Math. Control Signals Systems (1996) 9: 370–385

Mathematics of Control, Signals, and Systems

Finite-Dimensional Filters with Nonlinear Drift, VI: Linear Structure of Ω*

Jie Chen† and Stephen S.-T. Yau†

Abstract. Ever since the concept of estimation algebra was first introduced by Brockett and Mitter independently, it has been playing a crucial role in the investigation of finite-dimensional nonlinear filters. Researchers have classified all finite-dimensional estimation algebras of maximal rank with state space less than or equal to three. In this paper we study the structure of quadratic forms in a finite-dimensional estimation algebra. In particular, we prove that if the estimation algebra is finite dimensional and of maximal rank, then the $\Omega = (\partial f_j/\partial x_i - \partial f_i/\partial x_j)$ matrix, where f denotes the drift term, is a linear matrix in the sense that all the entries in Ω are degree one polynomials. This theorem plays a fundamental role in the classification of finite-dimensional estimation algebra of maximal rank.

Key words. Finite-dimensional filters, Nonlinear drift, Estimation algebra of maximal rank.

1. Introduction

There are several approaches to nonlinear filtering. The basic approach was via the "innovations method," originally proposed by Kailath in 1967 and subsequently rigorously developed by Fujisaki *et al.* [FKK] in 1972. As pointed out by Mitter [M3], the difficulty with this approach is that the innovations process is not, in general, explicitly computable. In the late 1970s and early 1980s, Brockett and Clark [BC], Brockett [B3], and Mitter [M3] proposed the idea of using estimation algebras to construct finite-dimensional nonlinear filters. This Lie algebra approach has several merits. First, it takes into account 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.* [HMS]. 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 con-

* Date received: October 4, 1995. Date revised: August 22, 1996. This research was supported by Army Research Office Grants DAAH 04-93-0006 and DAAH 04-1-0530.

† Control and Information Laboratory, MSCS, M/C 249, University of Illinois at Chicago 851 South Morgan Street, Chicago, Illinois 60607-7045, U.S.A. u32790@UICVM.UIC.EDU or yau@uic.edu.

structed is universal in the sense of [CM] (see Section 2 for the definition). 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. Even in the case of linear filtering with non-Gaussian initial condition, the number of sufficient statistics needed in Makowski's method [M1] or Haussman and Pardoux's method [H-P] is a polynomial of degree two. So the Lie algebraic method is of practical importance.

Although the concept of estimation algebra has proven to be an invaluable tool in the study of nonlinear filtering problems [M2], until recently very little was known about estimation algebras. Beginning in the late 1980s, however, the structure and classification of finite-dimensional exact estimation algebras were studied in detail [TWY], [DTWY]. In [W] the concept of Ω was introduced, which is defined as the matrix whose (i, j) element is $\partial f_j/\partial x_i - \partial f_i/\partial x_j$, where f is the drift term of the state evolution equation. For the class of exact filtering systems, Ω is identically zero. More recently, Yau [Y] has studied filtering systems such that all entries of Ω are constants. He was able to classify all finite-dimensional estimation algebras of maximal rank in such filtering systems. The general concept of an estimation algebra of maximal rank was first introduced by Chiou and Yau [CY]. Most of the known finite-dimensional estimation algebras are of maximal rank. For example, if (2.0) below is linear, i.e., $f(x) = Ax$, $g(x) = B$, and $h(x) = Cx$, and if (A, B, C) is minimal, then the corresponding estimation algebra is maximal (see p. 114 of [H]). Chiou and Yau [CY] and Chen *et al.* [CLY] have shown respectively that if the dimension of the state space is two or three, all entries of Ω are constants as long as the estimate algebra is of maximal rank and finite dimensional. Thus, finite-dimensional estimation algebra of maximal rank is completely classified if the dimension of state space is at most three. The novelty of their theorems is that there are no *a priori* assumptions on the drift term of the nonlinear filtering system.

Our approach for the complete classification of finite-dimensional estimation algebras of maximal rank consists of two steps. The first step is to prove that, for such an estimation algebra, all the entries in the Ω-matrix are degree one polynomials. The second step is to prove that in fact all the entries in Ω are constants. Then we can apply the result of Yau [Y] to give a complete classification of finite-dimensional estimation algebras of maximal rank. The purpose of this paper is to complete the first step. The following is our main theorem.

Main Theorem. *If E is a finite-dimensional estimation algebra of maximal rank, then all the entries $\omega_{ij} = \partial f_j/\partial x_i - \partial f_i/\partial x_j$ of Ω are degree one polynomials. Let k be the maximal rank of quadratic forms in E. Then there exists an orthogonal change of coordinates such that ω_{ij} are constants for $1 \le i, j \le k$; ω_{ij} are degree one polynomials in $x_1, \ldots, x_k$ for $1 \le i \le k$ or $1 \le j \le k$; and ω_{ij} are degree one polynomials in $x_{k+1}, \ldots, x_n$ for $k + 1 \le i, j \le n$.*

Let n be the dimension of the state space. In case $n = 3$, there are three unknowns: ω_{12}, ω_{13}, and ω_{23}. It is easy to see that they are all degree two polynomials in view of Ocone's theorem. In [YL] Leung and Yau showed that the coefficients of the quadratic parts of ω_{12}, ω_{13}, and ω_{23} have to satisfy 90 quadratic equations. It was also shown in that paper that this system of 90 quadratic equations has only a trivial solution. Later Chen *et al.* proved that Ω is a matrix of constants [CLY]. The novelty of our main theorem is that it holds for arbitrary n. Thus, it is the fundamental step in the classification of finite-dimensional estimation algebras of maximal rank.

