

HESSIAN MATRIX NON-DECOMPOSITION THEOREM

STEPHEN S.-T. YAU*, XI WU*, AND WING-SHING WONG†

ABSTRACT. In his 1983 invited lecture at the International Congress of Mathematics, Roger Brockett proposed to classify finite dimensional estimation algebras. The following problem arises from the first author's classification theory of finite dimensional estimation algebras with maximal rank. Can the Hessian matrix of a homogeneous polynomial of degree 4 be decomposed in the form $\Delta(x)\Delta(x)^T$ where $\Delta(x)$ is an anti-symmetric linear matrix (i.e., entries of $\Delta(x)$ are linear in x)? In this short note, we show that this cannot be true, in other words, the Hessian matrix is nondecomposable in this form.

1. Introduction

In his 1983 invited lecture at the International Congress in Mathematics, Roger Brockett [2] proposed to classify finite dimensional estimation algebras. The concept of an estimation algebra (Lie algebra) was first introduced by Brockett and Clark [3], Brockett [1] and Mitter [10] independently. This concept plays a crucial role in the investigation of finite-dimensional nonlinear filters which are important in both commercial and military industries. Since 1990, Yau [15], [16] has launched a program to study Brockett's problem. He first considered Wong's anti-symmetric matrix $\Omega = (\omega_{ij})$ [13]. He solved the Brockett's problem when Ω matrix has only constant entries. Yau's program is to show that Ω matrix must have constant entries for a finite dimensional estimation algebra. Recently Chen and Yau [4] studied the structure of quadratic forms in a finite dimensional estimation algebra. Let k be the quadratic rank of the estimation algebra and n be the dimension of the state space of the estimation algebra. They showed that the left upper corner (ω_{ij}) , $1 \leq i, j \leq k$, of Ω matrix is a matrix with constant coefficients. In 1997, the first and second authors and Hu showed that the lower right corner (ω_{ij}) , $k + 1 \leq i, j \leq n$, of Ω matrix is also a matrix with constant coefficients. In fact, they proved the weak form of Hessian matrix nondecomposition theorem, i.e., the matrix equation $\Delta\Delta^T = H(\eta_4)$ has no non-trivial solution where $\Delta = (\beta_{ij})$ is an anti-symmetric linear matrix (i.e., entries are linear functions), $\frac{\partial\beta_{ij}}{\partial x_\ell} + \frac{\partial\beta_{\ell i}}{\partial x_j} + \frac{\partial\beta_{j\ell}}{\partial x_i} = 0$ (cyclic condition) and η_4 is a homogeneous polynomial of degree 4. This result together with the result of [17] solve the Brockett's problem of classification of finite dimensional estimation algebras

Received September 21, 1999.

* Research partially supported by U.S. Army Research Office DAAD19-99-1-0203.

† Research partially supported by Hong Kong RGC.

with maximal rank. Since early 1990, there have been a lot of interest on the strong form of Hessian matrix nondecomposition. The purpose of this paper is to prove that this is indeed true.

Main Theorem. *Let $\eta_4(x_1, \dots, x_n)$ be a homogeneous polynomial of degree 4 in x_1, \dots, x_n over \mathbb{R} . Let $H(\eta_4) = (\frac{\partial^2 \eta_4}{\partial x_i \partial x_j})_{1 \leq i, j \leq n}$ be the Hessian matrix of η_4 . Then $H(\eta_4)$ cannot be decomposed as $\Delta(x)\Delta(x)^T$, 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., $H(\eta_4)(x) = \Delta(x)\Delta(x)^T$ implies $\Delta = 0$ and $\eta_4 = 0$.*

The anti-symmetry of $\Delta(x)$ is essential in the Theorem because of the following example:

