International Journal of Control

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tcon20

Finite-dimensional estimation algebra on arbitrary state dimension with nonmaximal rank: linear structure of Wong matrix

Xiaopei Jiao & Stephen S.-T. Yau

INTERNATIONAL JOURNAL OF CONTROL
https://doi.org/10.1080/00207179.2023.2291402

Finite-dimensional estimation algebra on arbitrary state dimension with nonmaximal rank: linear structure of Wong matrix

Xiaopei Jiao[a,b] and Stephen S.-T. Yau[a,b]

[a]Yanqi Lake Beijing Institute of Mathematical Science and Applications, Beijing, People's Republic of China; [b]Department of Mathematical Sciences, Tsinghua University, Beijing, People's Republic of China

ABSTRACT

Ever since Brockett, Clark and Mitter introduced the estimation algebra method, it becomes a powerful tool to classify the finite-dimensional filtering system. In this paper, we investigate finite-dimensional estimation algebra with non-maximal rank. The structure of Wong matrix Ω will be focused on since it plays a critical role in the classification of finite-dimensional estimation algebras. In this paper, we first consider general estimation algebra with non-maximal rank and determine the linear structure of the submatrix of Ω by using rank condition and property of Euler operator. In the second part, we proceed to consider the case of linear rank $n-1$ and prove the linear structure of Ω. Finally, we give the structure of finite-dimensional filters which implies the drift term must be a quadratic function plus a gradient of a smooth function.

ARTICLE HISTORY
Received 17 February 2023
Accepted 30 November 2023

KEYWORDS
State estimation; finite-dimensional filter; estimation algebra; classification

1. Introduction

Filtering problem refers to estimating the state of a stochastic dynamical system by using the information of observation history. In the 1960s, a breakthrough in estimation technique which is a well-known Kalman–Bucy filter (Kalman & Bucy, 1961) appeared. Linear Kalman filtering motivated numerous research in the study of nonlinear filtering. In the sense of mean square error, optimal estimate of nonlinear filtering is the conditional expectation $\mathbb{E}[\phi(X_t) \mid \mathcal{Y}_t]$, where X_t denotes the state of system and $\mathcal{Y}_t$ denotes the history of observations. ϕ is a smooth function. Obviously, conditional density $\rho(t,x)$ contains full information of nonlinear filtering. In the 1960s. Kushner–Stratonovich equation was proposed which describes the evolution of conditional density. However, Kushner equation is hard to solve directly because it is a nonlinear stochastic partial differential equation. In the late 1960s, Duncan–Mortensen–Zakai (DMZ) equation (Zakai, 1969) was proposed and described the evolution of unnormalised conditional density $\sigma(t,x)$. DMZ equation is a linear stochastic partial differential equation and it is easier to deal with.

Currently there are basically four ways of solving nonlinear filtering problems. The first one is based on the projection method, e.g. extended Kalman filter and geometric projection filter (Brigo et al., 1998). Projection algorithms aim to project the nonlinear system to a well-studied system approximately. The second approach is the particle evolution method. Here different particles $\{X_i\}$ will be created and evolved according to the governing stochastic differential equation. Then empirical distribution is used to approximate the conditional density $\rho \approx \frac{1}{N}\sum_{i=1}^{N}\delta(x-X_i)$. Typical algorithms include ensemble Kalman filter (Epstein, 1969) and feedback particle filter (Yang et al., 2013). The third aspect is to reformulate the original filtering system to an optimisation framework. Most recently developed and popular techniques include optimal control and optimal transportation (Kim et al., 2019; Zhang et al., 2021). The fourth aspect is based on the solution of DMZ which will be explained in detail below.

By a reversible exponential mapping (Rozovsky, 1972), a path-wise robust version of DMZ equation was proposed which is a deterministic PDE with time-varying coefficients. In recent decades, numerous works were developed based on the solution of robust DMZ equation. Yau and Yau (1996) discussed the general Kolmogorov equation when observation is always zero and find a fundamental solution of Kolmogorov equation. Yau and Yau (2004) proceeded and restricted the system to finite-dimensional case and obtained the fundamental solution of Schrodinger equation. Subsequently, this fundamental solution technique was extended to a time-varying filtering system (Chen et al., 2017) and is named 'Direct method' (Chen et al., 2017; Shi et al., 2018). In 2008, Yau-Yau filtering algorithm (Yau & Yau, 2008) was proposed and robust DMZ equation was transformed to a series of Kolmogorov equations. Yau–Yau algorithm separated the filtering process to on-line and off-line parts. Later, Hermite spectral method (HSM) was proposed to effectively solve forward Kolmogorov equation (Luo & Yau, 2013), which is an off-line part of Yau–Yau algorithm. HSM is an effective method and has complete theoretical analysis for filtering problems.

It is noted that DMZ equation is a linear SPDE and Wei–Norman approach can be applied to solve it in principle.

CONTACT Stephen S.-T. Yau yau@uic.edu Yanqi Lake Beijing Institute of Mathematical Science and Applications, Beijing 101408, People's Republic of China; Department of Mathematical Sciences, Tsinghua University, Beijing 100084, People's Republic of China

This motivated the establishment of estimation algebra method proposed by Brockett (1981); Brockett and Clark (1980); Mitter (1980) in the 1970s. Once estimation algebra of system is a finite-dimensional Lie algebra, Wei–Norman approach will reduce the solution of DMZ equation to a Kolmogorov equation, a system of ordinary differential equations (ODEs) and several first-order linear partial differential equations (PDEs). Therefore, DMZ equation can be solved completely and universal recursive filters can be constructed successfully.

In the International Congress of Mathematicians of 1983, Brockett proposed the program of classifying all finite-dimensional estimation algebra. In the 1990s, through persistent efforts, Yau et al. finished the complete classification of maximal rank estimation algebras (Chen & Yau, 1996; Chiou & Yau, 1994; Yau, 1994, 2003; Yau & Hu, 2005). Since the beginning of twentieth century, Yau et al. have been devoted to the study of nonmaximal rank finite-dimensional estimation algebra. In the classification of full-rank estimation algebra, a critical step is that Yau can prove that entries of Ω are all affine functions. This linear structure of Ω is a key towards Mitter conjecture. This linear structure is also found in the low-dimensional non-maximal rank estimation algebra. More precisely, Wu and Yau (2006) have finished the classification of estimation algebra with state dimension 2 and rank 1 in 2006. In 2017, Shi and Yau studied the situation of state dimension 3 and rank 2 and proved the linear structure of Ω (Shi & Yau, 2017). Based on linear structure of Ω, they succeeded to prove Mitter conjecture (Shi & Yau, 2017). However, in the high-dimensional situation with state dimension larger than 3, whether Ω can keep linear structure remains unknown and is an important open problem in the field of finite-dimensional filters. In this paper, we will give a positive answer for case with state dimension n and linear rank $n-1$.