In Section 3 we classify quadratic forms in finite-dimensional estimation algebras (without maximal rank assumption). In Section 4 we use the results of Section 3 to study the Ω matrix.

2. Basic Concepts

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

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

Here x, v, y, and w are respectively $\mathbf{R}^n$, $\mathbf{R}^p$, $\mathbf{R}^m$, and $\mathbf{R}^m$ valued processes, and v and w have components that 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 [DM], for example) that $\rho(t, x)$ is given by normalizing $\sigma(t, x)$, which satisfies the following Duncan–Mortensen–Zakai (DMZ) equation:

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

where

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

and, for $i = 1, \ldots, m$, 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 (2.1) is a stochastic partial differential equation. In real applications, we are interested in constructing state estimators from observed sample paths with some property of robustness. Based on Rozovsky's transformation [R],

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

Davis [D] considered the following robust DMZ equation:

$$\frac{\partial\xi}{\partial t}(t, x) = L_0\xi(t, x) + \sum_{i=1}^{m} y_i(t)[L_0, L_i]\xi(t, x) + \frac{1}{2}\sum_{i,j=1}^{m} y_i(t)y_j(t)[[L_0, L_i], L_j]\xi(t, x),$$

$$\xi(0, x) = \sigma_0, \tag{2.2}$$

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

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

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

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

$$\xi(t, x) = v(z(t), t, x). \tag{2.4}$$

Following [CM], we say that system (2.0) has a robust universal finite-dimensional filter if, for each initial probability density σ_0, there exists a z_0 such that (2.3) and (2.4) hold if $z(0) = z_0$ and μ_i, w_i are independent of σ_0.

The method of Wei and Norman [WN] 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 the A_i's are $n \times n$ matrices and the a_i's are scalar-valued functions. Let $B_1, \ldots, B_l$ be a basis of the Lie algebra generated by $A_1, \ldots, A_m$. Then the Wei–Norman theorem states that, locally in t, $X(t)$ has a representation of the form

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

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

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

for all i. The function c_i's in the above equation are determined by the structure constants of the Lie algebra (generated by the A_i's) relative to the basis $\{B_1, \ldots, B_l\}$.

The extension of Wei and Norman's approach to the nonlinear 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 partial differential equation. For this purpose, we need to introduce the estimation algebra of (2.0) and examine its algebraic structure.

Definition. The estimation algebra E of a filtering problem (2.0) is defined to be the Lie algebra generated by $\{L_0, L_1, \ldots, L_m\}$. 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.

In [W] the concept of Ω is introduced, which is defined as the matrix whose

(i, j)-element ω_{ij} is $\partial f_j/\partial x_i - \partial f_i/\partial x_j$. Define

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

and

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

Then

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

The following theorem, proved in [Y], plays a fundamental role in the classification of finite-dimensional estimation algebras.

Theorem 2.1. *Let E be a finite-dimensional estimation algebra of* (2.0) *such that $\omega_{ij} = \partial f_j/\partial x_i - \partial f_i/\partial x_j$ are constant functions. If E is of maximal rank, then E is a real vector space of dimension $2n + 2$ with basis given by $1, x_1, x_2, \ldots, x_n, D_1, D_2, \ldots, D_n$, and L_0.*

We need the following basic result [O] for later discussion.

Theorem 2.2 (Ocone). *Let E be a finite-dimensional estimation algebra. If a function ξ is in E, then ξ is a polynomial of degree ≤ 2.*

For the convenience of the readers, we also list the following elementary lemmas without proof. The lemmas were proven in [Y] and [CY].

Lemma 2.3.

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

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

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

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

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

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

(vii) $[D_k^2, hD_iD_j] = 2(\partial h/\partial x_k)D_kD_iD_j + 2h\omega_{jk}D_iD_k + 2h\omega_{ik}D_kD_j + (\partial^2 h/\partial x_k^2)D_iD_j + 2h(\partial\omega_{jk}/\partial x_i)D_k + h(\partial\omega_{jk}/\partial x_k)D_i + h(\partial\omega_{ik}/\partial x_k)D_j + h(\partial^2\omega_{jk}/\partial x_i\partial x_k)$.

(viii) $[gD_iD_j, hD_k] = g(\partial h/\partial x_j)D_iD_k + g(\partial h/\partial x_i)D_jD_k + gh\omega_{kj}D_i + gh\omega_{ki}D_j + g(\partial^2 h/\partial x_i\,\partial x_j)D_k + gh(\partial\omega_{kj}/\partial x_i) - h(\partial g/\partial x_k)D_iD_j$.

Lemma 2.4.

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

The following theorem, which was proven in detail in [Y], shows in particular how to construct finite-dimensional filters from finite-dimensional estimation algebras.