Let $\eta(x_1, x_2) = \frac{1}{6}x_1^4 + \frac{2}{3}x_1^3x_2 + x_1^2x_2^2 + \frac{2}{3}x_1x_2^3 + \frac{1}{6}x_2^4$, then

$$\begin{aligned} H(\eta) &= \begin{pmatrix} 2x_1^2 + 4x_1x_2 + 2x_2^2 & 2x_1^2 + 4x_1x_2 + 2x_2^2 \\ 2x_1^2 + 4x_1x_2 + 2x_2^2 & 2x_1^2 + 4x_1x_2 + 2x_2^2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 + x_2 & x_1 + x_2 \\ x_1 + x_2 & x_1 + x_2 \end{pmatrix} \begin{pmatrix} x_1 + x_2 & x_1 + x_2 \\ x_1 + x_2 & x_1 + x_2 \end{pmatrix} = \Delta\Delta^T, \end{aligned}$$

where

$$\Delta = \begin{pmatrix} x_1 + x_2 & x_1 + x_2 \\ x_1 + x_2 & x_1 + x_2 \end{pmatrix}.$$

Let us write $\Delta(x) = A_1x_1 + A_2x_2 + \dots + A_nx_n$ in our Main Theorem, where A_ℓ is a $n \times n$ anti-symmetric matrix with real constant coefficients. Then the equation $H(\eta_4)(x) = \Delta(x)\Delta(x)^T$ will give us a lot of quadratic equations in $A_\ell(i, j)$ ((i, j) entry of the matrix A_ℓ), $1 \leq i, j, \ell \leq n$ (see (3.1) below for example). Although it is possible to prove that these quadratic equations can have only trivial solution for $n \leq 4$ (see [6], pp. 1137–1138), it has been a challenging problem to algebraic geometers whether this system of quadratic equations in $A_\ell(i, j)$ can only admit trivial solution over \mathbb{R} even for $n = 5$.

Our Main Theorem above has led us to formulate the following conjecture which may be proved by generalizing the technique developed in this paper:

Conjecture. *For $l \geq 2$, let $\eta_{2l}(x_1, \dots, x_n)$ be a homogeneous polynomial of degree $2l$ in x_1, \dots, x_n over \mathbb{R} . Let $H(\eta_{2l})$ be the Hessian matrix of η_{2l} . Then $H(\eta_{2l})$ cannot be decomposed as $\Delta(x)\Delta(x)^T$, where $\Delta(x) = (\beta_{ij})_{1 \leq i, j \leq n}$ is an anti-symmetric matrix with β_{ij} homogeneous polynomials in x of degree $l - 1$, unless η_{2l} and Δ are trivial, i.e., $H(\eta_{2l})(x) = \Delta(x)\Delta(x)^T$ implies $\Delta = 0$ and $\eta_{2l} = 0$.*

We thank the referee and Daniel Stroock for their useful suggestions of revising this paper.

2. Classification of finite dimensional estimation algebras with maximal rank

In this section, we recall some basic concepts and results from previous papers [4][7] [14][15][16]. The filtering problem considered here is based on the signal observation model

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

in which x, v, y , and w are respectively $\mathbb{R}^n, \mathbb{R}^p, \mathbb{R}^m$, and \mathbb{R}^m valued processes and v and w have components that are independent, standard Brownian processes. We further assume that $n = p; f, g$ and h are vector-valued, orthogonal matrix-valued and vector-valued C^∞ smooth functions. We shall refer to $y(t)$ as the observation at time t .

Let $\rho(t, x)$ denote the conditional probability density of the state given the observation $\{y(s) : 0 \leq s \leq t\}$. It is well known (see [9], for example) that $\rho(t, x)$ is given by normalizing a function $\sigma(t, x)$ that satisfies the following DMZ equation:

$$(2.2) \quad \begin{cases} d\sigma(t, x) = L_0\sigma(t, x)dt + \sum_{i=1}^m L_i\sigma(t, x)dy_i(t), \\ \sigma(0, x) = \sigma_0(x), \end{cases}$$

where

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

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

Equation (2.1) is a stochastic partial differential equation. In real applications, we are interested in constructing robust state estimators from observed sample paths with some property of robustness. Davis in [8] 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 (2.2) to the following time-varying partial differential equation, which is called the robust DMZ equation:

$$(2.3) \quad \begin{cases} \frac{\partial u}{\partial t}(t, x) = L_0u(t, x) + \sum_{i=1}^m y_i(t)[L_0, L_i]u(t, x) \\ \quad + \frac{1}{2} \sum_{i,j=1}^m y_i(t)y_j(t)[[L_0, L_i], L_j]u(t, x), \\ u(0, x) = \sigma_0(x), \end{cases}$$

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

Definition 2.1. If X and Y are differential operators, the Lie bracket of X and Y , $[X, Y]$, is

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

for any C^∞ function ψ .

