m'-1}\xi/\partial x_{i_1}\dots\partial x_{i_{m'-1}}$ is a linear function of $x_1,\dots,x_k$ with coefficients in C^∞-functions of $x_{k+1},\dots,x_n$. Now by induction, we conclude that ξ is a polynomial of degree at most m' in $x_1,\dots,x_k$ variables with coefficients in C^∞-functions of $x_{k+1},\dots,x_n$. $\square$

Theorem 17: *Let $E_k=x_1(\partial/\partial x_i)+\dots+x_k(\partial/\partial x_k)$ be an Euler operator in $x_1,\dots,x_k$ variables. Suppose that m is a positive constant and ξ is a C^∞ function on $\mathbb{R}^n$ such that $E_k(\xi)+m\xi$ is a polynomial of degree r in $x_1,\dots,x_n$ variables. Then ξ is a polynomial of degree r in $x_1,\dots,x_n$ variables.*

Proof: By Theorem 16, $\xi=\sum_{0\le|\alpha|\le r}a_\alpha(x_{k+1},\dots,x_n)x_1^{\alpha_1}\dots x_k^{\alpha_k}$, where $\alpha=(\alpha_1,\dots,\alpha_k)$ and $|\alpha|=\alpha_1+\dots+\alpha_k$ and $a_\alpha(x_{k+1},\dots,x_n)$ is C^∞.

Hence we have

$$
\begin{aligned}
E_k(\xi) + m\xi &= \sum_{0<|\alpha|\le r} |\alpha| a_\alpha(x_{k+1},\dots,x_n) x_1^{\alpha_k}\dots x_k^{\alpha_k} \\
&\quad + m \sum_{0\le|\alpha|\le r} a_\alpha(x_{k+1},\dots,x_n) x_1^{\alpha_1}\dots x_k^{\alpha_k} \\
&= \sum_{0<|\alpha|\le r} (|\alpha|+m) a_\alpha(x_{k+1},\dots,x_n) x_1^{\alpha_1}\dots x_k^{\alpha_k} \\
&\quad + m a_0(x_{k+1},\dots,x_n) \\
&= \sum_{0\le|\alpha|\le r} p_\alpha(x_{k+1},\dots,x_n) x_1^{\alpha_1}\dots x_k^{\alpha_k}
\end{aligned}
$$

where $p_\alpha(x_{k+1},\dots,x_n)$s are polynomials in $x_{k_1},\dots,x_n$ (because $E_k(\xi)+m\xi$ is a polynomial in $x_1,\dots,x_n$, so we may assume that it is a polynomial in $x_1,\dots,x_k$ with coefficients being polynomials in $x_{k+1},\dots,x_n$). Now, looking at both sides, we conclude that $(|\alpha|+m)a_\alpha = p_\alpha$, for all $\alpha=(\alpha_1,\dots,\alpha_k)$, $0<|\alpha|\le r$; in other words all a_α, $0<|\alpha|\le r$ are polynomials and also $a_0 = (1/m)p_0$ is a polynomial, and hence ξ is a polynomial. $\square$

Remark: Theorem 17 is false if $m=0$. It is possible that $E_k(\xi)$ is a polynomial of degree r in $x_1,\dots,x_n$ variables, but ξ is not a degree r polynomial in $x_1,\dots,x_n$ variables. For example, we can simply take ξ to be any degree r polynomial in $x_1,\dots,x_n$ variables plus a transcendental function in $x_{k+1},\dots,x_n$ variables.

Theorem 18: *Let*

$$E_k = x_1\frac{\partial}{\partial x_1} + \cdots + x_k\frac{\partial}{\partial x_k}$$

be an Euler operator in $x_1,\dots,x_k$ variables. Suppose that ξ is a C^∞ function on $\mathbb{R}^n$ such that $E_k(\xi)$ is a polynomial of degree r in $x_1,\dots,x_n$ variables. Then $\xi = P_r(x_1,\dots,x_n) + a(x_{k+1},\dots,x_n)$ where $P_r(x_1,\dots,x_n)$ is a polynomial of degree r and $a(x_{k+1},\dots,x_n)$ is a C^∞ function in $x_{k+1},\dots,x_n$.

Proof: In view of Theorem 16, $\xi = \sum_{0\le|\alpha|\le r} a_\alpha(x_{k+1},\dots,x_n)x_1^{\alpha_1}\dots x_k^{\alpha_k}$, where $\alpha=(\alpha_1,\dots,\alpha_k)$ and $|\alpha| = \alpha_1+\cdots\alpha_k$ and $a_\alpha(\alpha_{k+1},\dots,x_n)$ is C^∞. Then $E_k(\xi) = \sum_{0<|\alpha|\le r}|\alpha|a_\alpha(x_{k+1},\dots,x_n)x_1^{\alpha_1}\dots x_k^{\alpha_k}$, which is a polynomial of degree r in $x_1,\dots,x_n$ variables. Therefore $a_\alpha(x_{k+1},\dots,x_n)$ for $|\alpha|\ge 1$, are polynomials. Theorem 18 follows immediately. $\square$

Lemma 9: *Let E be a finite-dimensional estimation algebra of maximal rank. Let k be the quadratic rank of E. For $1\le i$, $j\le n$, ω_{ij} and $\alpha_i = \sum_{j=1}^k x_j\omega_{ij}\in E$ are polynomials of degree 2 in $x_1,\dots,x_n$ variables. Furthermore, we have the following relationships:*

(i) $E_k(\omega_{ij}) + 2\omega_{ij} = \dfrac{\partial\alpha_i}{\partial x_j} - \dfrac{\partial\alpha_j}{\partial x_i}, \quad \forall 1\le i,j\in k;$

(ii) $E_k(\omega_{ij}) + \omega_{ij} = \dfrac{\partial\alpha_i}{\partial x_j} - \dfrac{\partial\alpha_j}{\partial x_i}, \quad \forall 1\le i\le k,\ k+1\le j\le n;$

(iii) $E_k(\omega_{ij}) + \omega_{ij} = \dfrac{\partial\alpha_i}{\partial x_j} - \dfrac{\partial\alpha_j}{\partial x_i}, \quad \forall 1\le j\le k, k+1\le i\le n;$

(iv) $E_k(\omega_{ij}) = \dfrac{\partial\alpha_i}{\partial x_j} - \dfrac{\partial\alpha_j}{\partial x_i}, \quad \forall k+1\le i,j\le n$