Theorem 2.5 [Y]. *Let E be an estimation algebra of* (2.0) *satisfying* $\partial f_j/\partial x_i - \partial f_i/\partial x_j = c_{ij}$, *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, $x_1, \dots, x_n, D_1, \dots, D_n$, *and* L_0, *and* $\eta := \sum_{i=1}^n \partial f_i/\partial x_i + \sum_{i=1}^n f_i^2 + \sum_{i=1}^m h_i^2$ *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* (2.2) *has a solution for all* $t \ge 0$ *of the form*

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

where $T(t), r_1(t), \dots, r_n(t), s_1(t), \dots, s_n(t)$ *satisfies the following ordinary differential equations*:

$$\frac{ds_i}{dt}(t) = r_i(t) + \sum_{j=1}^n s_j(t) c_{ji} + \sum_{k=1}^n h_{ki} y_k(t), \qquad 1 \le i \le n, \tag{2.5}$$

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

$$\frac{dr_j}{dt}(t) = \frac{1}{2}\sum_{i=1}^n s_i(t)(a_{ij} + a_{ji}), \qquad 1 \le j \le n; \tag{2.6}$$

and

$$\begin{aligned}\frac{dT}{dt}(t) = &-\frac{1}{2}\sum_{i=1}^n r_i^2(t) - \frac{1}{2}\sum_{i=1}^n s_i^2(t)\left(\sum_{j=1}^n c_{ij}^2 - a_{ii}\right) + \sum_{i=1}^n r_i(t) - \sum_{j=2}^n \sum_{i=1}^j s_j(t) c_{ij} \\ &+ \sum_{1 \le i < k \le n} s_i(t) s_k(t)\left[\sum_{j=1}^n c_{ij} c_{jk} + \frac{1}{2}(a_{ik} + a_{ki})\right] + \frac{1}{2}\sum_{i=1}^n s_i(t) b_i \\ &+ \frac{1}{2}\sum_{i,j=1}^m y_i(t) y_j(t) \sum_{k=1}^n h_{ik} h_{jk} - \sum_{i,j=1}^n s_i(t) r_j(t) c_{ij}.\end{aligned} \tag{2.7}$$

It follows that a universal finite-dimensional filter exists for (2.7).

3. Structure of Quadratic Forms in Estimation Algebra

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

Definition. For any quadratic form $p \in Q$, there exists a symmetric matrix $A \in M_n(\mathbf{R})$ such that

$$p(x) = X^T A X. \tag{3.1}$$

The rank of the quadratic form p is denoted by $r(p)$ and is defined to be the rank of the matrix A.

Let $p_0 \in E \cap Q$ be an element with the greatest positive rank, namely, $r(p_0) \geq r(p)$ for any $p \in E \cap Q$. After an orthogonal transformation on x, p_0 can be written as

$$p_0 = c_1 x_1^2 + c_2 x_2^2 + \cdots + c_k x_k^2, \qquad c_i \neq 0, \quad 0 \leq k \leq n. \tag{3.2}$$

From p_0, we can construct a sequence of quadratic forms in $E \cap Q$ as follows:

$$q_0 = p_0,$$
$$q_j = [[L_0, q_{j-1}], q_0] = \sum_{i=1}^{k} 4^j c_i^{j+1} x_i^2.$$

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

$$p_0 = x_1^2 + x_2^2 + \cdots + x_k^2 \in E. \tag{3.3}$$

Lemma 3.1. *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 $k + 1 \leq j \leq n$.*

Proof. Suppose on the contrary that $\partial p/\partial x_j \neq 0$ for some $j > k$. Let A be a symmetric matrix such that $p = X^T A X$. A can be written as

$$A = \begin{pmatrix} A_1 & A_2 \\ A_2^T & A_4 \end{pmatrix}, \tag{3.4}$$

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 and an $(n-k) \times (n-k)$ orthogonal matrix S_1 such that $S^T A_1 S$ and $S_1^T A_4 S_1$ are diagonal matrices. So, we can assume that A_1 and A_4 are diagonal matrices. $\partial p/\partial x_j \neq 0$ for some $j > k$ implies $A_2 \neq 0$ or $A_4 \neq 0$. Since

$$r(\lambda p_0 + \sigma p) = \operatorname{rank}\begin{pmatrix} \lambda I + \sigma A_1 & \sigma A_2 \\ \sigma A_2^T & \sigma A_4 \end{pmatrix}, \tag{3.5}$$

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

$$r(\lambda p_0 + \sigma p) > k. \tag{3.6}$$

This contradicts the greatest positive rank assumption of p_0. ∎

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

$$p_1 = \sum_{i=1}^{k_1} x_i^2 \in E, \qquad 1 \le k_1 \le k. \tag{3.7}$$

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

$$S_1 = \{1, 2, \ldots, k_1\} \subseteq S = \{1, \ldots, k\} \tag{3.8}$$

and

$$Q_1 = \text{real vector space spanned by } \{x_i x_j : k_1 + 1 \le i \le j \le k\} \subseteq Q. \tag{3.9}$$

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

$$p_2 = \sum_{i=k_1+1}^{k_2} x_i^2 \in E \cap Q_1 \tag{3.10}$$

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

$$S_i = \{k_{i-1} + 1, \ldots, k_i\}, \qquad k_0 = 0, \qquad k_i \le k, \tag{3.11}$$

$$Q_i = \text{linear span}\{x_i x_j : k_i + 1 \le i \le j \le k\}, \tag{3.12}$$

$$p_i = \sum_{j=k_{i-1}+1}^{k_i} x_j^2 = \sum_{j \in S_i} x_j^2, \qquad i > 0, \tag{3.13}$$

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

Lemma 3.2. *If $p \in E \cap Q$, then*

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

Proof. In view of Lemma 2.3 and the fact that $[L_0, p_i] \in E$, $[L_0, p_0 - p_i] \in E$, we have

$$\sum_{j \in S_i} x_j D_j \in E, \qquad \sum_{j \in S - S_i} x_j D_j \in E. \tag{3.14}$$

Hence

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

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

$$p(0, \ldots, 0, x_{k_{i-1}+1}, \ldots, x_{k_i}, 0, \ldots, 0) = \lambda p_i. \quad \blacksquare \tag{3.15}$$

Similarly we also have the following lemma.