Complex structure of Ω matrix has been investigated in many papers (Jiao & Yau, 2020; Shi et al., 2017). It is a known fact that under non-maximal rank case estimation algebra, Ω could possess polynomial or even smooth function. In this paper, we will focus on the special case of non-maximal rank estimation algebra, i.e. $r = n-1$. Our ultimate goal in this paper is to prove entries of $\Omega \in \mathbb{R}^{n\times n}$ are all affine functions. To this end, we will finish in two major steps as below.

Step 1. General finite-dimensional estimation algebra with non-maximal rank $r(\leq n-1)$ will be investigated. We will prove the partially linear structure of Ω, i.e. upper-left corner submatrix with size $r \times r$ possesses linear structure. The techniques used here include linear rank and quadratic rank properties as well as the property of Euler operator theory. This result will provide a common starting point for further study on nonmaximal rank cases.

Step 2. With the partially linear structure of Ω, we consider the case of linear rank $r = n-1$. We show that remaining part of Ω has an affine structure. The main method used here is to construct infinite sequences in the estimation algebra. Once infinite sequence is obtained, certain restriction about Ω will be obtained due to the assumption of finite dimensionality of estimation algebra. Following this procedure and utilising contradiction, we will reduce the degree of Ω step by step. Finally, under the help of a lot of useful intermediate conclusions, the degree of Ω eventually can be reduced to just 1.

This paper is organised as follows. In Section 2, we introduce some basic concepts of nonlinear filtering and preliminary results about estimation algebras. In Section 3, the general nonmaximal rank case is considered and it demonstrates that submatrix of Ω is an affine matrix. In Section 4, nonmaximal rank estimation algebra with linear rank $n-1$ is considered and linear structure of Ω is proven. In Section 5, the structure of drift function in the filtering system is shown. In Section 6, we will give a summary. Appendix contains the detailed proofs of related results.

2. Basic concepts and preliminaries

Basic notations: The set of real numbers is denoted by $\mathbb{R}$. $\mathbb{R}^k$ refers to k-dimensional Euclidean space. $A = (a_{ij})$ denotes a matrix A with i,j-entry a_{ij}. $rank(A)$ denotes the rank of matrix A. I_n denotes the identity matrix of size $n \times n$. δ_{ij} denotes Kronecker symbol which means $\delta_{ij} = 1$ if $i = j$ otherwise $\delta_{ij} = 0$. $\mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ represents a diagonal matrix . $A_{1\leq i,j\leq r}$ denotes submatrix consisting of first r rows and columns of A. Let $C^\infty(U)$ be the set of smooth function defined on U. $span\{v_1, \dots, v_k\}$ refers to a linear space spanned by vectors $\{v_1, v_2, \dots, v_k\}$. $P_k(x_{i_1}, \dots, x_{i_m})$ denotes the set of polynomial of degree no more than k in variable $x_{i_1}, \dots, x_{i_m}$. $pol_k(x_{i_1}, \dots, x_{i_m})$ represents a polynomial in set $P_k(x_{i_1}, \dots, x_{i_m})$. For an polynomial ϕ, $\phi^{(k)}$ denotes the homogeneous degree k part of ϕ. $\deg(\phi)$ denotes the degree of a polynomial ϕ. $J_\xi = (\frac{\partial^2 \xi}{\partial x_i \partial x_j})$ denotes the Hessian matrix of function ξ. *const* denotes a constant number.

In this paper, we consider the following time-invariant nonlinear filtering system:

$$\begin{cases} \mathrm{d}x(t) = f(x(t))\,\mathrm{d}t + g(x(t))\,\mathrm{d}w(t), & x(0) = x_0 \in \mathbb{R}^n, \\ \mathrm{d}y(t) = h(x(t))\,\mathrm{d}t + \mathrm{d}v(t), & y(0) \in \mathbb{R}^m, \end{cases} \tag{1}$$

where $x(t) = (x_1, \dots, x_n) \in \mathbb{R}^n, y(t) = (y_1, \dots, y_m) \in \mathbb{R}^m$ represents the state and observation vectors in Euclidean space. $f : \mathbb{R}^n \to \mathbb{R}^n$ denotes the drift mapping. $h : \mathbb{R}^n \to \mathbb{R}^m$ denotes the observation function. $g : \mathbb{R}^n \to \mathbb{R}^{n\times p}$ represents the diffusion coefficient. $f = (f_i), h = (h_i), g = (g_{ij})$ are all assumed smooth vector fields. $w(t) \in \mathbb{R}^p, v(t) \in \mathbb{R}^m$ are mutually independent standard Wiener process, i.e. $\mathbb{E}[dwdw^T] = I_n dt, \mathbb{E}[dvdv^T] = I_m dt$. Define coefficient matrix $C = (C_{ij}) := gg^T \in \mathbb{R}^{n\times n}$.

For a continuous filtering system, the ultimate goal is to determine the conditional expectation $\mathbb{E}[\phi(x_t) \mid \mathcal{Y}_t]$, where ϕ is a C^∞ function and $\mathcal{Y}_t := \sigma\{y_s : 0 \leq s \leq t\}$ is the filtration of observation. Under the sense of mean square error, conditional expectation $\mathbb{E}[x_t \mid \mathcal{Y}_t]$ is the optimal estimate for the state of system. Therefore, conditional density $\rho(t, x)$ given the observation history includes complete information of the filtering system.

Mathematically, unnormalised conditional density $\sigma(t, x)$ is described by the following Duncan–Mortensen–Zakai (DMZ) equation:

$$\mathrm{d}\sigma(t, x) = L_0\sigma\,\mathrm{d}t + \sigma(t, x)h_t^T \circ \mathrm{d}y_t, \tag{2}$$

where

$$L_0(\circ) := \frac{1}{2}\sum_{i,j=1}^{n} C_{ij}\frac{\partial^2(\circ)}{\partial x_i \partial x_j} - \sum_{i=1}^{n}\frac{\partial(f_i\circ)}{\partial x_i} - \frac{1}{2}h^T h(\circ). \tag{3}$$

Note that DMZ equation is formulated by the form of Stratonovich stochastic integral. We can find the singularity of matrix $C = (C_{ij})$ will influence the second-order differential operator term in L_0. In this paper, we assume diffusion coefficient g is an orthogonal matrix which will lead to $C = I_n$. It means that we consider the case of non-degenerated state noise.

Next we can reformulate forward differential operator L_0 as

$$L_0 = \frac{1}{2}\sum_{i=1}^{n}\frac{\partial^2}{\partial x_i^2} - \sum_{i=1}^{n} f_i\frac{\partial}{\partial x_i} - \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} - \frac{1}{2}\sum_{i=1}^{m} h_i^2. \tag{4}$$

And we define $L_i := h_i, 1 \le i \le m$ is the zero degree differential operator of multiplication by h_i.

