



Distributed order estimation for continuous-time stochastic systems

Xinghua Zhu^{1,2} · Zhixin Liu^{1,2} · Xiaoming Hu³

Received: 10 September 2023 / Revised: 2 October 2023 / Accepted: 5 October 2023
© The Author(s) 2024

Abstract

In this paper, we investigate the distributed estimation problem of continuous-time stochastic dynamic systems over sensor networks when both the system order and parameters are unknown. We propose a local information criterion (LIC) based on the L_0 penalty term. By minimizing LIC at the diffusion time instant and utilizing the continuous-time diffusion least squares algorithm, we obtain a distributed estimation algorithm to simultaneously estimate the unknown order and the parameters of the system. By dealing with the effect of the system noises and the coupling relationship between estimation of system orders and parameters, we establish the almost sure convergence results of the proposed distributed estimation algorithm. Furthermore, we give a simulation example to verify the effectiveness of the distributed algorithm in estimating the system order and parameters.

Keywords Stochastic differential equations · Sensor networks · Distributed order estimation · Cooperative excitation condition · Convergence

1 Introduction

With the rapid development of computer and communication technology, sensor networks are widely applied in engineering systems due to high flexibility, robustness, and ease of placement. One of the important issues is how to design distributed algorithms to estimate unknown parameters in dynamical systems by efficiently using the data from sensor networks. Compared with the centralized algorithms,

each sensor in distributed algorithms can exchange information with its neighbors via network structure, and complete the estimation task cooperatively. It is clear that distributed algorithms can greatly reduce the communication and computation burden, and have better robustness to node attacks. Distributed estimation algorithms have a wide range of practical applications such as target tracking, health monitoring, environmental monitoring and pollution prevention [1–4]. Meanwhile, the theoretical investigation of distributed algorithms has attracted much attention of researchers in various fields, e.g., control theory, system identification, signal processing. (cf., [5–8]).

For discrete-time dynamical systems, some distributed algorithms were proposed and convergence results of the algorithms were also established for time-invariant parameters (cf., [9, 10]), time-varying parameters (cf., [11, 12]), order estimation (cf., [13]), sparse signals (cf., [8]). However, in many practical scenarios, continuous time signals such as electrical signals, speech signals, attitude control, seismic waves, electromagnetic waves, etc. are widely used, and the dynamical systems are often modeled by (stochastic) differential equations based on specific laws of physics [14–20]. In recent years, some theoretical results for distributed estimation problems of continuous-time systems are given. For example, Chen and Wen et al. in [21] studied the consensus-type distributed identification algorithm

This work was supported by the National Key R&D Program of China (No. 2018YFA0703800), the Natural Science Foundation of China (No. T2293770), the Strategic Priority Research Program of Chinese Academy of Sciences (No. XDA27000000) and the National Science Foundation of Shandong Province (No. ZR2020ZD26).

✉ Xiaoming Hu
hu@kth.se

Xinghua Zhu
zxh@amss.ac.cn

Zhixin Liu
lzx@amss.ac.cn

¹ Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Beijing 100190, China

² School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100190, China

³ Department of Mathematics, KTH Royal Institute of Technology, Stockholm 10044, Sweden

for continuous-time systems where regressors are uniformly bounded and satisfy the cooperative persistent excitation (PE) condition. Nascimento and Sayed in [22] investigated the exponential stability of diffusion-type LMS algorithm with PE regressors. Papusha et al. in [23] investigated the asymptotic parameter convergence of consensus-type distributed gradient algorithm under PE condition. We see that in most of existing results, the regressors are required to be deterministic and satisfy the PE condition, which is hard to be satisfied for the stochastic feedback systems.

The linear regressions are often used to model the engineering systems. Generally speaking, the regression model with higher order and more parameters can fit the data better. However, the model with high order may lead to parameter redundancy as well as computational pressure. How to choose the proper order of a regression model is an important topic in the fields of statistics, machine learning and system identification. For discrete-time linear regression models, some algorithms are constructed by optimizing some information criteria, such as AIC (Akaike Information Criterion) [24], BIC (Bayesian Information Criterion) [25], and CIC (the first letter ‘‘C’’ refers to the information criterion designed for feedback control systems) [26] and their variants, and some theoretical results for the algorithms are obtained. For example, Hannan and Kavalieris in [27] used the AIC criterion to design an algorithm to estimate the system order and unknown parameters, and gave the convergence analysis of the algorithm under stable input signals. Ninness in [28] established the Cramér-Rao lower bound for the order estimation problem of the stable stochastic observation model. Chen and Guo in [29] proposed a least-squares based order estimation algorithm by introducing an information criterion, and the consistent results for the system order and parameters were given under non-persistent excitation condition when the upper bounds of orders are known. Later, Guo and Huang in [30] established the convergence result of the order estimation problem with unknown upper order bounds. Chen and Zhao in [31] proposed a recursive order estimation algorithm with independent and identical distributed input signals. In recent years, some studies have been carried out based on genetic algorithms and neural networks [32, 33]. However, only simulation experiments were conducted to validate the algorithms, no corresponding rigorous theoretical analysis was given. For the order estimation problem of continuous-time models, Victor et al. in [34] designed a two-stage estimation algorithm to estimate the unknown parameters and orders of a noise-free dynamic system and verified the effectiveness of the algorithm by simulation experiments. Belkhatir and Laleg-Kirati in [35] studied local convergence for order estimation and parameter estimation of noisy systems under bounded input–output assumption. Subsequently, Stark et al. in [36] analyzed local convergence results for system order estimation with obser-

vation noise. So far, little attention was paid to the design and global convergence of the distributed order estimation problem of continuous-time stochastic systems.