Proof: By Lemma 2, we have $\omega_{ij}\in E$ and $\alpha_i = \frac{1}{2}[[L_0,D_j],p_0]\in E$ where p_0 is defined by (11). Theorem 2 implies that ω_{ij} and α_i are polynomials of degree 2 in $x_1,\dots,x_2$ variables. The relationships (i)–(iv) follow immediately from the definition of $E_k(\omega_{ij})$ and α_i. For example, we give the proof of (i) here

$$\frac{\partial\alpha_i}{\partial x_j} = \sum_{l=1}^k \frac{\partial(x_l\omega_{il})}{\partial x_j} = \omega_{ij} + \sum_{l=1}^k x_l\frac{\partial\omega_{il}}{\partial x_j}$$

$$\frac{\partial\alpha_j}{\partial x_i} = \sum_{l=1}^k \frac{\partial(x_l\omega_{jl})}{\partial x_i} = \omega_{ji} + \sum_{l=1}^k x_l\frac{\partial\omega_{jl}}{\partial x_i}$$

$$
\begin{aligned}
\frac{\partial\alpha_j}{\partial x_i} - \frac{\partial\alpha_i}{\partial x_j} &= 2\omega_{ji} + \sum_{l=1}^k x_l\left(\frac{\partial\omega_{jl}}{\partial x_i} - \frac{\partial\omega_{il}}{\partial x_j}\right) \\
&= 2\omega_{ji} + \sum_{l=1}^k x_l\left(\frac{\partial\omega_{jl}}{\partial x_i} + \frac{\partial\omega_{li}}{\partial x_j}\right) \\
&= 2\omega_{ji} + \sum_{l=1}^k x_l\frac{\partial\omega_{ji}}{\partial x_l} = 2\omega_{ji} + E_k(\omega_{ji})
\end{aligned}
$$
$\square$

Corollary 4: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then*

$$\Omega = (\omega_{ij}) = \left(\begin{array}{c|c} P_1(x_1,\dots,x_n) & P_1(x_1,\dots,x_n) \\ \hline P_1(x_1,\dots,x_n) & P_1(x_1,\dots,x_n) + P_2(x_{k+1},\dots,x_n) \end{array}\right)$$

i.e. ω_{ij}s are polynomials of degree 1 in $x_1,\dots,x_n$ variables for $1\le i\le k$ or $1\le j\le k$ and ω_{ij} are polynomials of degree 1 in $x_1,\dots,x_n$ variables plus polynomials of degree 2 in $x_{k+1},\dots,x_n$ variables for $k+1\le i$, $j\le n$.

Proof: This follows from Theorems 17 and 18 and Lemma 9 $\square$

Theorem 19: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then*

$$\Omega = (\omega_{ij}) = \left(\begin{array}{c|c} P_1(x_1,\dots,x_k) & P_1(x_1,\dots,x_k) \\ \hline P_1(x_1,\dots,x_k) & P_1(x_{k+1},\dots,x_n) \end{array}\right)$$

i.e.

(i) *ω_{ij} is a polynomial of degree 1 in $x_1,\dots,x_k$ for $1\le i\le k$ or $1\le j\le k$*

(ii) *ω_{ij} is a polynomial degree 1 in $x_{k+1},\dots,x_n$ for $k+1 \le i,j \in n$.*

Proof: Since $\alpha_i = \sum_{j=1}^{k} x_j\omega_{ij}$ is a quadratic polynomial in E by Lemma 9, it cannot depend on $x_{k+1},\dots,x_n$ variables for $1 \le i \le n$ according to Lemma 4, (i) follows immediately. If $k+1 \le i,\ j \le n$, by using the cyclic relationship

$$\frac{\partial\omega_{ij}}{\partial x_l} + \frac{\partial\omega_{li}}{\partial x_j} + \frac{\partial\omega_{jl}}{\partial x_i} = 0$$

we have $\partial\omega_{ij}/\partial x_l = 0$ for $1 \le l \le k$. This means that ω_{ij} are independent of $x_1,\dots,x_k$ for $k+1 \le i,\ j \le n$. Now $\omega_{ij} = p_1(x_{k+1},\dots,x_n) + p_2(x_{k+1},\dots,x_n)$ for $k+1 \le i$, $j \le n$. Since $\omega_{ij}^{(2)} \in E$ as a quadratic polynomial in E cannot depend on $x_{k+1},\dots,x_n$ variables for $k+1 \le i$, $j \le n$ according to Theorem 2, it follows that $p_2(x_{k+1},\dots,x_n) = 0$. □

Lemma 10: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the same notation as in* (39), *if*

$$\sum_{i\in S_l} x_i\alpha_i = 0 \tag{47}$$

where α_is are homogeneous polynomials of degree 2 *in E, then $\alpha_i = 0$ for all $i \in S_l$.*

Proof: Let $X_i = (x_{k_{i-1}+1}, x_{k_{i-1}+2},\dots,x_{k_i})^{\mathrm T}$ and $X = x_1, x^2,\dots,x_n)^{\mathrm T}$. Without loss of generality, we assume that $l = 1$. Let $X^{\mathrm T} = (X_1^{\mathrm T}, \bar X_1^{\mathrm T})$ where $\bar X_1$ is the complementing variable of X_1 in X. Write

$$\alpha_i(X) = \alpha_i(X_1,0) + \alpha_i(0,\bar X_1) + [\alpha_i - \alpha_i(X_1,0) - \alpha_i(0,\bar X_1)] \tag{48}$$

Hence (47) is still true if we replace α_i in (47) by one of the three terms on the right-hand side of (48). We see immediately that

$$\alpha_i(0,\bar X_i) = 0 \qquad \forall i \in S_1 \tag{49}$$

By Lemma 5, we have

$$\alpha_i(X_1,0) = \lambda_i p_1 \tag{50}$$

So the corresponding equation of (47) for $\alpha_i(X_1,0)$ gives

$$\sum_{i\in S_1} x_i\lambda_i p_1 = 0 \tag{51}$$

It follows that $\lambda_i = 0$, that is,

$$\alpha_i(X_1,0) = 0 \qquad \forall i \in S_1 \tag{52}$$

Finally, $\alpha_i - \alpha_i(X_1,0) - \alpha_i(0,\bar X_1)$ is a sum of $2X_1^{\mathrm T}R_{il}X_l$ for $l \ge 2$ and R_{il} is a constant multiple of an orthogonal matrix. Therefore the corresponding equation of (47) for $\alpha_i - \alpha_i(X_1,0) - \alpha_i(0,\bar X_1)$ gives

$$\sum_{l\ge 2} X_1^{\mathrm T}\left(\sum_{i\in S_1} 2x_iR_{il}\right)X_l = \sum_{i\in S_1} x_i \sum_{l\ge 2} 2X_1^{\mathrm T}R_{il}X_l = 0 \tag{53}$$

This implies

$$X_1^{\mathrm T}\left(\sum_{i\in S_1} 2x_iR_{il}\right) = 0 \qquad \forall l \ge 2 \tag{54}$$

Fix $i_0 \in S_1$, and let $x_{i_0} = 1$ and $x_i = 0$ for $i \ne i_0$. Then (54) becomes

$$(0,\dots,0,1,0,\dots,0)R_{i_0l} = 0 \qquad \forall l \ge 2 \tag{55}$$

Since R_{i_0l} is a constant multiple of an orthogonal matrix, we see that $R_{i_0l} = 0$, $\forall l \ge 2$. This is true for all $i_0 \in S_1$. Thus

$$\alpha_i - \alpha_i(X_1,0) - \alpha_i(0,\bar X_1) = 0 \tag{56}$$

So we have proved $\alpha_i = 0$ by (49), (52) and (56) □

Theorem 20: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the same notation as in* (39), *if $p \ne q$ and $i \in S_p$, $j \in S_q$, then ω_{ij} is a constant.*