Lemma 3.3. *If $p \in E \cap Q$, then*

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

Proof. The lemma follows immediately from the following formula:

$$p(x_1, \ldots, x_{k_{i-1}}, 0, \ldots, 0, x_{k_i+1}, \ldots, x_k)$$
$$= p - \left[\sum_{j \in S - S_i} x_j D_j, \left[\sum_{j \in S_i} x_j D_j, p\right]\right] - p(0, \ldots, 0, x_{k_{i-1}+1}, \ldots, x_{k_i}, 0, \ldots, 0). \qquad \blacksquare$$

Lemma 3.4. *Let $p = \sum_{i \in S_{l_1}} \sum_{j \in S_{l_2}} 2a_{ij} x_i x_j \in E$, where $a_{ij} \in \mathbf{R}$ and $l_1 < l_2$. Let $X_i = (x_{k_{i-1}+1}, \ldots, x_{k_i})^T$ be a $(k_i - k_{i-1})$-vector. Under this notation, p can be written as*

$$p = (X_{l_1}^T, X_{l_2}^T)\begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}\begin{pmatrix} X_{l_1} \\ X_{l_2} \end{pmatrix}. \tag{3.16}$$

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

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

Since $[[L_0, p], p] \in E$, from Lemma 3.2, we have

$$\sum_{i,m \in S_{l_1}} \left(\sum_{j \in S_{l_2}} a_{ij} a_{mj}\right) x_i x_m = \lambda_1 p_{l_1}, \tag{3.17}$$

$$\sum_{j,l \in S_{l_2}} \left(\sum_{i \in S_{l_1}} a_{ij} a_{il}\right) x_j x_l = \lambda_2 p_{l_2}. \tag{3.18}$$

Equations (3.17) and (3.18) 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 length, it follows that A is a constant multiple of an orthogonal matrix. $\blacksquare$

4. Structure of $\Omega = (\omega_{ij})$

In this section we investigate the possible structure of the matrix Ω. Throughout this section we assume that the estimation algebra E is of maximal rank. From Section 2 we know that x_i and D_i are in E for $1 \leq i \leq n$.

Lemma 4.1. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then $\omega_{ij} \in E$ is a polynomial with degree at most two.*

Proof. Since D_i, $D_j \in E$, we have $\omega_{ij} = [D_i, D_j] \in E$. Then lemma follows from Theorem 2.2. ∎

Let α_{ij}, β_{ij}, and γ_{ij} be respectively the quadratic part, linear part, and constant part of ω_{ij}, that is, $\omega_{ij} = \alpha_{ij} + \beta_{ij} + \gamma_{ij}$.

Lemma 4.2. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then $\alpha_{ij} \in E$ and*

(i) $\partial\alpha_{jl}/\partial x_i + \partial\alpha_{li}/\partial x_j + \partial\alpha_{ij}/\partial x_l = 0$,
(ii) $\partial\beta_{jl}/\partial x_i + \partial\beta_{li}/\partial x_j + \partial\beta_{ij}/\partial x_l = 0$,
(iii) *α_{ij} depends only on $x_1, \ldots, x_k$ for $i \le k$ or $j \le k$, and*
(iv) *$\alpha_{ij} = 0$ for $k + 1 \le i, j \le n$.*

Proof. Since E is finite dimensional of maximal rank and $\omega_{ij} \in E$, it follows that $\alpha_{ij} \in E$. Hence $\partial\alpha_{ij}/\partial x_l = 0$ for $l > k$ by Lemma 3.1. Recall that $\omega_{ij} = \partial f_j/\partial x_i - \partial f_i/\partial x_j$. So, for every i, j, l,

$$\frac{\partial \omega_{ij}}{\partial x_l} + \frac{\partial \omega_{li}}{\partial x_j} + \frac{\partial \omega_{jl}}{\partial x_i} = 0. \tag{4.1}$$

Parts (i) and (ii) of Lemma 4.2 follow from (4.1). Let $k + 1 \le i, j \le n$, and $l \le k$. Then (i) gives $\partial\alpha_{ij}/\partial x_l = 0$. Note that α_{ij} is a quadratic form in $x_1, \ldots, x_k$. Hence $\alpha_{ij} = 0$ for $k + 1 \le i, j \le n$. ∎

Theorem 4.3. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then ω_{ij} is a polynomial of degree at most one for all $1 \le i, j \le n$.*

Proof. Since $p_l \in E$ and $D_j \in E$, we have

$$\sum_{i \in S_l} x_i \omega_{ji} = -[[L_0, p_l], D_j] - \sum_{i \in S_l} \delta_{ij} D_i \in E.$$