In this paper, we study the distributed estimation problems of continuous-time dynamical systems described by stochastic differential equation with unknown order and parameters over sensor networks. Motivated by the discrete-time distributed order estimation problem (cf., [13]), we introduce a local information criterion based on the L_0 penalty term, and propose a two-step distributed algorithm to simultaneously estimate the unknown orders and parameters of continuous-time stochastic regression models. To be specific, the order estimates are obtained by minimizing the local information criterion at discrete diffusion points. Based on this, the unknown parameters are estimated by using diffusion least square algorithm. Under the cooperative excitation conditions of the regression signals, almost sure global convergence are established for the distributed order estimation as well as the distributed parameter estimation, where some mathematical tools such as stochastic Lyapunov function method, Itô’s theorem, and continuous time martingale estimation theorem are used to deal with the noise effect and the coupling relationship between the estimate of the system order and unknown parameters. Finally, simulation examples are given to verify the effectiveness of the proposed distributed algorithm. Compared with most existing results on distributed adaptive estimation algorithms, the convergence results are established without relying on the commonly used independency, stationarity or ergodicity assumptions on regression signals, which makes it possible to apply our results to the feedback control systems.

The rest of this paper is organized as follows. In Sect. 2, we give the problem formulation. The main results of the proposed algorithm are provided in Sect. 3. A simulation example is given in Sect. 4, and the concluding remarks are made in Sect. 5.

2 Problem formulation

2.1 Some notations and preliminaries

In this section, we will introduce some notation and preliminary results on graph theory and probability theory.

The communication between sensors are modeled as an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, \dots, N\}$ is composed of all sensors and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the set of edges. The weighted adjacency matrix $\mathcal{A} = (a_{ij})_{n \times n}$ is used to describe the interaction weights between sensors, where the weight $a_{ij} > 0$ if and only if $(i, j) \in \mathcal{E}$. For simplicity of analysis, we assume that the matrix \mathcal{A} is symmetric and doubly stochastic, i.e., $a_{ij} = a_{ji}$ and $\sum_{i=1}^N a_{ij} = \sum_{i=1}^N a_{ji} = 1$ for all i and j . A path of length s in the graph \mathcal{G} is defined as a

sequence of labels $\{i_1, \dots, i_s\}$ satisfying $(i_j, i_{j+1}) \in \mathcal{E}$ for all $1 \leq j \leq s - 1$. The diameter of graph \mathcal{G} , denoted as $D_{\mathcal{G}}$, is defined as the maximum length of the path between any two sensors. We define $\mathcal{N}_i = \{j \in \mathcal{V} | (j, i) \in \mathcal{E}\}$ as the neighbor set of the graph \mathcal{G} . See [37] for more information about graph theory.

Let (Ω, \mathcal{F}, P) be a probability space, and $\{\mathcal{F}_t, t \geq 0\}$ be a nondecreasing family of sub- σ -algebras of \mathcal{F} . The process $\{X_t, \mathcal{F}_t; 0 \leq t < \infty\}$ is said to be a martingale if we have $E(X_t | \mathcal{F}_s) = X_s$ almost surely (a.s.) for $0 \leq s < t < \infty$, where $E[\cdot | \cdot]$ is the conditional mathematical expectation operator. The Wiener process $\{w(t), \mathcal{F}_t\}$ is an independent incremental process with $E[w(t) | \mathcal{F}_s] = 0$ and $E[w^2(t)] < \infty, t > s \geq 0$, where $E[\cdot]$ is the mathematical expectation operator. For the continuous-time martingale, the following martingale estimation theorem is often used to deal with the continuous-time stochastic noise.