Proof: Recall that from (42), we have $\sum_{i\in S_p} x_iD_i$ and $\sum_{j\in S_q} x_jD_j$ in E. Hence

$$\sum_{i\in S_p}\sum_{j\in S_q} x_ix_j\omega_{ij} = -\left[\sum_{i\in S_p} x_iD_i, \sum_{j\in S_q} x_jD_j\right] \in E \tag{57}$$

In view of Theorems 2 and 19, equation (57) implies

$$\sum_{i\in S_p}\sum_{j\in S_q} x_ix_j\omega_{ij}^{(1)} = \sum_{i\in S_p} x_i\left(\sum_{j\in S_q} x_j\omega_{ij}^{(1)}\right) = \sum_{j\in S_q} x_j\left(\sum_{i\in S_p} x_i\omega_{ij}^{(1)}\right) = 0 \tag{58}$$

Hence $\omega_{ij}^{(1)}$ depends only on x_m, where $m \in S_p \cup S_q$ for $i \in S_p$ and $j \in S_q$. Since E is of maximal rank, $D_j \in E$ for any j. In particular, $[\sum_{i\in S_p} x_iD_i, D_j] \in E$ for $j \in S_q$, and $[\sum_{i\in S_q} x_jD_j, D_i] \in E$ for $i \in S_p$. In view of (iii) of Lemma 1, we have

$$\sum_{i\in S_p} x_i\omega_{ij}^{(1)} \in E \quad \text{for } j \in S_q \quad \text{and} \quad \sum_{j\in S_q} x_j\omega_{ij}^{(1)} \in E \quad \text{for } i \in S_p \tag{59}$$

Equations (58), (59) and Lemma 10 simply

$$\sum_{i\in S_p} x_i\omega_{ij}^{(1)} = 0 \quad \text{for } j \in S_q \quad \text{and} \quad \sum_{j\in S_q} x_j\omega_{ij}^{(1)} = 0 \quad \text{for } i \in S_p \tag{60}$$

The first equation of (60) says that, for $i \in S_p$ and $j \in S_q$, $\omega_{ij}^{(1)}$ does not depend on the variable x_m for $m \in S_q$. The second equation of (60) says that, for $i \in S_p$ and $j \in S_q$, $\omega_{ij}^{(1)}$ does not depend on the variable x_m for $m \in S_p$. Hence $\omega_{ij}^{(1)} = 0$ $\square$

Theorem 21: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the same notation as in (39), if $i, j \in S_l$, then ω_{ij} is a constant.*

Proof: Without loss of generality, we shall assume that $l = 1$. For $1 \le i \le k_1$, $\alpha_i = \sum_{j=1}^{k} x_j \omega_{ij}$ is in E by Lemma 9. In view of Theorem 20, we have

$$\alpha_i = \sum_{j=1}^{k_1} x_j \omega_{ij} \in E \quad \Rightarrow \quad \alpha_i(x_k, \ldots, x_{k_1} 0, \ldots, 0)$$

$$= \sum_{j=1}^{k_1} x_j \omega_{ij}(x_1, \ldots, x_{k_1}, 0, \ldots, 0) \in E \tag{61}$$

Since ω_{ij} is a degree one polynomial in $x_1, \ldots, x_k$ for $1 \le i, j \le k_1$, we can write

$$\omega_{ij}^{(1)} = \sum_{l=1}^{k} A_l(i,j) x_l \tag{62}$$

Equations (61) and (62) imply $\sum_{l,j=1}^{k_1} x_j x_l A_l(i,j) \in E$ for $1 \le i, j \le k_1$. By Lemma 5, $\sum_{l,j=1}^{k_1} x_j x_l A_l(i,j) = \lambda \sum_{i=1}^{k_1} x_i^2$. This implies

$$A_l(i,j) = 0 \quad \text{for} \quad 1 \le l \ne j < k_1, \qquad 1 \le i \le k_1 \tag{63}$$

and

$$A_1(i,1) = A_2(i,2) = \cdots = A_{k_1}(i,k_1) \tag{64}$$

We claim that all the terms in (64) are also zero. Choose l so that $1 \le l \le k_1$ and $l \ne i$. Then $A_l(i,l) = -A_l(l,i) = 0$ by (63). In view of (64) and (63), we have

$$A_l(i,j) = 0 \quad \text{for } 1 \le l, i, j \le k_1 \tag{65}$$

Observe that $A_l(i,j) = \partial \omega_{ij}^{(1)} / \partial x_l$. Therefore (iv) of Lemma 8 implies

$$A_l(i,j) + A_j(l,i) + A_i(j,l) = 0$$
$$\text{for } 1 \le i, j \le k_1, \quad k_1 + 1 \le l \le k$$

Since $A_j(l,i) = \partial \omega_{li}^{(1)} / \partial x_j = 0$ and $A_i(j,l) = \partial \omega_{jl}^{(1)} / \partial x_i = 0$ by Theorem 20, we have

$$A_l(i,j) = 0 \quad \text{for } 1 \le i, j \le k_1, \quad k_1 + 1 \le l \le k \tag{66}$$

Therefore we have shown that $\omega_{ij}^{(1)} = 0$ for $1 \le i$, $j \le k_1$. $\square$

Theorem 22: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then*

$$\Omega = (\omega_{ij}) = \left(\begin{array}{c|c} \text{Constants} & P_1(x_1, \ldots, x_k) \\ \hline P_1(x_1, \ldots, x_k) & P_1(x_{k+1}, \ldots, x_n) \end{array} \right)$$

(i) *ω_{ij} is a constant for $1 \le i, j \le k$,*

(ii) *ω_{ij} is a polynomial of degree one in $x_1, \ldots, x_k$ for $1 \le i \le k$, $k+1 \le j \le n$ or $k+1 \le i \le n$, $1 \le j \le k$*

(iii) *ω_{ij} is a polynomial of degree one in $x_{k+1}, \ldots, x_n$ for $k+1 \le i, j \le n$.*

Proof: This is an immediate consequence of Theorems 19, 20 and 21. $\square$

6. Hessian matrix non-decomposition theorem

In this section, we are going to prove that ω_{ij} is a constant for $k+1 \le i$, $j \le n$. We shall see that this statement follows from the weak Hessian matrix non-decomposition theorem which is a general theorem and has nothing to do with estimation algebras. The weak Hessian matrix non-decomposition theorem was first proved by Wu *et al.* (2002). In this section, we shall prove the Hessian matrix non-decomposition theorem, which is a stronger result than weak Hessian matrix non-decomposition theorem.

Lemma 11: *Suppose that E is a finite dimensional estimation algebra of maximal rank. Then*

(i) $\displaystyle\sum_{l=1}^{n} \omega_{jl}\omega_{il} - \frac{1}{2}\frac{\partial^2 \eta}{\partial x_j x_i} \in E$ *for any $1 \le i$, $j \le n$*

(ii) *η is a polynomial of degree 4.*

Proof: (i) follows from (vi) of Lemma 2 and Theorem 19. From (i) and Theorem 19 $\partial^2\eta/\partial x_i \partial x_j$ is a degree two polynomial for all $1 \le i$, $j \le \eta$. Therefore η is a polynomial of degree 4.