By Ocone's theorem (Theorem 2.2), we deduce that

$$\sum_{i \in S_l} x_i \alpha_{ji} = 0. \tag{4.2}$$

Without loss of generality, we assume that $l = 1$. Let $X^T = (X_1^T, \overline{X}_1^T)$. $\overline{X}_1$ is the complementing variable of X_1 in X. Write

$$\alpha_{ji}(X) = \alpha_{ji}(X_1, 0) + \alpha_{ji}(0, \overline{X}_1) + (\alpha_{ji} - \alpha_{ji}(X_1, 0) - \alpha_{ji}(0, \overline{X}_1)). \tag{4.3}$$

Hence (4.2) is still true if we replace α_{ji} by one of the three terms in the right-hand side of (4.2). We see immediately that

$$\alpha_{ji}(0, \overline{X}_1) = 0. \tag{4.4}$$

Also, by Lemmas 4.2 and 3.2, we have

$$\alpha_{ji}(X_1, 0) = \lambda_i p_1, \tag{4.5}$$

so (4.2) gives

$$\sum_{i\in S_1} x_i\lambda_i p_1 = 0. \tag{4.6}$$

It follows that $\lambda_i = 0$, that is, $\alpha_{ji}(X_1, 0)$, $\forall i \in S_1$. Finally, $\alpha_{ji} - \alpha_{ji}(X_1, 0) - \alpha_{ji}(0, \overline{X}_1)$ is a linear combination of $2X_1^T A_{ji} X_l$ for $l \geq 2$, and A_{ji} is a constant multiple of an orthogonal matrix. Therefore (4.2) gives

$$\sum_{l\geq 2} \beta_l X_1^T\left(\sum_{i\in S_1} 2x_i A_{ji}\right) X_l = 0, \qquad \beta_l = \text{constant}. \tag{4.7}$$

This implies that whenever $\beta_l \neq 0$, we have

$$X_1^T\left(\sum_{i\in S_1} 2x_i A_{ji}\right) = 0. \tag{4.8}$$

Fix $i_0 \in S_1$, and let $x_{i_0} = 1$ and $x_i = 0$ for $i \neq i_0$. Then (4.8) becomes

$$(0, \ldots, 0, 1, 0, \ldots, 0)A_{ji_0} = 0. \tag{4.9}$$

Since A_{ji_0} is a constant multiple of an orthogonal matrix, we see that $A_{ji_0} = 0$. Thus

$$\alpha_{ji} - \alpha_{ji}(X_1, 0) = \alpha_{ji}(0, \overline{X}_1) = 0. \tag{4.10}$$

So we have proved that $\alpha_{ji} = 0$. ∎

We need some more results before we can finish the proof of the main theorem stated in Section 1.

Proposition 4.4. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the same notation as before, if $l_1 \neq l_2$ and $i \in S_{l_1}, j \in S_{l_2}$, then ω_{ij} is a constant.*

Proof. Recall that from (3.14), we have $\sum_{i\in S_{l_1}} x_i D_i$ and $\sum_{i\in S_{l_2}} x_i D_i$ in E. Hence

$$\sum_{i\in S_{l_1}}\sum_{j\in S_{l_2}} x_i x_j \omega_{ij} = -\left[\sum_{i\in S_{l_1}} x_i D_i, \sum_{j\in S_{l_2}} x_j D_j\right] \in E. \tag{4.11}$$

In view of Theorem 2.2, (4.11) implies

$$\sum_{i\in S_{l_1}}\sum_{j\in S_{l_2}} x_i x_j \beta_{ij} = \sum_{i\in S_{l_1}} x_i\left(\sum_{j\in S_{l_2}} x_j\beta_{ij}\right) = \sum_{j\in S_{l_2}} x_j \sum_{i\in S_{l_1}} x_i\beta_{ij} = 0. \tag{4.12}$$

Hence β_{ij} depends only on x_m, where $m \in S_{l_1} \cup S_{l_2}$. Since E is of maximal rank, $D_j \in E$ for any j. In particular, $[\sum_{i\in S_{l_1}} x_i D_i, D_j] \in E$ for $j \in S_{l_2}$, and $[\sum_{j\in S_{l_2}} x_j D_j, x_i]$ $\in E$ for $i \in S_{l_1}$. In view of (iii) of Lemma 2.3, we have

$$\sum_{i\in S_{l_1}} x_i\beta_{ij} \in E \quad \text{for } j \in S_{l_2} \quad \text{and} \quad \sum_{j\in S_{l_2}} x_j\beta_{ij} \in E \quad \text{for } i \in S_{l_1}. \tag{4.13}$$

From (4.12), a similar argument as in the proof of Theorem 4.3, we have

$$\sum_{i\in S_{l_1}} x_i\beta_{ij} = 0 \quad \text{for } j \in S_{l_2} \quad \text{and} \quad \sum_{j\in S_{l_2}} x_j\beta_{ij} = 0 \quad \text{for } i \in S_{l_1}. \tag{4.14}$$

The first equation of (4.14) says that, for $i \in S_{l_1}$ and $j \in S_{l_2}$, β_{ij} does not depend on the variable x_m for $m \in S_{l_2}$. The second equation of (4.14) says that, for $i \in S_{l_1}$ and $j \in S_{l_2}$, β_{ij} does not depend on the variable x_m for $m \in S_{l_1}$. Hence $\beta_{ij} = 0$. ∎