Let

$$\begin{aligned} D_i &:= \frac{\partial}{\partial x_i} - f_i,\ 1 \le i \le n, \\ \eta &:= \sum_{i=1}^{n}\frac{\partial f_i}{\partial x_i} + \sum_{i=1}^{n} f_i^2 + \sum_{i=1}^{m} h_i^2, \end{aligned} \tag{5}$$

then we can obtain a more compact form of L_0,

$$L_0 = \frac{1}{2}\left(\sum_{i=1}^{n} D_i^2 - \eta\right). \tag{6}$$

Next we give some basic concepts related to Lie algebra.

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

Definition 2.2: A vector space $\mathcal{F}$ with the Lie bracket operation $\mathcal{F} \times \mathcal{F} \to \mathcal{F}$ denoted by $(x, y) \longmapsto [x, y]$ is called a Lie algebra if the following axioms are satisfied:

(1) The Lie bracket operation is bilinear.
(2) $[x, x] = 0$ for all $x \in \mathcal{F}$.
(3) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0, \quad x, y, z \in \mathcal{F}$.

Definition 2.3: Let $\mathfrak{g}$ and $\tilde{\mathfrak{g}}$ be two Lie algebras. An isomorphism $f : \mathfrak{g} \to \tilde{\mathfrak{g}}$ is a linear map and satisfies

(1) f is a bijection.
(2) f is a homomorphism of Lie algebras, i.e. $f[g_1, g_2] = [f(g_1), f(g_2)]$ for any $g_1, g_2 \in \mathfrak{g}$.

If there exists an isomorphism, we denote $\mathfrak{g}$ is isomorphic to $\tilde{\mathfrak{g}}$, i.e. $\mathfrak{g} \cong \tilde{\mathfrak{g}}$.

Remark 2.1: If two Lie algebras are isomorphic, then they have the same Lie algebra structure.

Next we introduce the concept of estimation algebra related to the filtering system.

Definition 2.4: The estimation algebra E of a filtering system (1) is defined to be the Lie algebra generated by $\{L_0, L_1, \dots, L_m\}$, i.e. $E = \langle L_0, h_1, \dots, h_m\rangle_{L.A.}$.

Remark 2.2: In the whole paper, we assume E is a finite-dimensional Lie algebra.

Definition 2.5: Let $L(E) \subset E$ be the vector space consisting of all the homogeneous degree 1 polynomials in E. Then the linear rank of estimation algebra E is defined by $r := \dim L(E)$. If $r = n$, we call E has maximal rank. Otherwise, E has non-maximal rank.

Definition 2.6: For a given function $h \in E$, we consider homogenous quadratic part $h^{(2)} = x^T Ax$. We define quadratic rank of h is $\lambda(h) : rank(A)$. Then quadratic rank of estimation algebra E is defined as the maximal rank of function in E, i.e. $\lambda(E) := \max_{h\in E} \lambda(h)$.

Especially, we should note that the structure of linear rank and quadratic rank play quite important roles in the classification of known non-maximal rank estimation algebra.

Definition 2.7: The Wong's Ω-matrix is the matrix $\Omega = (\omega_{ij})$, where

$$\omega_{ij} = \frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j}, \quad \forall 1 \le i, j \le n. \tag{7}$$

Obviously, $\omega_{ij} = -\omega_{ji}$, i.e. Ω is an antisymmetric matrix.

It is worth noting that elements of Ω matrix satisfy the following cyclical condition, which can be obtained by direct calculations.

$$\frac{\partial \omega_{ij}}{\partial x_l} + \frac{\partial \omega_{jl}}{\partial x_i} + \frac{\partial \omega_{li}}{\partial x_j} = 0, \quad \text{for } 1 \le i, j, l \le n. \tag{8}$$

Remark 2.3: The structure of Ω matrix influences the form of drift term f. If $\Omega = 0$, it corresponds to Benes filter, i.e. $f = \nabla\phi$ where $\phi \in C^\infty(\mathbb{R}^n)$. In Benes filtering system, drift vector field has a potential which is a meaningful nonlinear case such as electronic field case. If we consider more general situation Ω is a constant matrix, then drift vector field corresponds to $f(x) = Lx + l + \nabla\phi$, where $L \in \mathbb{R}^{n\times n}, l \in \mathbb{R}^n, \phi \in C^\infty(\mathbb{R}^n)$. This type of filter is called Yau filtering system which contains Kalman–Bucy filter and Benes filter as special cases. Yau filtering system plays an important role in the study of maximal and non-maximal rank estimation algebras.

Definition 2.8: Let U be the vector space of differential operators in the form

$$A = \sum_{(i_1, i_2, \dots, i_n)\in I_A} a_{i_1, i_2, \dots, i_n} D_1^{i_1} D_2^{i_2} \cdots D_n^{i_n}, \tag{9}$$

where functions $a_{i_1, i_2, \dots, i_n} \in C^\infty(\mathbb{R}^n)$ are smooth and $I_A \subset N^n$ is the finite set of A. For $i = (i_1, i_2, \dots, i_n) \in N^n$, denote $|i| := \sum_{k=1}^{n} i_k$. The order of A is defined by $ord(A) := \max_i |i|$. Let U_k denote differential operator in E with order no more than k. Especially, U_0 denotes smooth function in E.

Basic notations related to Lie bracket: Let $A, B \in E$ and V is a subspace of E. Then we define an equivalence relation $A = B$, mod V if $A - B \in V$. We define adjoint map $Ad : E \times E \to E$ by $Ad_A B = [A, B]$ and $Ad_A^k B = [A, Ad_A^{K-1} B]$. Euler operator is $E_S := \sum_{l \in S} x_l \frac{\partial}{\partial x_l}$, where S is an index subset of $\{1, 2, \ldots, n\}$.

Estimation algebra is an operator algebra. The following calculation rule is useful in exploring the structure of estimation algebra.

Lemma 2.9 (Yau, 1994): *Let E be an estimation algebra for the filtering problem* (1). *$\Omega = (\omega_{ij})$ is defined as in 2.7. Assume $X, Y, Z \in E$ and $g, h \in C^\infty(\mathbb{R}^n)$. Then*