Lemma 12: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Let k be the quadratic rank. Let $\eta = \eta_4(x_{k+1}, \ldots, x_n) +$ polynomial of degree 3 in $x_{k+1}, \ldots, x_n$ variables with coefficients degree at most 4 polynomials in $x_1, \ldots, x_k$ variables. Then for any $k+1 \le i, j \le n$*

$$\sum_{l=k+1}^{n} \omega_{jl}^{(1)} \omega_{il}^{(1)} = \frac{1}{2}\frac{\partial^2 \eta_4}{\partial x_j \partial x_i}$$

where $\eta_4 = \eta_4(x_{k+1}, \ldots, x_n)$ is a homogeneous polynomials of degree 4 in $x_{k+1}, \ldots, x_n$ variables.

Proof: From Theorem 22 and Lemma 11, we know that for $k+1 \le i, j \le n$

$$\sum_{l=k+1}^{n} \omega_{jl}^{(1)} \omega_{il}^{(1)} - \frac{1}{2}\frac{\partial^2 \eta_4}{\partial x_j \partial x_i}$$

is the homogeneous polynomial of degree 2 part of

$$\sum_{l=1}^{n} \omega_{jl}\omega_{il} - \frac{1}{2}\frac{\partial^2 \eta}{\partial x_j \partial x_i}$$

in $x_{k+1}, \ldots, x_n$ variables. The result follows immediately from Lemma 4. □

The following notations and Lemma 13 were used and observed by Chen *et al.* (1997). Define

$$\begin{aligned}\Delta := (\omega^{(1)})_{il}, k+1 \leq i,\ l \leq n, \text{an } (n-k) \\ \times (n-k) \text{ anti-symmetric matrix} \\ = \sum_{j=k+1}^{n} A_j x_j\end{aligned}$$

where $A_j = (A_j(p,q)),\ k+1 \leq p,\ q \leq n,$ are $(n-k)\times(n-k)$ anti-symmetric matrix with constant coefficients. The anti-symmetry of Δ and A_j follows directly from that of Ω.

Lemma 13: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the notations as above, then*

(i) $\Delta\Delta^{\mathrm{T}} = \frac{1}{2}H(\eta_4)$, *where* $H(\eta_4) = (\partial^2\eta_4/\partial x_i \partial x_j)$, $k+1 \leq i,\ j \leq n$, *is the Hessian matrix of* $\eta_4 = \eta_4(x_{k+1}, \ldots, x_n)$.

(ii) $A_i(j,l) + A_l(i,j) + A_j(l,i) = 0$.

Proof: (i) follows from Lemma 12 while (ii) is a consequence of Lemma 8 (iv). □

The following weak Hessian matrix non-decomposition theorem is a general mathematical theorem which has independent interest besides non-linear filtering theory. For a $(n-k)\times(n-k)$ matrix with $n-k$ less than or equal to 4, the theorem was proved in Chen *et al.* (1997).

Theorem 23: *Let* $\Delta = \sum_{j=k+1}^{n} A_j x_j$ *be an* $(n-k)\times(n-k)$ *anti-symmetric matrix where* $A_j = (A_j(p,q))$, $k+1 \leq p,\ q \leq n$, *is an anti-symmetric matrix with constant coefficients. Suppose*

$$A_i(j,l) + A_l(i,j) + A_j(l,i) = 0 \quad \text{for all } k+1 \leq i,j,l \leq n$$

Let $\eta_4 = \eta_4(x_{k+1}, \ldots, x_n)$ *be a homogeneous polynomial of degree* 4 *in* $x_{k+1}, \ldots, x_n$. *Let* $H(\eta_4) = (\partial^2\eta_4/\partial x_i \partial x_j)$, $k+1 \leq i,\ j \leq n$, *be the Hessian matrix of* η_4. *If* $\Delta\Delta^{\mathrm{T}} = \frac{1}{2}H(\eta_4)$, *then* $\Delta \equiv 0$, *i.e.* $A_j = 0$ *for all* $k+1 \leq j \leq n$.

The weak Hessian matrix non-decomposition theorem is a consequence of the following Hessian matrix non-decomposition theorem.

Theorem 24: *Let* $\eta_4(x_1, \ldots, x_n)$ *be a homogeneous polynomial of degree* 4 *in* $x_1, \ldots, x_n$ *over* $\mathbf{R}$. *Let* $H(\eta_4) = (\partial^2\eta_4/\partial x_i \partial x_j)_{1 \leq i,\ j \leq n}$ *be the Hessian matrix of* η_4. *Then* $H(\eta_4)$ *cannot be decomposed as* $\Delta(x)\Delta(x)^{\mathrm{T}}$, *where* $\Delta(x) = (\beta_{ij})_{1 \leq i,\ j \leq n}$ *is an anti-symmetric matrix with* β_{ij} *linear functions in* x, *unless* η_4 *and* Δ *are trivial, i.e.* $H(\eta_4)(x) = \Delta(x)\Delta(x)^{\mathrm{T}}$ *implies* $\Delta = 0$ *and* $\eta_4 = 0$.

Let us write $\Delta(x) = A_1x_1 + A_2x_2 + \cdots + A_nx_n$ where A_l is a $n \times n$ antisymmetric matrix with real constant coefficients. Then the equation $H(\eta_4)(x) = \Delta(x)\Delta(x)^{\mathrm{T}}$ will give us a lot of quadratic equations in $A_l(i,j)$ ((i,j) entry of the matrix A_l.), $1 \leq i,j,l \leq n$. Although it is possible to prove that these quadratic equations can have only trivial solution for $n \leq 4$ (see Chen *et al.* (1997), pp. 1137–1138), it has been a challenging problem to algebraic geometors whether this system of quadratic equations in $A_l(i,j)$ can only admit trivial solution over $\mathbf{R}$ even for $n = 5$.

To prove Theorem 24, we need two lemmas.