Proposition 4.5. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the same notation as before, then*

(i) *for* $j \in S_l$,

$$\begin{bmatrix} \beta_{jk_{l-1}+1} \\ \vdots \\ \beta_{jk_l} \end{bmatrix} = A_1^{(j)} X_l \qquad \textit{with} \quad A_1^{(j)} = -A_1^{(j)T},$$

where $X_l = (x_{k_{l-1}+1}, \ldots, x_{k_l})$ *and A is a* $k_l \times k_l$ *matrix; and*

(ii) *for* $j > k$,

$$\begin{bmatrix} \beta_{jl} \\ \vdots \\ \beta_{jk_l} \end{bmatrix} = \lambda X_l + A_2^{(j)} \tilde{X}_l,$$

where $\tilde{X}_l$ *is the complementary variable vector of* X_l *in* $(x_1, \ldots, x_k)^T$; *that is,* $\tilde{X}_l = (x_1, \ldots, x_{k_{l-1}}, x_{k_l+1}, \ldots, x_k)$ *and* $A_2^{(j)}$ *is a* $k_l \times (k - k_l)$ *matrix.*

Proof. Since $[[L_0, p_l], D_j] \in E$, we have $\sum_{i \in S_l} x_i \beta_{ji} \in E$. There exist two matrices $A_1^{(j)}$, $A_2^{(j)}$ such that

$$\begin{bmatrix} \beta_{jk_{l-1}+1} \\ \vdots \\ \beta_{jk_l} \end{bmatrix} = A_1^{(j)} X_l + A_2^{(j)} \overline{X}_l, \tag{4.15}$$

where $\tilde{X}_l = (x_1, \ldots, x_{k_{l-1}}, x_{k_l+1}, \ldots, x_n)$. Hence,

$$\sum_{i \in S_l} x_i \beta_{ji} = X_l^T A_1^{(j)} X_l + X_l^T A_2^{(j)} \tilde{X}_l \in E. \tag{4.16}$$

Therefore, by Lemma 3.2, each of the above terms in (4.16) belongs to E. In fact $X_l^T A_1^{(j)} X_l \in E$ implies

$$X_l^T A_1^{(j)} X_l = \lambda p_l \tag{4.17}$$

by Lemma 3.2. In view of Lemma 3.1, $X_l^T A_2^{(j)} \tilde{X}_l \in E$ is independent of $x_{k+1}, \ldots, x_n$ variables. Hence

$$A_2^{(j)} = (B_1, B_2, \ldots, B_m, 0, \ldots, 0), \tag{4.18}$$

where $B_1, \ldots, B_m$ are constants multiple of some orthogonal matrices by Lemma 3.4.

Case (i): $j \in S_l$. In this case since $\beta_{jj} = 0$, the λ in (4.17) has to be zero. It follows that $A_1^{(j)} = -A_1^{(j)T}$ (note that $A_1^{(j)}$ is not a symmetric matrix). In view of (4.16), since $X_l^T A_1^{(j)} X_l = 0$, we conclude that β_{ji} is independent of the x_i variable for all $i \in S_l$. Recall that $\beta_{ji} = -\beta_{ij}$. So we have shown that β_{ji} is independent of x_j and x_i for i,

$j \in S_l$. As $\sum_{i \in S_l} x_i \beta_{ji}$ is independent of x_j and B_i is either nonsingular or zero, it follows that $B_i = 0$ for all $1 \le i \le m$. Thus $A_2^{(j)} = 0$. This completes the proof of (i).

Case (ii): $j > k$. Let $i_1, i_2 \in S_l$. Then

$$\frac{\partial \beta_{ji_2}}{\partial x_{i_1}} + \frac{\partial \beta_{i_2 i_1}}{\partial x_j} + \frac{\partial \beta_{i_1 j}}{\partial x_{i_2}} = \frac{\partial \omega_{ji_2}}{\partial x_{i_1}} + \frac{\partial \omega_{i_2 i_1}}{\partial x_j} + \frac{\partial \omega_{i_1 j}}{\partial x_{i_2}} = 0. \tag{4.19}$$

Case (i) above implies that $\beta_{i_2 i_1}$ is independent of $x_{k+1}, \ldots, x_n$ for $i_1, i_2 \in S_l$. Therefore (4.19) reduces to

$$\frac{\partial \beta_{ji_2}}{\partial x_{i_1}} = \frac{\partial \beta_{ji_1}}{\partial x_{i_2}}. \tag{4.20}$$

Equation (4.20) shows that $A_1^j(i_1, i_2) = A_1^{(j)}(i_2, i_1)$. Therefore, from (4.17) we can take $A_1^{(j)} = \lambda I$. ∎

Proposition 4.6. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. With the same notation as before, $\beta_{ij} = 0$ for $i, j \in S_l = \{k_{l-1} + 1, \ldots, k_l\}$. That is, $A_1^{(j)} = 0$ in* (i) *of Proposition* 4.5.

Proof. Without loss of generality, we assume that $l = 1$. If $|S_1| = 2$, then, by (i) of Proposition 4.5,

$$\begin{pmatrix} \beta_{11} \\ \beta_{12} \end{pmatrix} = \begin{pmatrix} 0 & a \\ -a & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \beta_{21} \\ \beta_{22} \end{pmatrix} = \begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Since $\beta_{11} = 0$ and $\beta_{22} = 0$, we conclude that $a = b = 0$. Hence, $\beta_{12} = 0 = \beta_{21}$.