(1) $[XY, Z] = X[Y, Z] + [X, Z]Y$;
(2) $[gD_i, h] = g\frac{\partial h}{\partial x_i}$;
(3) $[gD_i, hD_j] = gh\omega_{ji} + g\frac{\partial h}{\partial x_i}D_i - h\frac{\partial g}{\partial x_j}D_i$;
(4) $[gD_i^2, h] = 2g\frac{\partial h}{\partial x_i}D_i + g\frac{\partial^2 h}{\partial x_i^2}$;
(5) $[D_i^2, hD_j] = 2\frac{\partial h}{\partial x_i}D_iD_j + 2h\omega_{ji}D_i + \frac{\partial^2 h}{\partial x_i^2}D_j + h\frac{\partial \omega_{ji}}{\partial x_i}$;
(6) $[D_i^2, D_j^2] = 4\omega_{ji}D_jD_i + 2\frac{\partial \omega_{ji}}{\partial x_j}D_i + 2\frac{\partial \omega_{ji}}{\partial x_i}D_j + \frac{\partial^2 \omega_{ji}}{\partial x_i \partial x_j} + 2\omega_{ji}^2$;
(7) $[D_k^2, hD_iD_j] = 2\frac{\partial h}{\partial x_k}D_kD_iD_j + 2h\omega_{jk}D_iD_k + 2h\omega_{ik}D_kD_j + \frac{\partial^2 h}{\partial x_k^2}D_iD_j + 2h\frac{\partial \omega_{jk}}{\partial x_i}D_k + h\frac{\partial \omega_{jk}}{\partial x_k}D_i + h\frac{\partial \omega_{ik}}{\partial x_k}D_j + h\frac{\partial^2 \omega_{jk}}{\partial x_i \partial x_k}$;
(8) $[gD_iD_j, hD_k] = g\frac{\partial h}{\partial x_j}D_iD_k + g\frac{\partial h}{\partial x_i}D_jD_k - h\frac{\partial g}{\partial x_k}D_iD_j + gh\omega_{kj}D_i + gh\omega_{ki}D_j + g\frac{\partial^2 h}{\partial x_i \partial x_j}D_k + gh\frac{\partial \omega_{kj}}{\partial x_i}$.

Following technical results of brackets appear frequently in estimation algebra.

Lemma 2.10: (1) $[L_0, x_i] = D_i$;
(2) $[[L_0, \phi], \phi] = |\nabla \phi|^2 = \sum_{i=1}^n (\frac{\partial \phi}{\partial x_i})^2$;
(3) $[L_0, D_j] = \sum_{i=1}^n \omega_{ji}D_i + \frac{1}{2}\frac{\partial \eta}{\partial x_j} + \frac{1}{2}\sum_{i=1}^n \frac{\partial \omega_{ji}}{\partial x_i}$;
(4) $[L_0, x_j^2] = 2x_jD_j + 1$.

Some known results used in this paper are listed.

Theorem 2.11 (Ocone, 1981): *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.

Following theorem proposed by Wu and Yau illustrates the coefficients of highest order term of a differential operator in E must be polynomials.

Theorem 2.12 (Wu & Yau, 2006): *Let E be a finite-dimensional estimation algebra. If $l \geq 0$ and*

$$A = \sum_{|(i_1, i_2, \ldots, i_n)| = l+1} a_{i_1, i_2, \ldots, i_n} D_1^{i_1} D_2^{i_2} \cdots D_n^{i_n}, \quad \text{mod } U_l \in E, \tag{10}$$

then coefficients $a_{i_1, i_2, \ldots, i_n}$ are polynomials.

A trivial extension of Theorems A2 and A3 in Yau and Rasoulian (1999) about Euler operator can be written down.

Theorem 2.13: *Let Euler operator $E_S := \sum_{l \in S} x_l \frac{\partial}{\partial x_l}$, where S is an index subset of $\{1, 2, \ldots, n\}$. Set m is a positive constant. Suppose $E_S(\zeta) + m\zeta \in P_k(x)$. Then $\zeta \in P_k(x)$.*

Theorem 2.14: *Let Euler operator $E_S := \sum_{l \in S} x_l \frac{\partial}{\partial x_l}$, where S is an index subset of $\{1, 2, \ldots, n\}$. Suppose $E_S(\zeta) \in P_k(x)$. Then $\zeta \in P_k(x) + a(x_j, j \notin S)$, where a is a smooth function.*

3. Finite-dimensional estimation algebra of non-maximal linear rank

It has been proven that constant structure of Ω matrix plays an important role in maximal rank classification of estimation algebra. One key step in maximal rank classification is to determine the linear structure of Ω. As an extension, in this section, we will consider general finite-dimensional estimation algebra with non-maximal rank. Our goal in this section is to determine the linear structure of submatrix of Ω as shown in Section 3.6. The techniques used here include the structure of linear rank and quadratic rank and property of Euler operator. This result will provide a common starting point for further study on nonmaximal rank cases.

First, we make an assumption:

Assumption 3.1: *linear rank $\nu(E) = r \leq n - 1$.*

Without loss of generality, we assume $x_1, x_2, \ldots, x_r \in E$ and $x_{r+1}, \ldots, x_n \notin E$. It can be easily checked that $1 \in E$ since $[[L_0, x_1], x_1] = 1 \in E$.

Next due to the structure of linear rank, we can give a restriction for homogeneous quadratic part of a function in E. Detailed proof can be found in Appendix.

Lemma 3.1: *Let $\phi(x) \in E$ be a function in E. Then homogeneous quadratic part of ϕ must be a block diagonal form, i.e.*

$$\phi^{(2)}(x) = x^T \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} x, \tag{11}$$

where A_1 and A_2 are symmetric matrices with size $r \times r$ and $(n - r) \times (n - r)$, $x = (x_1, x_2, \ldots, x_n)^T$.

Brockett (1979) proved that if one performs a smooth nonsingular change of variable $z = F(x)$, this mapping will lead to an isomorphism of estimation algebra. Then if we consider making an orthogonal transformation and translation, the estimation algebra will extend an isomorphism.

Next we denote that quadratic rank of E is $\lambda(E) = k$. By definition, we can find a quadratic polynomial $\phi \in E$ such that

$$\phi^{(2)}(x) = x^T \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} x, \tag{12}$$

and $rank(A_1) + rank(A_2) = k$. Under an appropriate block diagonal orthogonal transformation

$$T = \begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}, \tag{13}$$

quadratic part $\phi^{(2)}$ can be diagonalised as

$$\phi^{(2)} = \begin{pmatrix} D_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & D_2 \end{pmatrix}, \tag{14}$$

where D_1, D_2 are diagonal matrices with non-zero diagonal elements. Wu and Yau (2006, Theorem 3.7) used the technique of translation of variable and Vandermonde matrix then proved that there exists quadratic function $p_0 = \sum_{i=1}^{k_1} x_i^2 + \sum_{i=n-k_2+1}^{n} x_i^2 \in E$, where $k_1 + k_2 = k, k_1 \le r$ and $k_2 \le n - r$. Note that p_0 has greatest quadratic rank k in E, i.e. $\lambda(p_0) = k$.

Remark 3.1: Note that such orthogonal transformation (13) and translation of variable $x_i \longmapsto x_i + const$ do not change the basis of $L(E)$.

For convenience, we define index set $S := \{1, \dots, k_1, n - k_2 + 1, \dots, n\}$. Here we note

(i) if $k_1 = 0$, $S = \{n - k_2 + 1, \dots, n\}$;
(ii) if $k_2 = 0$, $S = \{1, \dots, k_1\}$;
(iii) if $k_1 = k_2 = 0$, $S = \emptyset$.

Following lemma describes the structure of homogeneous part of any function in E. The proof appears in the Appendix.