Lemma 14: *Let* $\eta_4(x_1, \ldots, x_n)$ *be a homogeneous polynomial of degree* 4 *in* $x_1, \ldots, x_n$ *over* $\mathbf{R}$. *Let* $H(\eta_4) = (\partial^2\eta_4/\partial x_i \partial x_j)_{1 \leq i,\ j \leq n}$ *be the Hessian matrix of* η_4. *Let* $\Delta(x) = (\beta_{ij})_{1 \leq i,\ j \leq n} := A_1x_1 + \cdots + A_nx_n$ *where* $A_l = (A_l(i,j))_{1 \leq i,j \leq n}$ *are* $n \times n$ *antisymmetric matrices with coefficient in* $\mathbf{R}$. *Suppose that* $H(\eta_4)(x) = \Delta(x)\Delta(x)^{\mathrm{T}}$. *Then*

$$\begin{aligned}\sum_{l=1}^{n}[A_i(j,l)]^2 &= \sum_{l=1}^{n}[A_j(i,l)]^2 \\ &= \frac{1}{2}\sum_{l=1}^{n}[A_i(i,l)A_j(j,l) + A_i(j,l)A_j(i,l)] \qquad (67)\end{aligned}$$

Proof: Observe that $H(\eta_4)(x) = \Delta(x)\Delta(x)^{\mathrm{T}}$ implies

$$\frac{\partial^2 \eta}{\partial x_i \partial x_j} = \sum_{l=1}^{n} \beta_{il}\beta_{jl} \qquad (68)$$

Since

$$\frac{\partial^2}{\partial x_i^2}\left(\frac{\partial^2 \eta}{\partial x_j^2}\right) = \frac{\partial^2}{\partial x_j^2}\left(\frac{\partial^2 \eta}{\partial x_i^2}\right) = \frac{\partial^2}{\partial x_i \partial x_j}\left(\frac{\partial^2 \eta}{\partial x_i \partial x_j}\right)$$

we have

$$\frac{\partial^2}{\partial x_i^2}\left(\sum_{l=1}^{n}\beta_{jl}^2\right) = \frac{\partial^2}{\partial x_j^2}\left(\sum_{l=1}^{n}\beta_{il}^2\right) = \frac{\partial^2}{\partial x_i \partial x_j}\left(\sum_{l=1}^{n}\beta_{il}\beta_{jl}\right)$$

Notice that β_{ij} is linear in $x_1, \ldots, x_n$ for $1 \leq i,j \leq n$. This leads to

$$\begin{aligned}2\sum_{l=1}^{n}\left(\frac{\partial \beta_{jl}}{\partial x_i}\right)^2 &= 2\sum_{l=1}^{n}\left(\frac{\partial \beta_{il}}{\partial x_j}\right)^2 \\ &= \sum_{l=1}^{n}\left(\frac{\partial \beta_{il}}{\partial x_i}\frac{\partial \beta_{jl}}{\partial x_j} + \frac{\partial \beta_{il}}{\partial x_j}\frac{\partial \beta_{jl}}{\partial x_i}\right) \qquad (69)\end{aligned}$$

As $A_i(j,l) = \partial\beta_{jl}/\partial x_i$, we see that (67) is equivalent to (69) □

Lemma 15: *Let $\eta(x)$ be a C^∞ function of $\mathbf{R}^n$. Let $\tilde{\eta}(x) = \eta(Rx)$ where R is a $n \times n$ matrix. Then $H(\tilde{\eta})(x) = R^{\mathrm{T}} H(\eta)(Rx)R$.*

Proof: Let $y = Rx$ where $r = (r_{ij})$ is a $n \times n$ matrix. Then by chain rule, we have

$$\frac{\partial \tilde{\eta}}{\partial x_i(x)} = \sum_{p=1}^n \frac{\partial \eta}{\partial y_p}(Rx)\frac{\partial y_p}{\partial x_i} = \sum_{p=1}^n r_{pi}\frac{\partial \eta}{\partial y_p}(Rx)$$

$$\begin{aligned}\frac{\partial^2 \tilde{\eta}}{\partial x_i \partial x_j}(x) &= \sum_{p=1}^n r_{pi}\frac{\partial}{\partial x_j}\left[\frac{\partial \eta}{\partial y_p}(Rx)\right] \\ &= \sum_{p=1}^n r_{pi}\sum_{q=1}^n \frac{\partial^2 \eta}{\partial y_p \partial y_q}(Rx)\frac{\partial y_q}{\partial x_j} \\ &= \sum_{p,q=1}^n r_{pi}\frac{\partial^2 \eta}{\partial y_p y_q}(Rx) r_{qj}\end{aligned}$$

Therefore $H(\tilde{\eta})(x) = R^{\mathrm{T}} H(\eta)(Rx)R$. □

We are now ready to prove our main theorem by induction on n. For $n = 1$, the theorem is trivially true. For $n = 2$, by the antisymmetry of the matrice of A_1 and A_2, we only need to show that $A_1(2,1) = 0 = A_2(1,2)$. But this follows immediately from (67) with $(i,j) = (1,2)$.

We shall assume by induction hypothesis that our main theorem is true for $n-1$. For any $n \times n$ orthogonal matrix R, we have

$$\begin{aligned}&\Delta(x)\Delta(x)^{\mathrm{T}} = H(\eta)(x) \\ &\Rightarrow R^{\mathrm{T}}\Delta(Rx)RR^{\mathrm{T}}\Delta(Rx)^{\mathrm{T}}R = R^{\mathrm{T}}H(\eta)(Rx)R \\ &\Rightarrow \tilde{\Delta}(x)\tilde{\Delta}(x)^{\mathrm{T}} = H(\tilde{\eta})(x) \quad \text{by Lemma 12} \qquad (70)\end{aligned}$$

where

$$\tilde{\eta}(x) = \eta(Rx) \qquad (71)$$

$$\begin{aligned}\tilde{\Delta}(x) &= R^{\mathrm{T}}\Delta(Rx)R \\ &= R^{\mathrm{T}}[A_1(r_{11}x_1 + r_{12}x_2 + \cdots + r_{1n}x_n) + \cdots \\ &\quad + A_n(r_{n1}x_1 + r_{n2}x_2 + \cdots + r_{nn}x_n)]R \\ &= \tilde{A}_1 x_1 + \tilde{A}_2 x_2 + \cdots + \tilde{A}_n x_n \qquad (72)\end{aligned}$$

where

$$\tilde{A}_l = R^{\mathrm{T}}A_1Rr_{1l} + R^{\mathrm{T}}A_2Rr_{2l} + \cdots + R^{\mathrm{T}}A_nRr_{nl}, \quad 1 \le l \le n \qquad (73)$$

$$\tilde{A}_l^{\mathrm{T}} = -\tilde{A}_l \qquad (74)$$

If $(A_1(1,2), A_1(1,3), \ldots, A_1(1,n)) \neq 0$, then we shall take

$$R = \left(\begin{array}{c|cccc} 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & & & & \\ 0 & & \tilde{R} & & \\ \vdots & & & & \\ 0 & & & & \end{array}\right)$$

where $\tilde{R}$ is a $(n-1) \times (n-1)$ orthogonal matrix such that

$$(A_1(1,2), A_1(1,3), \ldots, A_1(1,n)) \cdot \tilde{R} = (a, 0, \ldots, 0), \quad a \neq 0$$