Definition 2.2. The estimation algebra E of a filtering problem (2.1) is defined as the Lie algebra generated by $\{L_0, L_1, \dots, L_m\}$. E is said to be an estimation algebra with maximal rank if, for any $1 \leq i \leq n$, there exists a constant c_i such that $x_i + c_i$ is in E .

The concept of estimation algebra was first introduced by Brockett[1], Brockett and Clark[3], and Mitter[10] independently. The definition of estimation algebra with maximal rank was introduced by Chiou and Yau[7]. In his 1983 invited lecture at the International Congress of Mathematics, Brockett proposed to classify finite dimensional estimation algebras. This is an important problem because as long as the estimation algebra is finite dimensional, one can use Weierstrass approach [12] to construct finite dimensional filter for (2.3) (see [15][16] for example). In fact, it was shown in [16] that (2.3) is reduced to a system of linear ordinary differential equations and a Kolmogorov equation independent of observation $y(t)$.

Let $\Omega = (\omega_{ij})$ be a $n \times n$ matrix with

$$\omega_{ij} = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}, \quad \forall 1 \leq i, j \leq n.$$

Clearly, Ω is skew symmetric.

Define

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

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

Then

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

The following basic results play a fundamental role in the classification of finite-dimensional estimation algebras.

Theorem 2.1 ([11]). *Let E be a finite-dimensional estimation algebra. If a function ψ is in E , then ψ is a polynomial of degree ≤ 2 .*

Theorem 2.2 ([16]). *Let E be a finite-dimensional estimation algebra of (2.1) such that ω_{ij} are constant functions. If E is with 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 .*

Corollary 2.1 ([16]). *Let E be a finite-dimensional estimation algebra with maximal rank. Then E contains the real vector space spanned by $1, x_1, \dots, x_n, D_1, \dots, D_n$ and L_0 .*

Definition 2.3. Let Q be the space of quadratic forms in n variables, namely, the real vector space spanned by $x_i x_j$, with $1 \leq i \leq j \leq n$. Let $X = (x_1, \dots, x_n)^T$. For any quadratic form $p \in Q$, there exists a symmetric matrix A such that $p(x) = X^T A X$. 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, $r(p_0) \geq r(p)$ for any $p \in E \cap Q$. The maximal rank of quadratic form in the estimation algebra E is defined to be $k = r(p_0)$ and is called the quadratic rank of E .

Lemma 2.1 ([4]). *If p is a quadratic form in estimation algebra E , then p is independent of x_j for $j > k$, that is, $\frac{\partial p}{\partial x_j} = 0$ for $k + 1 \leq j \leq n$, where k is the quadratic rank of E .*

Lemma 2.2 ([16]).

- (i) $[XY, Z] = X[Y, Z] + [X, Z]Y$, where X, Y , and Z are differential operators;
- (ii) $[gD_i, h] = g \frac{\partial h}{\partial x_i}$, where $D_i = \frac{\partial}{\partial x_i} - f_i$. g and h are functions defined on \mathbb{R}^n ;
- (iii) $[gD_i^2, h] = 2g \frac{\partial h}{\partial x_i} D_i + g \frac{\partial^2 h}{\partial x_i^2}$.

Lemma 2.3 ([14]). *E contains the following elements:*

- (i) $\omega_{ij} = [D_j, D_i] \in E, \quad \forall 1 \leq i, j \leq n$;
- (ii) $H_j := [L_0, D_j] = \sum_{i=1}^n \left(\omega_{ji} D_i + \frac{1}{2} \frac{\partial \omega_{ji}}{\partial x_i} \right) + \frac{1}{2} \frac{\partial \eta}{\partial x_j} \in E, \quad \forall 1 \leq j \leq n$;
- (iii) $[H_j, D_l] = \sum_{i=1}^n \omega_{ji} \omega_{li} - \frac{1}{2} \frac{\partial^2 \eta}{\partial x_i \partial x_j} - \sum_{i=1}^n \frac{\partial \omega_{ji}}{\partial x_i} D_i - \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 \omega_{ji}}{\partial x_i \partial x_i} \in E, \quad \forall 1 \leq j, l \leq n$.