Lemma 3.2: *If $p \in E$ is a quadratic function, then homogeneous quadratic part $p^{(2)}(x)$ is independent of x_j for $j \notin S$, i.e. $\frac{\partial p^{(2)}(x)}{\partial x_j} = 0$ for $j = k_1 + 1, \dots, n - k_2$.*

By using the technique of Lemma 3.2 and cyclic condition satisfied by Ω, we can further describe the structure of elements of Ω. The proof can be found in the Appendix.

Lemma 3.3: *Suppose E is a finite-dimensional estimation algebra of linear rank r and quadratic rank $k = k_1 + k_2$, where k_1, k_2 are defined in p_0. Then*

(i) $\omega_{ij}^{(2)}$ *only depends on* $x_1, \dots, x_{k_1}, x_{n-k_2+1}, \dots, x_n$ *for* $1 \le i, j \le r$.
(ii) $\omega_{ij}^{(2)}$ *only depends on* $x_{n-k_2+1}, \dots, x_n$ *for* $k_1 + 1 \le i, j \le r$.
(iii) $\frac{\partial \omega_{ij}^{(1)}}{\partial x_l} + \frac{\partial \omega_{jl}^{(1)}}{\partial x_i} + \frac{\partial \omega_{li}^{(1)}}{\partial x_j} = 0$ *for* $1 \le i, j, l \le r$.
(iv) $\frac{\partial \omega_{ij}^{(2)}}{\partial x_l} + \frac{\partial \omega_{jl}^{(2)}}{\partial x_i} + \frac{\partial \omega_{li}^{(2)}}{\partial x_j} = 0$ *for* $1 \le i, j, l \le r$.

To find more information about elements of Ω, we define following function for $1 \le i \le r$:

$$\begin{aligned} E \ni \alpha_i &:= \frac{1}{2}[[L_0, D_i], p_0] \\ &= \frac{1}{2}\left[\sum_{l=1}^{n} \omega_{il} D_l, \quad \text{mod } U_0, p_0\right] \\ &= \frac{1}{2}\sum_{l=1}^{n} [\omega_{il} D_l, p_0] \\ &= \frac{1}{2}\sum_{l=1}^{n} \omega_{il} \frac{\partial p_0}{\partial x_l} \\ &= \sum_{l \in S} x_l \omega_{il}. \end{aligned} \tag{15}$$

By Theorem 2.11 (Ocone), α_i's are polynomials of degree at most 2 in $x_1, \dots, x_n$. In the next lemma, we will find the relation between α_i and (ω_{ij}) by taking derivative in terms of x. Detailed proof can be found in the Appendix.

Lemma 3.4: *Assume $S \neq \emptyset$. For $1 \le i \le r$, $\alpha_i := \sum_{l \in S} \omega_{il} x_l \in E$ are degree at most 2 polynomials in x. Then*

(i) $E_S(\omega_{ij}) + 2\omega_{ij} = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i}$, *for* $1 \le i, j \le k_1$.
(ii) $E_S(\omega_{ij}) + \omega_{ij} = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i}$, *for* $1 \le i \le k_1, k_1 + 1 \le j \le r$.
(iii) $E_S(\omega_{ij}) + \omega_{ij} = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i}$, *for* $1 \le j \le k_1, k_1 + 1 \le i \le r$.
(iv) $E_S(\omega_{ij}) = \frac{\partial \alpha_i}{\partial x_j} - \frac{\partial \alpha_j}{\partial x_i}$, *for* $k_1 + 1 \le i, j \le r$.

Next based on Lemma 3.4, we utilise the tool of Euler operator to simplify (ω_{ij}). In the following, we use $\Omega_{1 \le i,j \le r}$ to represent a submatrix which is the intersection of the first r rows and columns of Ω. Proof appears in the Appendix.

Lemma 3.5: *Under Assumption 1, we have*

$$\Omega_{1 \le i,j \le r} = \left(\begin{array}{c|c} P_1(x) & P_1(x) \\ P_1(x) & P_1(x) + P_2(x_{k_1+1}, \dots, x_{n-k_2}) \end{array} \right), \tag{16}$$

i.e. ω_{ij} are polynomials of degree 1 in $x_1, x_2, \dots, x_n$ for $1 \le i, j \le k_1$ or $1 \le i \le k_1, k_1 + 1 \le j \le r$ or $1 \le j \le k_1, k_1 + 1 \le i \le r$. ω_{ij} are polynomials of degree 1 in $x_1, x_2, \dots, x_n$ plus polynomials of degree 2 in $x_{k_1+1}, \dots, x_{n-k_2}$ variables for $k_1 + 1 \le i, j \le r$.

In Lemma 3.5, $\Omega_{1 \le i,j \le r}$ has been partitioned to four blocks. From left to right and from top to bottom, sizes of submatrices are $k_1 \times k_1, k_1 \times (r - k_1), (r - k_1) \times k_1, (r - k_1) \times (r - k_1)$. To obtain a more meticulous structure, we improve Lemma 3.5 by using linear and quadratic rank structures. And finally it implies the linear structure of $\Omega_{1 \le i,j \le r}$. Proof can be checked in the Appendix.

Theorem 3.6: *Under Assumption 1, we have*

$$\Omega_{1 \le i,j \le r} = \left(\begin{array}{c|c} P_1(x_1, x_2, \dots, x_{k_1}) & P_1(x_1, x_2, \dots, x_{k_1}) \\ P_1(x_1, x_2, \dots, x_{k_1}) & P_1(x_{k_1+1}, \dots, x_r) \end{array} \right), \tag{17}$$

i.e. (i) *ω_{ij} are polynomials of degree 1 in $x_1, x_2, \dots, x_{k_1}$ for $1 \le i, j \le k_1$ or $1 \le i \le k_1, k_1 + 1 \le j \le r$ or $1 \le j \le k_1, k_1 + 1 \le i \le r$.*

(ii) *ω_{ij} are polynomials of degree 1 in $x_{k_1+1}, \dots, x_r$ for $k_1 + 1 \le i, j \le r$.*

Following result can be directly obtained by Theorem 3.6.

Corollary 3.7: *Under Assumption 1, submatrix $\Omega_{1 \le i,j \le r}$ has linear structure.*

4. Finite-dimensional estimation algebra of state dimension *n* and linear rank *n*−1

In this section, we discuss the situation of linear rank $r = n-1$, where n is state–space dimension. In Section 3, partial linear structure of Ω has been obtained. Under the assumption linear rank $r = n-1$, $\Omega_{1\le i,j\le n-1}$ has been proven linear structure. In this section, we will further investigate remaining part of Ω and prove the linear structure of Ω. The main method is infinite sequence technique.