Then

$$\tilde{A}_1 = R^{\mathrm{T}}A_1R = R^{\mathrm{T}}\left(\begin{array}{c|c} 0 & A_1(1,2)\cdots A_1(1,n) \\ \hline A_1(2,1) & \\ \vdots & B_1 \\ A_1(n,1) & \end{array}\right)$$

$$R = \left(\begin{array}{c|cccc} 0 & a & 0 & \cdots & 0 \\ \hline -a & & & & \\ 0 & & \tilde{R}^{\mathrm{T}}B_1R & & \\ \vdots & & & & \\ 0 & & & & \end{array}\right)$$

i.e. $(\tilde{A}_1(1,2), \tilde{A}_1(1,3), \ldots, \tilde{A}_1(1,n)) = (a, 0, \ldots, 0)$. By applying Lemma 11 to (70), we have

$$\begin{aligned}\sum_{l=1}^n [\tilde{A}_1(2,l)]^2 = \sum_{l=1}^n [\tilde{A}_2(1,l)]^2 &= \frac{1}{2}\sum_{l=1}^n [\tilde{A}_1(1,l)\tilde{A}_2(2,l) \\ &\quad + \tilde{A}_1(2,l)\tilde{A}_2(1,l)] \\ &= \frac{1}{2}\sum_{l=1}^n \tilde{A}_1(2,l)\tilde{A}_2(1,l) \\ &\le \frac{1}{4}\sum_{l=1}^n [\tilde{A}_1(2,l)]^2 + \frac{1}{4}\sum_{l=1}^n [\tilde{A}_2(1,l)]^2\end{aligned}$$

$$\Rightarrow \frac{3}{4}\sum_{l=1}^n [\tilde{A}_1(2,l)]^2 \le \frac{1}{4}\sum_{l=1}^n [\tilde{A}_2(1,l)]^2,$$

$$\frac{3}{4}\sum_{l=1}^n [\tilde{A}_2 1,l)^2 \le \frac{1}{4}\sum_{l=1}^n [\tilde{A}_1(2,l)]^2$$

$$\Rightarrow \sum_{l=1}^n [\tilde{A}_1(2,l)]^2 = 0$$

$$\Rightarrow \tilde{A}_1(1,2) = -\tilde{A}_1(2,1) = 0$$

This contradicts the fact that $\tilde{A}_1(1,2) = a \neq 0$. Therefore we conclude that $A_1(1,l) = 0,\ 1 \le l \le n$. Now we apply Lemma 11 with $i = 1,\ 2 \le j \le n$. Then we get

$$\sum_{l=1}^{n}[A_1(j,l)]^2 = \sum_{l=1}^{n}[A_j(1,l)]^2 = \frac{1}{2}\sum_{l=1}^{n}A_1(j,l)A_j(1,l)$$

$$\le \frac{1}{4}\sum_{l=1}^{n}[A_1(j,l)^2 + \frac{1}{4}\sum_{l=1}^{n}[A_j(1,l)]^2$$

$$\Rightarrow \frac{3}{4}\sum_{l=1}^{n}[A_1(j,l)]^2 \le \frac{1}{4}\sum_{l=1}^{n}[A_1(j,l)]^2,$$

$$\frac{3}{4}\sum_{l=1}^{n}[A_j(1,l)]^2 \le \frac{1}{4}\sum_{l=1}^{n}[A_1(j,l)]^2$$

$$\Rightarrow \sum_{l=1}^{n}[A_1(j,l)]^2 = 0 = \sum_{l=1}^{n}[A_j(1,l)]^2$$

$$\Rightarrow A_1 = 0 \quad \text{and}$$

$$A_l = \left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B_l & \\ 0 & & & \end{array}\right) \qquad 2 \le l \le n$$

where B_l is a $(n-1) \times (n-1)$ antisymmetric matrix.

Let $\bar{x} = (x_2, \ldots, x_n)$ and $\bar{\Delta}(\bar{x}) = B_2 x_2 + \cdots + B_l x_l$. Then

$$\Delta(x) = \left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \bar{\Delta}(\bar{x}) & \\ 0 & & & \end{array}\right)$$

Since

$$H(\eta_4) = \Delta(x)\Delta(x)^{\mathrm{T}} = \left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & \bar{\Delta}(\bar{x})\bar{\Delta}(\bar{x})^{\mathrm{T}} & \\ 0 & & & \end{array}\right)$$

we have

$$\frac{\partial^2 \eta_4}{\partial x_1 \partial x_l} = 0 \qquad 1 \le l \le n$$

Thus η_4 is independent of x_1 variable. Denote $\bar{\eta}_4 = \eta_4(x_2, \ldots, x_n)$. Then we have

$$H(\bar{\eta}_4) = \left(\frac{\partial^2 \eta}{\partial x_i \partial x_j}\right)_{2 \le i,\, j \le n} = \bar{\Delta}(\bar{x})\bar{\Delta}(\bar{x})^{\mathrm{T}}$$

By induction hypothesis, we have $\bar{\Delta}(\bar{x}) = 0$. Therefore $\Delta(x) = 0$. □

7. Proof of the classification theorem

In this last section, we shall only outline the proof that ω_{ij} is a constant for $1 \le i \le k,\ k+1 \le j \le n$ or $k+1 \le i \le n,\ 1 \le j \le k$. The details of the proof of the Lemmas and Propositions below can be found in Yau and Hu (preprint). Let U_i be the space of differential operators with order at most i. The following Propositions and Lemmas will facilitate the proof of our classification theorem.

Lemma 16: *Let $D_i = (\partial/\partial x_i) - f_i$ and g, h be functions defined on $\mathbb{R}^n$. Then*

$$\begin{aligned}[gD_1^{i_1} \ldots D_s^{i_s}, hD_1^{j_1} \ldots D_t^{j_t}] &= i_1 g\frac{\partial h}{\partial x_1} D_1^{i_1-1} D_2^{i_2} \ldots D_s^{i_s} D_1^{j_1} \ldots D_t^{j_t} \\ &+ i_s g\frac{\partial h}{\partial x_s} D_1^{i_1} \ldots D_{s-1}^{i_{s-1}} D_{s-1}^{i_s-1} D_1^{j_1} \ldots D_t^{j_t} \\ &- j_1 h\frac{\partial g}{\partial x_1} D_1^{i_1} \ldots D_s^{i_s} D_1^{j_1-1} D_2^{j_2} \ldots D_t^{j_t} \\ &- \cdots - j_t h\frac{\partial g}{\partial x_t} D_1^{i_1} \ldots D_s^{i_s} D_1^{j_1} \ldots D_{t-1}^{j_{t-1}} D_t^{j_t-1} \\ &\pmod{U_{i_1+\cdots+i_s+j_1+\cdots+j_t-2}}\end{aligned}$$

Lemma 17: *Let E be a finite-dimensional estimation algebra with maximal rank. Let k be the quadratic rank of E. Then $\partial\omega_{il}/\partial x_j = \partial\omega_{jl}/\partial x_i$ for all $k+1 \le l \le n$ and $1 \le i, j \le k$.*

Proposition 1: *If $x_{k_{p-1}+1}^2 + \cdots + x_{k_p}^2$ is a basic quadratic form in E (cf. (41)) and $\partial\omega_{jl}/\partial x_i = 0$ for all $k+1 \le l \le n,\ k_{p-1}+1 \le i,\ j \le k_p$ and $i \neq j$, then $\partial\omega_{il}/\partial x_i = 0$ for all $k_{p-1}+1 \le i \le k_p$.*