Theorem 2.3 ([4]). *Suppose that E is a finite dimensional estimation algebra with maximal rank and k is the quadratic rank of E . Then*

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

- (i) ω_{ij} 's are constants for $1 \leq i, j \leq k$;
- (ii) ω_{ij} 's are polynomials of degree one in x_1, \dots, x_k for $1 \leq i \leq k$ or $1 \leq j \leq k$;
- (iii) ω_{ij} 's are polynomials of degree one in x_{k+1}, \dots, x_n for $k + 1 \leq i, j \leq n$.

Lemma 2.4 ([5][6]). *Suppose that E is a finite dimensional estimation algebra with maximal rank. Then*

- (i) $1, x_1, \dots, x_n, D_1, \dots, D_n, L_0 \in E$;

- (ii) $\sum_{l=1}^n \omega_{jl}\omega_{il} - \frac{1}{2} \frac{\partial^2 \eta}{\partial x_j \partial x_i} \in E$ for any $1 \leq i, j \leq n$;
- (iii) η is a polynomial of degree 4.

Lemma 2.5 ([14]). *Let $\eta = \eta_4(x_{k+1}, \dots, x_n) + \text{polynomial of degree 3 in } x_{k+1}, \dots, x_n \text{ variables (with degree 4 polynomial coefficients in } x_1, \dots, x_k \text{ variables), where } \eta_4 \text{ is a homogeneous polynomial of degree 4 in } x_{k+1}, \dots, x_n \text{ variables. Then for any } k + 1 \leq i, j \leq n,$*

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

where β_{ij} is the homogenous polynomial of degree one part of ω_{ij} .

As a corollary to our Main Theorem, we have the following result.

Theorem 2.4. *Suppose that E is a finite dimensional estimation algebra with maximal rank and k is the quadratic rank of E . Then*

$$\Omega = (\omega_{ij}) = \left(\begin{array}{c|c} \text{Constant} & P_1(x_1, \dots, x_k) \\ \hline \text{-----} & \text{-----} \\ P_1(x_1, \dots, x_k) & \text{Constant} \end{array} \right).$$

- (i) ω_{ij} 's are constants for $1 \leq i, j \leq k$;
- (ii) ω_{ij} 's are polynomials of degree one in x_1, \dots, x_k for $1 \leq i \leq k$ or $1 \leq j \leq k$;
- (iii) ω_{ij} 's are constants for $k + 1 \leq i, j \leq n$.

Proof. (i) and (ii) are in Theorem 2.3. (iii) follows from Lemma 2.5 and our Main Theorem. □

In [17], Yau and Hu proved also that ω_{ij} 's are constants for $1 \leq i \leq k$ or $1 \leq j \leq k$. Therefore the classification of finite dimensional estimation algebras with maximal rank is completed according to Theorem 2.2.

Theorem 2.5. *Let E be a finite dimensional estimation algebra with 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 .*

3. Proof of the main theorem

Lemma 3.1. *Let $\eta_4(x_1, \dots, x_n)$ be a homogeneous polynomial of degree 4 in x_1, \dots, x_n over \mathbb{R} . Let $H(\eta_4) = \left(\frac{\partial^2 \eta_4}{\partial x_i \partial x_j} \right)_{1 \leq i, j \leq n}$ be the Hessian matrix of η_4 . Let $\Delta(x) = (\beta_{ij})_{1 \leq i, j \leq n} := A_1 x_1 + \dots + A_n x_n$ where $A_\ell = (A_\ell(i, j))_{1 \leq i, j \leq n}$ are $n \times n$ anti-symmetric matrices with coefficients in \mathbb{R} . Suppose that $H(\eta_4)(x) =$*