By definition of linear rank, without loss of generality, we assume $x_1, \dots, x_{n-1} \in E$. It follows $[L_0, x_i] = D_i \in E$ for $1 \le i \le n-1$. By Theorem 3.6, ω_{ij}'s are affine functions in x for $1 \le i,j \le n-1$ and Ω must be of the partitioned form.

$$\Omega = \left(\begin{array}{cc|c} P_1(x_1,x_2,\dots,x_{k_1}) & P_1(x_1,x_2,\dots,x_{k_1}) & \omega_{1n} \\ P_1(x_1,x_2,\dots,x_{k_1}) & P_1(x_{k_1+1},\dots,x_{n-1}) & \omega_{2n} \\ & & \vdots \\ & & \omega_{n-1,n} \\ \hline \omega_{n1} \quad \omega_{n2} & \cdots \quad \omega_{n,n-1} & 0 \end{array}\right), \tag{18}$$

where $k_1 \le n-1$. Next we calculate $Y_i = [L_0, D_i] \in E$ for $1 \le i \le n-1$.

$$\begin{cases} Y_1 = \omega_{12}D_2 + \omega_{13}D_3 + \cdots + \omega_{1,n-1}D_{n-1} \\ \qquad + \omega_{1,n}D_n, \ \text{mod } U_0 \in E, \\ Y_2 = \omega_{21}D_1 + \omega_{23}D_3 + \cdots + \omega_{2,n-1}D_{n-1} \\ \qquad + \omega_{2,n}D_n, \ \text{mod } U_0 \in E, \\ \qquad \cdots \\ Y_{n-1} = \omega_{n-1,1}D_1 + \omega_{n-1,2}D_2 + \cdots + \omega_{n-1,n-2}D_{n-2} \\ \qquad + \omega_{n-1,n}D_n, \ \text{mod } U_0 \in E. \end{cases} \tag{19}$$

If we consider any differential operator in E,

$$A = \sum_{|(i_1,\dots,i_n)|=l+1} a_{i_1,\dots,i_n} D_1^{i_1} \cdots D_n^{i_n}, \quad \text{mod } U_l \in E. \tag{20}$$

Theorem 2.12 stated that the coefficients of the highest order term of operator A must be polynomials. Therefore $Y_{n-1} \in E$ implies that ω_{in}'s are polynomials in variables x for $1 \le i \le n-1$. And our goal is transformed to show in fact ω_{in}'s are affine functions for $1 \le i \le n-1$.

To deal with $\{Y_i\}$ uniformly, we define differential operator,

$$Y := w_1D_1 + w_2D_2 + \cdots + w_{n-1}D_{n-1} + w_nD_n, \ \text{mod } U_0 \in E, \tag{21}$$

where w_i are affine functions in $x_1, \dots, x_{n-1}$ for $1 \le i \le n-1$ and w_n is a polynomial of $x = (x_1, x_2, \dots, x_n)$. If we can prove $\deg(w_n) \le 1$, since $Y_1, Y_2, \dots, Y_{n-1}$ are all special cases of Y, then $\omega_{i,n}$ are degree 1 polynomials for $1 \le i \le n-1$. Our main theorem is as below.

Theorem 4.1 (Main theorem): *Let E be the finite-dimensional estimation algebra with state dimension n and linear rank $n-1$. Assume coefficients of operator Y in (21) satisfy that w_i are affine function in $x_1, x_2, \dots, x_{n-1}$ for $1 \le i \le n-1$ and w_n is a polynomial of $x_1, x_2, \dots, x_n$. Then w_n is a degree at most 1 polynomial of x.*

Above main theorem can directly imply the following linear structure of Ω.

Corollary 4.2 (Linear structure of Ω): *Let E be the finite-dimensional estimation algebra with state dimension n and linear rank $n-1$. Then Wong's Ω-matrix has linear structure, i.e. all the entries in the Ω-matrix are degree 1 polynomials. Furthermore, $\omega_{ij} \in P_1(x_1, \dots, x_{n-1})$ for $1 \le i,j \le n-1$.*

In the following, we will start the proof of main theorem.

First we prepare the following useful lemma about calculations of differential operators. Proof can be found in the Appendix.

Lemma 4.3: *Suppose E is finite dimensional and*

$$\begin{cases} K := const \cdot D_n^{l_1+2} + (B_1\bar{x} + const)D_1D_n^{l_1+1} \\ \qquad + (B_2\bar{x} + const)D_2D_n^{l_1+1} + \cdots \\ \qquad + (B_{n-1}\bar{x} + const)D_{n-1}D_n^{l_1+1} \\ \qquad + \text{terms with lower order in } D_n, \text{mod } U_{l_1+1} \in E, \\ Z_1 := (B_1\bar{x} + const)D_n^{l_2+1} + \text{terms with lower order} \\ \qquad \text{in } D_n, \text{mod } U_{l_2} \in E, \\ Z_2 := (B_2\bar{x} + const)D_n^{l_2+1} + \text{terms with lower order} \\ \qquad \text{in } D_n, \text{mod } U_{l_2} \in E, \\ \qquad \vdots \\ Z_{n-1} := (B_{n-1}\bar{x} + const)D_n^{l_2+1} + \text{terms with lower order} \\ \qquad \text{in } D_n, \text{mod } U_{l_2} \in E, \end{cases} \tag{22}$$

where const means a constant number, $\bar{x} = (x_1, x_2, \dots, x_{n-1})^T \in \mathbb{R}^{(n-1)}$. $l_1, l_2 \ge 0$ are nonnegative integers. $B_i \in \mathbb{R}^{1\times(n-1)}$ are constant row vectors for $1 \le i \le n-1$. Define block matrix $B = (B_{ij})$ as below,

$$B := \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_{n-1} \end{pmatrix}. \tag{23}$$

If B is a real symmetric matrix, then $B = 0$.

Remark 4.1: Note that Lemma 4.3 holds only under the assumption of finite dimensionality of E and independent of linear rank condition. Then it can be applied in any nonmaximal rank estimation algebra. This lemma is quite an important tool in the calculation of estimation algebra.

It is noted that $x_1, \dots, x_{n-1}$ play the same role in Lemma 4.3 and we do not use the special information of x_n. Therefore, by symmetry, if we simply replace index n by any $1 \le \alpha \le n-1$, $\bar{x}$ by $(x_1, \dots, x_{\alpha-1}, x_{\alpha+1}, \dots, x_n)$, same results hold as follows.