If $|S_1| \ge 3$, then by (i) of Proposition 4.5 there exist matrices $A_1, A_2, A_3, \ldots, A_{k_l}$ such that

$$\begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \beta_{13} \\ \vdots \\ \beta_{1k_1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \cdots & * \\ 0 & 0 & a & \cdots & * \\ 0 & -a & 0 & \cdots & * \\ \vdots & \vdots & \vdots & & \\ * & * & * & \cdots & * \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix} = A_1 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix},$$

$$\begin{bmatrix} \beta_{21} \\ \beta_{22} \\ \beta_{23} \\ \vdots \\ \beta_{2k_1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & b & \cdots & * \\ 0 & 0 & 0 & \cdots & * \\ -b & 0 & 0 & \cdots & * \\ \vdots & \vdots & \vdots & & \\ * & * & * & \cdots & * \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix} = A_2 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix},$$

$$\begin{bmatrix} \beta_{31} \\ \beta_{32} \\ \beta_{33} \\ \vdots \\ \beta_{3k_1} \end{bmatrix} = \begin{bmatrix} 0 & c & 0 & \cdots & * \\ -c & 0 & 0 & \cdots & * \\ 0 & 0 & 0 & \cdots & * \\ \vdots & \vdots & \vdots & & \\ * & * & * & \cdots & * \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix} = A_3 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix},$$

$$\vdots$$

$$\begin{bmatrix} \beta_{k_1 1} \\ \beta_{k_1 2} \\ \beta_{k_1 3} \\ \vdots \\ \beta_{k_1 k_1} \end{bmatrix} = \begin{bmatrix} * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \\ * & * & * & \cdots & * \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix} = A_{k_1} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{k_1} \end{bmatrix}.$$

Notice that $\beta_{12} = -\beta_{21}$ and $\beta_{13} = -\beta_{31}$. So $b = c = -a$. On the other hand,

$$\frac{\partial \beta_{12}}{\partial x_3} + \frac{\partial \beta_{23}}{\partial x_1} + \frac{\partial \beta_{31}}{\partial x_2} = a + (-b) + c = 3a = 0.$$

Hence, $a = 0$. Let $A_j^{(l_1, l_2, l_3)}$ be the 3×3 matrix obtained by restricting A_j on l_1, l_2, l_3 columns and l_1, l_2, l_3 rows. We have shown that $A_j^{1,2,3} = 0$ for $j = 1, 2, 3$. Similarly, by the cyclic relation (ii) of Lemma 4.2, we can prove that $A_j^{(l_1, l_2, l_3)} = 0$ for $j = l_1$, l_2, l_3. By considering $\{l_1, l_2, l_3\} = \{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 2, 5\}$, $\ldots$, $\{1, 2, k_1\}$, we see that the first two column vectors and the first two row vectors of A_1 are zero vectors because of the fact that $A_1^{(1,2,3)} = A_1^{(1,2,4)} = \cdots = A_1^{(1,2,k_1)} = 0$. Similarly $A_1^{(1,3,4)} = A_1^{(1,3,5)} = \cdots = A_1^{(1,3,k_1)} = 0$ imply that the first three column vectors and the first three row vectors of A_1 are zero vectors. By induction, we see that $A_1 = 0$.

Observe that the first row of A_2 is equal to the negative of the second row of A_1. So the first row vector and the first column vector of A_2 are zero. $A_2^{(2,3,4)} = A_2^{(2,3,5)} = \cdots = A_2^{(2,3,k_1)} = 0$ imply that the first three column vectors and the first three row vectors of A_2 are zero vectors. Similarly $A_2^{(2,4,5)} = A_2^{(2,4,6)} = \cdots = A_2^{(2,4,k_1)} = 0$ imply that the first four row vectors and the first four column vectors of A_2 are zero vectors. By induction, we see that $A_2 = 0$.

Suppose $A_1 = A_2 = \cdots = A_{j-1}$. We want to prove $A_j = 0$. Observe that the l row of A_j is equal to the negative of the j row of A_l. So we conclude that the first $j - 1$ row vectors and the first $j - 1$ column vectors of A_j are zero. $A_j^{j,j+1,j+2} = A_j^{j,j+1,j+3} = \cdots = A_j^{j,j+1,k_1} = 0$ imply that the first $j + 1$ row vectors and the first $j + 1$ column vectors of A_j are zero vectors. Similarly $A_j^{j,j+2,j+3} = A_j^{j,j+2,j+4} = \cdots = A_j^{j,j+2,k_1} = 0$ imply that the first $j + 2$ row vectors and the first $j + 2$ column vectors of A_j are zero vectors. By induction we see that $A_j = 0$.

Since we have shown by induction that $A_1, \ldots, A_{k_1}$ are zero, we conclude that $\beta_{ij} = 0$ for $1 \leq i, j \leq k_1$. ■

To finish the proof of the main theorem, we need the following proposition.