$\Delta(x)\Delta(x)^T$. Then

$$(3.1) \quad \sum_{\ell=1}^n [A_i(j, \ell)]^2 = \sum_{\ell=1}^n [A_j(i, \ell)]^2 \\ = \frac{1}{2} \sum_{\ell=1}^n [A_i(i, \ell)A_j(j, \ell) + A_i(j, \ell)A_j(i, \ell)].$$

Proof. Observe that $H(\eta_4)(x) = \Delta(x)\Delta(x)^T$ implies

$$(3.2) \quad \frac{\partial^2 \eta_4}{\partial x_i \partial x_j} = \sum_{\ell=1}^n \beta_{i\ell} \beta_{j\ell}.$$

Since $\frac{\partial^2}{\partial x_i^2} \left(\frac{\partial^2 \eta_4}{\partial x_j^2} \right) = \frac{\partial^2}{\partial x_j^2} \left(\frac{\partial^2 \eta_4}{\partial x_i^2} \right) = \frac{\partial^2}{\partial x_i \partial x_j} \left(\frac{\partial^2 \eta_4}{\partial x_i \partial x_j} \right)$, we have

$$\frac{\partial^2}{\partial x_i^2} \left(\sum_{\ell=1}^n \beta_{j\ell}^2 \right) = \frac{\partial^2}{\partial x_j^2} \left(\sum_{\ell=1}^n \beta_{i\ell}^2 \right) = \frac{\partial^2}{\partial x_i \partial x_j} \left(\sum_{\ell=1}^n \beta_{i\ell} \beta_{j\ell} \right).$$

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

$$(3.3) \quad 2 \sum_{\ell=1}^n \left(\frac{\partial \beta_{j\ell}}{\partial x_i} \right)^2 = 2 \sum_{\ell=1}^n \left(\frac{\partial \beta_{i\ell}}{\partial x_j} \right)^2 = \sum_{\ell=1}^n \left(\frac{\partial \beta_{i\ell}}{\partial x_i} \frac{\partial \beta_{j\ell}}{\partial x_j} + \frac{\partial \beta_{i\ell}}{\partial x_j} \frac{\partial \beta_{j\ell}}{\partial x_i} \right).$$

As $A_i(j, \ell) = \frac{\partial \beta_{j\ell}}{\partial x_i}$, we see that (3.1) is equivalent to (3.3). \square

Lemma 3.2. Let $\eta(x)$ be a C^∞ function on \mathbb{R}^n . Let $\tilde{\eta}(x) = \eta(Rx)$ where R is a $n \times n$ matrix. Then $H(\tilde{\eta})(x) = R^T H(\eta)(Rx)R$.

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

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

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

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 anti-symmetry of the matrices 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 (3.1) 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

$$\begin{aligned}
 (3.4) \quad \Delta(x)\Delta(x)^T &= H(\eta)(x) \\
 &\implies R^T \Delta(Rx) R R^T \Delta(Rx)^T R \\
 &= R^T H(\eta)(Rx) R \\
 &\implies \tilde{\Delta}(x)\tilde{\Delta}(x)^T = H(\tilde{\eta})(x) \quad \text{by Lemma 3.2}
 \end{aligned}$$

where

$$(3.5) \quad \tilde{\eta}(x) = \eta(Rx),$$

$$\begin{aligned}
 (3.6) \quad \tilde{\Delta}(x) &= R^T \Delta(Rx) R \\
 &= R^T [A_1(r_{11}x_1 + r_{12}x_2 + \cdots + r_{1n}x_n) + \cdots \\
 &\quad + A_n(r_{n1}x_1 + r_{n2}x_2 + \cdots + r_{nn}x_n)] R \\
 &= \tilde{A}_1x_1 + \tilde{A}_2x_2 + \cdots + \tilde{A}_nx_n,
 \end{aligned}$$

where

$$(3.7) \quad \tilde{A}_\ell = R^T A_1 R r_{1\ell} + R^T A_2 R r_{2\ell} + \cdots + R^T A_n R r_{n\ell}, \quad 1 \leq \ell \leq n,$$