Lemma 4.4: *Suppose E is finite-dimensional estimation algebra and $1 \le \alpha \le n$,*

$$K := const \cdot D_\alpha^{l_1+2} + (B_1\bar{x} + const)D_1D_\alpha^{l_1+1} + \cdots$$

$$+ (B_{\alpha-1}\bar{x} + const)D_{\alpha-1}D_\alpha^{l_1+1} + (B_{\alpha+1}\bar{x} + const)D_{\alpha+1}D_\alpha^{l_1+1} + (B_n\bar{x} + const)D_nD_\alpha^{l_1+1} + \text{terms with lower order in } D_\alpha, \text{ mod } U_{l_1+1} \in E,$$

$$Z_1 := (B_1\bar{x} + const)D_\alpha^{l_2+1} + \text{terms with lower order in } D_\alpha, \text{ mod } U_{l_2} \in E,$$

$$\cdots$$

$$Z_{\alpha-1} := (B_{\alpha-1}\bar{x} + const)D_\alpha^{l_2+1} + \text{terms with lower order in } D_\alpha, \text{ mod } U_{l_2} \in E,$$

$$\cdots$$

$$Z_{\alpha+1} := (B_{\alpha+1}\bar{x} + const)D_\alpha^{l_2+1} + \text{terms with lower order in } D_\alpha, \text{ mod } U_{l_2} \in E,$$

$$\cdots$$

$$Z_n := (B_n\bar{x} + const)D_\alpha^{l_2+1} + \text{terms with lower order in } D_\alpha, \text{ mod } U_{l_2} \in E, \tag{24}$$

where const means constant number, $\bar{x} = (x_1, \ldots, x_{\alpha-1}, x_{\alpha+1} \cdots, x_n)^T \in \mathbb{R}^{n-1}$. $l_1, l_2 \geq 0$ *are nonnegative integers.* $B_i \in \mathbb{R}^{1\times(n-1)}$ *are constant row vectors. Define block matrix B as below,*

$$B := \begin{pmatrix} B_1 \\ \cdots \\ B_{\alpha-1} \\ \cdots \\ B_{\alpha+1} \\ \cdots \\ B_{n-1} \end{pmatrix}. \tag{25}$$

If B is a symmetric matrix, then $B = 0$.

To explore the algebraic structure of w_n, we first expand w_n to a polynomial in terms of x_n with polynomial coefficients of $x_1, \ldots, x_{n-1}$. Following Lemma 4.5 shows that the coefficient of the highest degree term of x_n must be degree 1 polynomial of $x_1, \ldots, x_{n-1}$. Proof can be found in the Appendix.

Lemma 4.5: *Since* w_n *is a polynomial of* $x_1, x_2, \ldots, x_n$*, we may assume that*

$$w_n = a_l x_n^l + \cdots + a_1 x_n + a_0, \tag{26}$$

where a_i *are polynomials of* $x_1, x_2, \ldots, x_{n-1}$ *for* $0 \leq i \leq l$*. Assume* $a_l \neq 0$*. If* $l \geq 1$*, then* $a_l \in P_1(x_1, x_2, \ldots, x_{n-1})$.

Next we consider the simplified situation of our main Theorem 4.1. By using the method developed in Lemma 4.3, we will prove main theorem holds when $w_i = 0$ for $1 \leq i \leq n-1$ in the following Theorem 4.6. The proof appears in the Appendix and is long and full of techniques.

Theorem 4.6: *Suppose E is a finite-dimensional estimation algebra of dimension n and linear rank* $n-1$*. If the following differential operator is contained in estimation algebra,*

$$M_0 = \alpha(x_1, x_2, \ldots, x_n)D_n, \text{ mod } U_0 \in E, \tag{27}$$

where α *is a polynomial of* $x_1, x_2, \ldots, x_n$*. Then* α *is an affine function in variables* $x_1, x_2, \ldots, x_n$.

Different from expanding w_n in terms of x_n in Lemma 4.5, following Lemma 4.7 shows that the coefficient of the highest degree is also an affine function if we consider expanding w_n in terms of x_j for $1 \leq j \leq n-1$. Lemma 4.7 is a direct result from Theorem 4.6. Proof appears in the Appendix.

Lemma 4.7: *Suppose* $1 \leq j \leq n-1$ *and*

$$w_n = \alpha_{k_j} x_j^{k_j} + \cdots + \alpha_1 x_j + \alpha_0, \quad k_j \geq 1,\ \alpha_{k_j} \neq 0, \tag{28}$$

where α_i *are polynomials of* $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ *for* $0 \leq i \leq k_j$*. Then* $\alpha_{k_j} \in P_1(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)$.

In the following, we proceed to extend Lemma 4.5. We will prove the degree of w_n with respect to x_n is no more than 1 in the following two lemmas.

Lemma 4.8: *Suppose that*

$$w_n = a_l x_n^l + \cdots + a_1 x_n + a_0, \quad l \geq 1,\ a_l \neq 0, \tag{29}$$

where a_i *are polynomials of* $x_1, x_2, \ldots, x_{n-1}$ *for* $0 \leq i \leq l$*. If there exists* $1 \leq j \leq n-1$*,* w_n *contains* x_j *component, i.e.*

$$w_n = \alpha_{k_j} x_j^{k_j} + \cdots + \alpha_1 x_j + \alpha_0, \quad k_j \geq 1,\ \alpha_{k_j} \neq 0, \tag{30}$$

then $l = 1$.

Lemma 4.9: *Suppose that*

$$w_n = a_l x_n^l + \cdots + a_1 x_n + a_0, \quad l \geq 1,\ a_l \neq 0, \tag{31}$$

where a_i *are polynomials of* $x_1, x_2, \ldots, x_{n-1}$ *for* $0 \leq i \leq l$*. Then* $l = 1$.

Based on Lemma 4.9, we can assume

$$w_n = a_1(x_1, \ldots, x_{n-1})x_n + a_0(x_1, \ldots, x_{n-1}),$$

where $a_1 \in P_1(x_1, \ldots, x_{n-1})$. Remaining question is to reduce the degree of a_0. In the next lemma, we can prove a_0 is a polynomial of degree at most 2. Proof appears in the Appendix.

Lemma 4.10: *Assume* $w_n = a_1 x_n + a_0$*, where* $a_1 \in P_1(x_1, \ldots, x_{n-1})$*. Then* $a_0 \in P_2(x_1, \ldots, x_{n-1})$.

Up to now, we have proved that w_n is a polynomial of degree at most 2. Next we will further reduce degree of w_n to at most 1 and this will finish the proof of main theorem. Proof can be found in the Appendix.

Theorem 4.11 (Main theorem): *w_n is degree at most 1 polynomial of x.*

Corollary 4.12 (Linear structure of Ω): *Let E be the finite-dimensional estimation algebra with state dimension n and linear rank* $n-1$*. Then Wong's* Ω*-matrix has a linear structure; i.e. all the entries in the* Ω*-matrix are degree 1 polynomials.*

5. Structure of finite-dimensional filtering systems

In this section, we will proceed based on the results of Section 4. We give the structure of drift function in the following theorem. Proof appears in the Appendix.

Theorem 5.1: *$\frac{\partial f_j}{\partial x_i} - \frac{\partial f_i}{\partial x_j} = c_{ij} + D_{ij}^T x$, where $D_{ij} \in \mathbb{R}^n$, $x = (x_1, x_2, \dots, x_n)^T$ for all $1 \le i, j \le n$ if and only if*

$$(f_1, \dots, f_n) = (l_1, \dots, l_n) + \left(\frac{\partial \psi}{\partial x_1}, \dots, \frac{\partial \psi}{\partial x_n} \right), \qquad (32)$$

where $l_1, \dots, l_n$ are degree at most 2 polynomials and ψ is a C^∞ function.