Proposition 4.7. *Suppose that E is a finite-dimensional estimation algebra of maximal rank. Then* (i) *ω_{ij} are degree one polynomials in $x_1, \ldots, x_k$ for $1 \leq i \leq k$ or $1 \leq j \leq k$, and* (ii) *ω_{ij} are degree one polynomials in $x_{k+1}, \ldots, x_n$ for $k + 1 \leq i, j \leq n$.*

Proof. (i) For $1 \leq i, j \leq k$, we know from Propositions 4.4 and 4.6 that ω_{ij} are constants. For $j > k$, $\sum_{i=1}^k x_i \omega_{ij} = -[\sum_{i=1}^k x_i D_i, D_j] \in E$. Since $\sum_{i=1}^k x_i \omega_{ij}$ as quadratic polynomial in E cannot depend on $x_{k+1}, \ldots, x_n$ variables, it follows that ω_{ij}, for $1 \leq i \leq k$, $j > k$, are degree one polynomial independent of $x_{k+1}, \ldots, x_n$. This completes the proof (i).

(ii) Let $k, j > k$ and $l \leq k$. We have

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

By (i), $\partial \omega_{li}/\partial x_j = \partial \omega_{jl}/\partial x_i = 0$. Hence, we have $\partial \omega_{ij}/\partial x_l = 0$. This means that ω_{ij} are independent of $x_1, x_2, \ldots, x_k$ for $i, j \geq k + 1$. ∎

Remark. The main theorem in Section 1 follows from Theorem 4.3 and Propositions 4.4, 4.6, and 4.7.

References

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

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

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

[CLY] J. Chen, C. W. Leung, and S. S.-T. Yau, Finite-dimensional filters with nonlinear drift, IV: Classification of finite-dimensional estimation algebra of maximal rank with state space dimension 3, *SIAM J. Control Optim.* **34**(1) (1996), 179–198.

[CY] W. L. Chiou and S. S.-T. Yau, Finite-dimensional filters with nonlinear drift, II: Brockett's problem on classification of finite-dimensional estimation algebras, *SIAM J. Control Optim.* **32**(1) (1994), 297–310.

[C] M. Cohen De Lara, Contribution des methods geometriques au filtrage de dimension finie, Ph.D. thesis, Ecole des Mines de Paris, 1991.

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

[DM] M. H. A. Davis and S. I. Marcus, An introduction to nonlinear filtering, in *The Mathematics of Filtering and Identification and Applications*, M. Hazewinkel and J. C. Willems, eds., Reidel, Dordrecht, 1981, pp. 53–75.

[DTWY] R. T. Dong, L. F. Tam, W. S. Wong, and S. S.-T. Yau, Structure and classification theorems of finite dimensional exact estimation algebras, *SIAM J. Control Optim.* **29**(4) (1991), 866–877.

[DWY] R. T. Dong, W. S. Wong, and S. S.-T. Yau, Filtering systems with finite dimensional estimation algebras, Preprint, 1995.

[FKK] M. Fujisaki, G. Kallianpur, and H. Kunita, Stochastic differential equations for the nonlinear filtering problem, *Osaka J. Math.* **1** (1972), 19–40.

[HP] U. G. Haussman and E. Pardoux, A conditionally almost linear filtering problem with non-Gaussian initial condition, *Stochastics* **23** (1988), 241–275.

[H] H. Hazewinkel, Lecture on linear and nonlinear filtering, in *Analysis and Estimation of Stochastic Mechanical Systems*, CISM Courses and Lectures, vol. 303, W. Shiehlen and W. Wedig, eds., Springer-Verlag, Vienna, 1988, pp. 103–135.

[HMS] M. Hazewinkel, S. I. Marcus, and H. J. Sussmann, Nonexistence of finite dimensional filters for conditional statistics of the cubic sensor problem, *Systems Control Lett.* **3** (1983), 331–340.

[M1] A. Makowski, Filtering formula for partially observed linear systems with non-Gaussian initial conditions, *Stochastics* **16** (1986), 1–24.

[M2] S. I. Marcus, Algebraic and geometric methods in nonlinear filtering, *SIAM J. Control Optim.* **22** (1984), 817–844.

[M3] S. K. Mitter, On the analogy between mathematical problems of non-linear filtering and quantum physics, *Ricerche Automat.* **10** (1979), 163–216.

[O] D. Ocone, Finite dimensional estimation algebras in nonlinear filtering, in *The Mathematics of Filtering and Identification and Applications*, M. Hazewinkel and J. C. Willems, eds., Reidel, Dordrecht, 1981, pp. 629–636.

[R] B. L. Rozovsky, Stochastic partial differential equations arising in nonlinear filtering problem (in Russian), *Uspekhi Mat. Nauk.* **27** (1972), 213–214.

[TWY] L. F. Tam, W. S. Wong, and S. S.-T. Yau, On a necessary and sufficient condition for finite dimensionality of estimation algebras, *SIAM J. Control Optim.* **28** (1990), 173–185.

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

[Y] S. S.-T. Yau, Finite-dimensional filters with nonlinear drift, I: A class of filters including both Kalman–Bucy filters and Benes filters, *J. Math. Systems Estim. Control* **4**(2) (1994), 181–203.

[YL] S. S.-T. Yau and C. W. Leung, Recent result on classification of finite dimensional maximal rank estimation algebras with state space dimension 3, *Proceedings of the 31st Conference on Decision and Control*, Tucson, Arizona, Dec. 1992, pp. 2247–2250.