$$(3.8) \quad \tilde{A}_\ell^T = -\tilde{A}_\ell.$$

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

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

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

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

Then

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

i.e., $(\tilde{A}_1(1, 2), \tilde{A}_1(1, 3), \dots, \tilde{A}_1(1, n)) = (a, 0, \dots, 0)$. By applying Lemma 3.1 to (3.4), we have

$$\begin{aligned} \sum_{\ell=1}^n [\tilde{A}_1(2, \ell)]^2 &= \sum_{\ell=1}^n [\tilde{A}_2(1, \ell)]^2 = \frac{1}{2} \sum_{\ell=1}^n [\tilde{A}_1(1, \ell)\tilde{A}_2(2, \ell) + \tilde{A}_1(2, \ell)\tilde{A}_2(1, \ell)] \\ &= \frac{1}{2} \sum_{\ell=1}^n \tilde{A}_1(2, \ell)\tilde{A}_2(1, \ell) \\ &\leq \frac{1}{4} \sum_{\ell=1}^n [\tilde{A}_1(2, \ell)]^2 + \frac{1}{4} \sum_{\ell=1}^n [\tilde{A}_2(1, \ell)]^2 \\ \implies \frac{3}{4} \sum_{\ell=1}^n [\tilde{A}_1(2, \ell)]^2 &\leq \frac{1}{4} \sum_{\ell=1}^n [\tilde{A}_2(1, \ell)]^2 \\ \frac{3}{4} \sum_{\ell=1}^n [\tilde{A}_2(1, \ell)]^2 &\leq \frac{1}{4} \sum_{\ell=1}^n [\tilde{A}_1(2, \ell)]^2 \\ \implies \sum_{\ell=1}^n [\tilde{A}_1(2, \ell)]^2 &= 0 \\ \implies \tilde{A}_1(1, 2) &= -\tilde{A}_1(2, 1) = 0. \end{aligned}$$

This contradicts the fact that $\tilde{A}_1(1, 2) = a \neq 0$. Therefore we conclude that $A_1(1, \ell) = 0, 1 \leq \ell \leq n$. Now we apply Lemma 3.1 with $i = 1, 2 \leq j \leq n$. Then we get

$$\begin{aligned} \sum_{\ell=1}^n [A_1(j, \ell)]^2 &= \sum_{\ell=1}^n [A_j(1, \ell)]^2 = \frac{1}{2} \sum_{\ell=1}^n A_1(j, \ell)A_j(1, \ell) \\ &\leq \frac{1}{4} \sum_{\ell=1}^n [A_1(j, \ell)]^2 + \frac{1}{4} \sum_{\ell=1}^n [A_j(1, \ell)]^2 \\ \implies \frac{3}{4} \sum_{\ell=1}^n [A_1(j, \ell)]^2 &\leq \frac{1}{4} \sum_{\ell=1}^n [A_j(1, \ell)]^2 \\ \frac{3}{4} \sum_{\ell=1}^n [A_j(1, \ell)]^2 &\leq \frac{1}{4} \sum_{\ell=1}^n [A_1(j, \ell)]^2 \\ \implies \sum_{\ell=1}^n [A_1(j, \ell)]^2 &= 0 = \sum_{\ell=1}^n [A_j(1, \ell)]^2 \\ \implies A_1 &= 0 \text{ and } A_\ell = \left(\begin{array}{c|ccc} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & & \\ 0 & & & B_\ell \end{array} \right) \quad 2 \leq \ell \leq n \end{aligned}$$

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

Let $\bar{x} = (x_2, \dots, x_n)$ and $\bar{\Delta}(\bar{x}) = B_2x_2 + \dots + B_nx_n$. Then

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

Since

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

we have

$$\frac{\partial^2 \eta_4}{\partial x_1 \partial x_\ell} = 0, \quad 1 \leq \ell \leq n.$$

Thus η_4 is independent of x_1 variable. Denote $\bar{\eta}_4 = \eta_4(x_2, \dots, x_n)$. Then we have $H(\bar{\eta}_4) = \left(\frac{\partial^2 \eta}{\partial x_i \partial x_j} \right)_{2 \leq i, j \leq n} = \bar{\Delta}(\bar{x})\bar{\Delta}(\bar{x})^T$. By induction hypothesis, we have $\bar{\Delta}(\bar{x}) = 0$. Therefore $\Delta(x) = 0$. \square