6. Conclusion

This paper mainly focuses on the linear structure of Ω on nonmaximal rank estimation algebra. Section 3 gives a linear structure of submatrix of Ω. This would be useful and a starting point for the hereafter study of nonmaximal rank case. In Section 4, based on the general result of Section 3, the linear structure of Ω is obtained for the case with rank $n-1$. It is a critical step for overcoming the classification of nonmaximal rank case. And it provides a base for exploring the Mitter conjecture of nonmaximal rank case. Proving the Mitter conjecture for estimation algebra with rank $n-1$ is our future work. Finally, the structure of drift function in the case of linear rank $n-1$ is determined. This would provide a guidance for finding efficient numerical algorithms for finite-dimensional filter with linear rank $n-1$.

Acknowledgments

Professor Stephen Shing-Toung Yau is grateful to National Center for Theoretical Sciences (NCTS) for providing excellent research environment while part of this research was done.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work is supported by National Natural Science Foundation of China (NSFC) grant (11961141005), Tsinghua University start-up fund, and Tsinghua University Education Foundation fund (042202008).

References

Brigo, D., Hanzon, B., & LeGland, F. (1998). A differential geometric approach to nonlinear filtering: The projection filter. *IEEE Transactions on Automatic Control*, *43*(2), 247–252. https://doi.org/10.1109/9.661075

Brockett, R. W. (1979). Classification and equivalence in estimation theory. In *1979 18th IEEE Conference on Decision and Control including the Symposium on Adaptive Processes* (Vol. 2, pp. 172–175). IEEE.

Brockett, R. W. (1981). *Nonlinear systems and nonlinear estimation theory*. Springer.

Brockett, R. W., & Clark, J. M. C. (1980). The geometry of the conditional density function. In *Analysis and optimization of stochastic systems*. Academic Press.

Chen, X., Luo, X., & Yau, S. S. T. (2017). Direct method for time-varying nonlinear filtering problems. *IEEE Transactions on Aerospace and Electronic Systems*, *53*(2), 630–639. https://doi.org/10.1109/TAES.2017.2651650

Chen, J., & Yau, S. S. T. (1996). Finite-dimensional filters with nonlinear drift. VI: Linear structure of Ω. *Mathematics of Control, Signals and Systems*, *9*(4), 370–385. https://doi.org/10.1007/BF01211857

Chiou, W. L., & Yau, S. S. T. (1994). Finite-dimensional filters with nonlinear drift. II: Brockett's problem on classification of finite-dimensional estimation algebras. *SIAM Journal on Control and Optimization*, *32*(1), 297–310. https://doi.org/10.1137/S0363012991201660

Epstein, E. S. (1969). Stochastic dynamic prediction. *Tellus*, *21*(6), 739–759.

Jiao, X. P., & Yau, S. S. T (2020). New classes of finite dimensional filters with nonmaximal rank estimation algebra on state dimension n and linear rank n-2. *SIAM Journal on Control and Optimization*, *58*(6), 3413–3427. https://doi.org/10.1137/20M1350467

Kalman, R. E., & Bucy, R. S. (1961). New results in linear filtering and prediction theory.

Kim, J. W., Mehta, P. G., & Meyn, S. P. (2019). What is the Lagrangian for nonlinear filtering?. In *2019 IEEE 58th Conference on Decision and Control (CDC)* (pp. 1607–1614). IEEE.

Luo, X., & Yau, S. S. T. (2013). Hermite spectral method to 1-D forward Kolmogorov equation and its application to nonlinear filtering problems. *IEEE Transactions on Automatic Control*, *58*(10), 2495–2507. https://doi.org/10.1109/TAC.2013.2259975

Mitter, S. K. (1980). On the analogy between mathematical problems of non-linear filtering and quantum physics. Massachusetts Inst of Tech Cambridge Lab for Information and Decision Systems.

Ocone, D. (1981). *Finite dimensional estimation algebras in nonlinear filtering*. Springer.

Rozovsky, B. L. (1972). Stochastic partial differential equations arising in nonlinear filtering problems. *Uspekhi Matematicheskikh Nauk*, *27*(3), 213–214.

Shi, J., Chen, X., Dong, W., & Yau, S. S. T (2017). New classes of finite dimensional filters with non-maximal rank. *IEEE Control Systems Letters*, *1*(2), 233–237. https://doi.org/10.1109/LCSYS.2017.2713604

Shi, J., Yang, Z., & Yau, S. S. T. (2018). Direct method for Yau filtering system with nonlinear observations. *International Journal of Control*, *91*(3), 678–687. https://doi.org/10.1080/00207179.2017.1288300

Shi, J., & Yau, S. S. T. (2017). Finite dimensional estimation algebras with state dimension 3 and rank 2. I: Linear structure of Wong matrix. *SIAM Journal on Control and Optimization*, *55*(6), 4227–4246. https://doi.org/10.1137/16M1065471

Wu, X., & Yau, S. S. T. (2006). Classification of estimation algebras with state dimension 2. *SIAM Journal on Control and Optimization*, *45*(3), 1039–1073. https://doi.org/10.1137/S036301290443991X

Yang, T., Mehta, P. G., & Meyn, S. P. (2013). Feedback particle filter. *IEEE Transactions on Automatic Control*, *58*(10), 2465–2480. https://doi.org/10.1109/TAC.2013.2258825

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

Yau, S. S. T. (2003). Complete classification of finite-dimensional estimation algebras of maximal rank. *International Journal of Control*, *76*(7), 657–677. https://doi.org/10.1080/0020717031000098435

Yau, S. S. T., & Hu, G. Q. (2005). Classification of finite-dimensional estimation algebras of maximal rank with arbitrary state–space dimension and Mitter conjecture. *International Journal of Control*, *78*(10), 689–705. https://doi.org/10.1080/00207170500107596

Yau, S. S. T., & Rasoulian, A. (1999). Classification of four-dimensional estimation algebras. *IEEE Transactions on Automatic Control*, *44*(12), 2312–2318. https://doi.org/10.1109/9.811217

Yau, S. T., & Yau, S. S. T. (1996). Explicit solution of a Kolmogorov equation. *Applied Mathematics and Optimization*, *34*(3), 231–266. https://doi.org/10.1007/BF01182625

Yau, S. T., & Yau, S. S. T. (2004). Nonlinear filtering and time varying Schrödinger equation. *IEEE Transactions on Aerospace and Electronic Systems*, *40*(1), 284–292. https://doi.org/10.1109/TAES.2004.1292160

Yau, S. T., & Yau, S. S. T. (2008). Real time solution of the nonlinear filtering problem without memory II. *SIAM Journal on Control and Optimization*, *47*(1), 163–195. https://doi.org/10.1137/050648353

Zakai, M. (1969). On the optimal filtering of diffusion processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *11*(3), 230–243. https://doi.org/10.1007/BF00536382

Zhang, Q., Taghvaei, A., & Chen, Y. (2021). An optimal control approach to particle filtering. arXiv preprint arXiv:2110.03199.