Lemma 18: *Let $x_{k_{r-1}+1}^2 + \cdots + x_{k_r}^2$ and $x_{k_{s-1}+1}^2 + \cdots + x_{k_s}^2$ be the basic forms in E (cf. (41)), where $k_{r-1} < k_r \le k_{s-1} < k_s$. Let $\xi_{ij} = \sum_{l=k+1}^{n}(\partial\omega_{jl}/\partial x_i)D_l$. Suppose $\sum_{j=k_{s-1}+1}^{k_s} \xi_{pj}\xi_{qj} = 0$ for all $k_{r-1}+1 \le p,\ q \le k_r,\ p \neq q$. Then $\partial\omega_{jl}/\partial x_i = 0$ for all $k+1 \le l \le n,\ k_{r-1}+1 \le i \le k_r$ and $k_{s-1}+1 \le j \le k_s$.*

Lemma 19: *Let $x_{k_{r-1}+1}^2 + \cdots + x_{k_r}^2$ and $x_{k_{s-1}+1}^2 + \cdots + x_{k_s}^2$ be the basic quadratic forms in E (cf. (41)), where $k_{r-1} < k_r \le k_{s-1} < k_s$. Let $\xi_{ij} = \sum_{l=k+1}^{n}(\partial\omega_{jl}/\partial x_i)D_l$. Then $\sum_{j=k_{s-1}+1}^{k_s} \xi_{pj}\xi_{qj} = 0$ for all $k_{r-1}+1 \le p,\ q \le k_r$, $p \neq q$ if and only if $\sum_{j=k_{s-1}+1}^{k_s} a_{jl_1}^p a_{jl_2}^q = 0$ for all $k+1 \le l_1,\ l_2 \le n,\ k_{r-1}+1 \le p,\ q \le k_r,\ p \neq q$, where $a_{jl_1}^p = \partial\omega_{jl_1}/\partial x_p$.*

Lemma 20: *Let $x^2_{k_{r-1}+1} + \cdots + x^2_{k_r}$ and $x^2_{k_{s-1}+1} + \cdots + x^2_{k_s}$ be the basic quadratic forms in E (cf. (41)), where $k_{r-1} < k_r \le k_{s-1} < k_s$. Assume that $Q_l = \sum_{i=k_{r-1}+1}^{k_r} \sum_{j=k_{s-1}+1}^{k_s} a^i_{jl} x_i x_j \in E$ for all $k+1 \le l \le n$, where $a^i_{jl} = \partial\omega_{jl}/\partial x_i$. Then $\sum_{j=k_{s-1}+1}^{k_s} a^p_{jl_1} a^q_{jl_2} = 0$ for all $k+1 \le l_1, \; l_2 \le n$, $k_{r-1}+1 \le p, \; q \le k_r$.*

Proposition 2: *Let $x^2_{k_{r-1}+1} + \cdots + x^2_{k_r}$ and $x^2_{k_{s-1}+1} + \cdots + x^2_{k_s}$ be the basic quadratic forms in E (cf, (41)), where $k_{r-1} < k_r \le k_{s-1} < k_s$. Then $\partial\omega_{jl}/\partial x_i = 0$ for all $k+1 \le l \le n$, $k_{r-1}+1 \le i \le k_r$ and $k_{s-1}+1 \le j \le k_s$.*

Proposition 3: *Let $x^2_{k_{r-1}+1} + \cdots + x^2_{k_r}$ be a basic quadratic form in E (cf. (41)). Then $\partial\omega_{jl}/\partial x_i = 0$ for all $k+1 \le l \le n$, $k_{r-1}+1 \le i, j \le k_r$ and $i \ne j$.*

Theorem 25: *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then $\Omega = (\omega_{ij})$ is a matrix with constant coefficients.*

Proof: Theorem 24, we only need to prove ω_{ij} are constant functions $1 \le i \le k$, $k+1 \le j \le n$. This follows from Propositions 1–3. □

The following is the classification theorem of finite-dimensional estimation algebra of maximal rank.

Theorem 26: *Suppose that the state space of the filtering system* (1) *is of dimension n. If E is the finite-dimensional estimation algebra with maximal rank, then $f = \nabla\phi + (\alpha_1, \ldots, \alpha_n)$ where ϕ is a smooth function and α_i, $1 \le i \le n$, are affine functions and E is a real vector space of dimension $2n+2$ with basis given by* 1, *$x_1, \ldots, x_n$, $D_1, \ldots, D_n$ and L_0.*

Proof: This follows from Theorems 13 and 25. □

8. Conclusion

In this paper we explain why the theory of estimation algebras plays an important role in non-linear filtering. We show how to use the Wei–Norman approach to construct finite dimensional filters from finite dimensional estimation algebras. We survey some results in estimation algebras after 1984. We give a self-contained proof of complete classification of finite-dimensional estimation algebras of maximal rank in one place. The proof given here is simpler than those proofs scattering in several papers. This provides the readers with a complete coherent view of the important topic on classification of finite-dimensional estimation algebras.

Acknowledgements

Research partially supported by USA Army Research Office and National Science Foundation.

Finally, I would like to thank all the referees for their careful reading of this paper and many constructive comments which were very valuable in revising the paper.

References

Benés, V. E., 1981, Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics*, **5**, 65–92.

Benés. V. E. 1985, New exact nonlinear filters with large Lie algebras. *Systems Control Letters*, **5**, 217–221.

Blankenship, G. L., Liu, C. H., and Marcus, S., 1983, Asymptotic expansions and Lie algebras for some nonlinear filtering problem. *IEEE Transactions of Automatic Control*, **28**, 878–797.

Brockett, R. W., 1981, Nonlinear systems and nonlinear estimation theory. In M. Hazewinkel and J. C. Willems (Eds), *The Mathematics of Filtering and Identification and Applications* (Dordrecht: Reidel).

Brockett, R. W., 1979, Classification and equivalence in estimation theory. *Proceedings of the 18th IEEE Conference on Decision and Control*, Ft. Lauderdale, FL, USA, pp. 172–175.

Brockett, R. W., and Clark, J. M. C., 1980, The geometry of the conditional density functions. In O. L. R. Jacobs *et al.* (Eds) *Analysis and Optimization of Stochastic Systems* (New York: Academic Press), pp. 299–309.

Bucy, R. S., 1965, Nonlinear filter. *IEEE Transactions on Automatic Control*, **10**, 198.

Chaleyat-Maurel, M., and Michel, D., 1984, Des résultats de non existence de filtre de dimension finie. *Stochastics*, **13**, 83–102.

Chen, J., 1994, On uniquity of Yau filters. *Proceedings of the American Control Conference*, Baltimore, MD, USA, pp. 252–254.

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

Chen, J., and Yau, S. S.-T., 1997, Finite-dimensional filters with nonlinear drift VII: Mitter conjecture and structure of η. *SIAM Journal of Control and Optimization*, **35**, 1116–1131.

Chen, J., Yau, S. S.-T., and Leung, C. W., 1996, Finite-dimensional filters with non-linear drift IV: Classification of finite-dimensional estimation algebras of maximal rank with state space dimension 3. *SIAM Journal of Control and Optimization*, **34**, 179–198.