References

- [1] R.W. Brockett, *Nonlinear systems and nonlinear estimation theory*, in *Stochastic systems: the mathematics of filtering and identification and applications (Les Arcs, 1980)*, pp. 441–477, NATO Adv. Study Inst. Ser. C: Math. Phys. Sci., 78, Reidel, Dordrecht-Boston, Mass., 1981.
- [2] ———, *Nonlinear control theory and differential geometry*, Proceedings of the International Congress of Mathematics, Vol. 1, 2 (Warsaw, 1983), 1357–1368, PWN, Warsaw, 1984.
- [3] R.W. Brockett and J.M.C. Clark, *The geometry of the conditional density functions*, in *Analysis and optimization of stochastic systems*, (Proc. Internat. Conf., Univ. Oxford, Oxford, 1978), pp. 299–309, Academic Press, London-New York, 1980.
- [4] J. Chen and S.S.-T. Yau, *Finite dimensional filters with nonlinear drift VI: Linear structure of Ω* , Math. Control Signals Systems **9** (1996), 370–385.
- [5] ———, *Finite-dimensional filters with nonlinear drift VII: Mitter conjecture and structure of η* , SIAM J. Control Optim. **35** (1997), 1116–1131.
- [6] J. Chen, S.S.-T. Yau, and C.W. Leung, *Finite dimensional filters with nonlinear drift VIII: Classification of finite dimensional estimation algebra of maximal rank with state space dimension 4*, SIAM J. Control Optim. **35** (1997), 1132–1141.
- [7] W.L. Chiou and Stephen 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** (1994), 297–310.
- [8] M.H.A. Davis, *On a multiplicative functional transformation arising in nonlinear filtering theory*, Z. Wahrsch Verw. Gebiete **54** (1980), 125–139.
- [9] M.H.A. Davis and S.I. Marcus, *An introduction to nonlinear filtering*, in *The Mathematics of Filtering and Identification and Applications*, (M. Hazewinkel and J. S. Willems, eds.), Reidel, Dordrecht, 1981.
- [10] S.K. Mitter, *On the analogy between mathematical problems of nonlinear filtering and quantum physics*, Ricerche Automat. **10** (1979), 163–216.

- [11] D. Ocone, *Topics in nonlinear filtering theory*, Ph.D. thesis, Massachusetts Institute of Technology, 1980.
- [12] J. Wei and E. Norman, *On the global representation of the solutions of linear differential equations as a product of exponentials*, Proc. Amer. Math. Sci. *15* (1964), 327–334.
- [13] W.S. Wong, *On a new class of finite dimensional estimation algebras*, Systems Control Lett., **9** (1987), 79–83.
- [14] X. Wu, S.S.-T. Yau and G.Q. Hu, *Finite dimensional filters with nonlinear drift XII: Linear and constant structure of Ω matrix*, preprint, 1997.
- [15] S.S.-T. Yau, *Recent results on nonlinear filtering: New class of finite dimensional filters*, Proceedings of the 29th Conference on Decision and Decision and Control at Honolulu, Hawaii, Dec. (1990), 231–233.
- [16] _____, *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** (1994), 181–203.
- [17] S.S.-T. Yau and G.-Q. Hu, *Finite dimensional filters with nonlinear drift XIV: Classification of finite-dimensional estimation algebra of maximal rank with arbitrary state space dimension and Mitter conjecture*, preprint, 1998.

DEPARTMENT OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCES (M/C 249), UNIVERSITY OF ILLINOIS AT CHICAGO, 851 SOUTH MORGAN STREET, CHICAGO, IL 60607-7045
E-mail address: yau@uic.edu, wu@math.uic.edu

DEPARTMENT OF INFORMATION ENGINEERING, THE CHINESE UNIVERSITY OF HONG KONG, SHATIN, N.T., HONG KONG
E-mail address: wswong@ie.cuhk.edu.hk