Chen, J., Yau, S. S.-T., and Leung, C. W., 1997, 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**, 1132–1141.

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

Chiou, W. L., and Chiueh, W. R., (preprint) Classification of 5-dimensional estimation algebras on filtering systems.

Cohen de Lara, M., 1997, Characterization of a subclass of finite-dimensional estimation algebras with maximal rank. Application to filtering. *Mathematical Control Signals Systems*, **10**, 237–246.

Davis, M. H. A., 1980, On a multiplicative functional transformation arising in nonlinear filtering theory. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **54**, 125–139.

Davis, M. H. A., and Marcus, S. I., 1981, An introduction to nonlinear filter. In M. Hazewinkel and J. S. Willems (Eds) *The Mathematics of Filtering and Identification and Applications* (Dordrecht: Reidel).

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

DUNCAN, T. E., 1967, Probability densities for diffusion processes with applications to nonlinear filtering theory. PhD thesis, Standford.

FROST, P. A., and KAILATH, R., 1971, An innovations approach to least-squares estimation III. *IEEE Transactions Automatic Control*, **16**, 217–226.

FUJISAKI, M., KALLIANPUR, G., and KUNITA, H., 1972, Stochastic differential equations for the nonlinear filtering problem. *Osaka Journal of Mathematics*, **2**, 19–40.

HAUSSMAN, U. G., and PARDOUX, E., 1988, A conditionally almost linear filtering problem with non-Gaussian initial condition. *Stochastics*, **23**, 241–275.

HAZEWINKEL, M., 1988 a, Lecture on linear and nonlinear filtering. In W. Shiehlen and W. Wedig (Eds) *Analysis and Estimation of Stochastic Mechanical Systems, CISM Courses and Lectures 303* (Vienna: Springer).

HAZEWINKEL, M., 1988 b, Nongaussian linear filtering, identification of linear systems, and the symplectic group. Centrum Voor Wiskunde en Informatica, Report PM-R8810, October.

HAZEWINKEL, M., MARCUS, S. I., and SUSSMAN, H. J., 1983, Nonexistence of finite dimensional filters for conditional statistics of the cubic sensor problem. *Systems Control Letters*, **3**, 331–340.

HIJAB, B., 1980, Minimum energy estimation, PhD thesis, University of California, Berkeley, CA, USA.

HU, G. Q., YAU, S. S.-T., and CHIOU, W.-L., 2000, 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**, 915–932.

KAILATH, T., 1968, An innovations approach to least-squares estimation, Part I: Linear filtering in additive white noise. *IEEE Transactions on Automatic Control*, **13**, 646–655.

KALMAN, R. E., 1960, A new approach to linear filtering and predication problem. *Transactions on ASME, Ser. D. Journal of Basic Engineering*, **82**, 35–45.

KALMAN, R. E., and BUCY, R. S., 1961, New results in linear filtering and prediction theory. *Transactions on ASME Series D. Journal of Basic Engineering*, **83**, 95–108.

KRENER, A. J., 1973, On the equivalence of control systems and the linearization of nonlinear systems. *SIAM Journal of Control and Optimization*, **11**, 670–676.

KUSHNER, H. J., 1964, On the differential equations satisfied by conditional probability densities of Markov processes. *SIAM Journal of Control*, **2**, 106–119.

MAKOWSKI, A., 1986, Filtering formula for partially observed linear system with non-Gaussian initial conditions. *Stochastic*, **16**, 1–24.

MARCUS, S., 1984, Algebraic and geometric methods in nonlinear filtering. *SIAM Journal of Control and Optimization*, **22**, 817–844.

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

MITTER, S. K., 1978, Filtering theory and quantum fields. In C. Lobry (Ed), *Analyse des Systèmes*, Bordeaux, 1978; also, *Astérisque*, **76** (1980).

MORTENSEN, R. E., 1966, Optimal control of continuous time stochastic systems. PhD thesis, University of California, Berkeley, CA, USA.

OCONE, D., 1980, Topics in nonlinear filtering theory. PhD thesis, Massachusetts Institute of Technology, MA, USA.

RASOULIAN, A., and YAU, S. S.-T., 1997, Finite dimensional filters with nonlinear drift IX: Construction of finite dimensional estimation algebras of non-maximal rank. *Systems and Control Letters*, **30**, 109–118.

SHIRYAEV, A. N., 1967, Some new results in the theory of controlled stochastic processes [Russian]. *Transactions of the 4th Prague Conference on Information Theory*, Czech. Academy of Sciences, Prague.

STRATONOVICH, R. L., 1968, *Conditional Markov Processes and Their Application to the Theory of Optimal Control* (New York: Elsevier).

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

WEI, J., and NORMAN, E., 1964, On the global representation of the solutions of linear differential equations as a product of exponentials. *Proceedings of American Mathematical Science*, **15**, 327–334.

WONG, E., 1971, *Stochastic Processes in Information and Dynamical Systems* (New York: McGraw Hill).

WONG, W. S., 1987 a, Theorems on the structure of finite dimensional estimation algebras. *Systems and Control Letters*, **9**, 117–124.

WONG, W. S., 1987 b, On a new class of finite-dimensional estimation algebras. *Systems Control Letters*, **9**, 79–83.

WONG, W. S., and YAU, S. S.-T., 1998, The estimation algebra of nonlinear filtering system. In J. Baillieul and J. C. Willems (Eds) *Mathematical Control Theory, Special Volume Dedicated to 60th Birthday of Brockett* (Springer Verlag), 33–65.

WONHAM, W. M., 1965, Some applications of stochastic differential equations to optimal nonlinear filtering. *SIAM Journal of Control*, **2**, 347–369.

WU, X., YAU, S. S.-T., and HU, G. Q., 2002, Finite dimensional filters with nonlinear drift XII: linear and constant structure of Ω, stochastic theory and control. In B. Pasik-Duncan (Ed), *Proceedings of the Workshop held in Lawrence, Kansas, Lecture Notes in Control and Information Sciences #280* (Springer Verlag), pp. 507–518.

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

YAU, S. S.-T., 1994, Finite dimensional filters with nonlinear drift I: A class of filters including both Kalman–Bucy filters and Benes filters. *Journal of Mathematical Systems, Estimation and Control*, **4**, 181–203.

YAU, S. S.-T., and CHIOU, W. L., 1991, Recent results on classification of finite dimensional estimation algebras: Dimension of state space 2. *Proceedings of the 30th Conference on Decision and Control*, Brighton, England, December 11–13, 2758–2760.

YAU, S. S.-T., and HU, G.-Q., (preprint) Finite dimensional filters with nonlinear drift XIV: Classification of finite dimensional estimation algebras of maximal rank with arbitrary state space dimension and Mitter conjecture.

YAU, S. S.-T., and RASOULION, A., 1999, Classification of four-dimensional estimation algebras. *IEEE Transactions on Automatic Control*, **44**, 2312–2318.

ZAKAI, M., 1969, On the optimal filtering of diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, **11**, 230–243.