Lemma 1 [38] *Let (M_t, \mathcal{F}_t) be a measurable process satisfying $\int_0^t \|M_s\|^2 ds < \infty$, a.s. $\forall t > 0$. If $\{w_t, \mathcal{F}_t\}$ is a Wiener process, then as $t \rightarrow \infty$, $\int_0^t M_s dw_s = O(\sqrt{S(t) \log \log(S(t) + e)})$ a.s., where $S(t)$ is defined by $S(t) = \int_0^t \|M_s\|^2 ds$.*

The Itô formula plays a key role in dealing with continuous-time stochastic processes, which is described as follows:

Lemma 2 [39] *Assume that the stochastic process $\xi(t)$ obeys the equation $d\xi(t) = a(t, \xi(t))dt + B(t, \xi(t))dw(t)$, and $\{a(t, \xi(t)), \mathcal{F}_t\}$ is an l -dimensional adaptive process and $\{B(t, \xi(t)), \mathcal{F}_t\}$ is an $l \times m$ -dimensional adaptive matrix process satisfying $\|a(\cdot, \cdot)\|^{\frac{1}{2}} \in \mathcal{P}_t$ and $\|B(\cdot, \cdot)\| \in \mathcal{P}_t$. If the functions $f'_t(t, x)$, $f'_x(t, x)$ and $f''_{xx}(t, x)$ are continuous, then*

$$df(t, \xi(t)) = f'_t(t, \xi(t))dt + f'^T_{\xi(t)}(t, \xi(t))d\xi(t) + \frac{1}{2} \text{tr}\{f''_{\xi(t)\xi(t)}(t, \xi(t))B(t, \xi(t))(B(t, \xi(t)))^T\}dt.$$

2.2 Problem formulation

Consider a network consisting of N sensors. The dynamical system of each sensor $i \in \{1, \dots, N\}$ obeys the following continuous-time stochastic differential equation,

$$\begin{cases} y_i(t) = a_1 S y_i(t) + \dots + a_{\bar{p}} S^{\bar{p}} y_i(t) + b_1 S u_i(t) \\ \quad + \dots + b_{\bar{q}} S^{\bar{q}} u_i(t) + v_i(t), \\ u_i(t) = 0, y_i(t) = 0, t \leq 0, \end{cases} \quad (1)$$

where S is the Itô integral operator (i.e., $S y_i(t) = \int_0^t y_i(s) ds$), $y_i(t)$ and $u_i(t)$ are the scalar output and input of the sensor i at time t , $v_i(t)$ is the system noise and modeled as a standard Wiener process, both the order (\bar{p}, \bar{q}) and parameters $a_1, \dots, a_{\bar{p}}, b_1, \dots, b_{\bar{q}}$ ($a_{\bar{p}} \neq 0, b_{\bar{q}} \neq 0$) are unknown.

For the convenience of analysis, we assume that the system order has known the upper bounds, i.e., $(\bar{p}, \bar{q}) \in I, I \triangleq \{(p, q) \mid 0 \leq p \leq U_p, 0 \leq q \leq U_q\}$, where U_p and U_q are the upper bounds of the order. Denote the vector of regressor and unknown parameters in (1) as follows:

$$\begin{aligned} \phi_i(t; p, q) &= [y_i(t), \dots, S^{p-1} y_i(t), u_i(t), \dots, S^{q-1} u_i(t)]^T, \\ \theta(p, q) &= [a_1, \dots, a_p, b_1, \dots, b_q]^T, \\ a_j &= 0, b_\tau = 0, \bar{p} < j \leq p, \bar{q} < \tau \leq q. \end{aligned}$$

Then the system (1) can be rewritten into the following regression model,

$$\begin{aligned} y_i(t) &= S\phi_i^T(t; p, q)\theta(p, q) + v_i(t) \quad (p \geq \bar{p}, q \geq \bar{q}) \\ &= S\phi_i^T(t; \bar{p}, \bar{q})\theta(\bar{p}, \bar{q}) + v_i(t). \end{aligned} \quad (2)$$

In this paper, we put forward a two-step distributed algorithm to alternately estimate the system order (\bar{p}, \bar{q}) and the parameter $\theta(\bar{p}, \bar{q})$ in (2). First, a local information criterion is proposed, by minimizing which the estimate of the system order can be obtained; Then, the diffusion least squares algorithm will be adopted to estimate unknown parameters based on the order estimate due to fast convergence rate of least squares algorithm.

In the following, we first introduce the diffusion least squares algorithm. In this algorithm, the update of unknown parameters only occurs at discrete diffusion time instants which are denoted as $0 = t_0, t_1, t_2, \dots, t_k, \dots$ satisfying $t_k \rightarrow \infty$ as $k \rightarrow \infty$. For $t \in [t_n, t_{n+1})$, we introduce the following local objective function with fixed order (p, q) ,

$$\begin{aligned} J_{n,i}(t; p, q, \theta(p, q)) &= J_{n,i}(t_n; p, q, \theta(p, q)) \\ &\quad + \int_{t_n}^t \left\{ \frac{dy_i(s)}{ds} - \theta^T(p, q)\phi_i(s; p, q) \right\}^2 ds, \end{aligned} \quad (3)$$

$$i = 1, 2, \dots, N,$$

and

$$\begin{aligned} J_{n+1,i}(t_{n+1}; p, q, \theta(p, q)) &= \sum_{j \in \mathcal{N}_i} a_{ij} J_{n,j}(t_{n+1}; p, q, \theta(p, q)), \quad i = 1, 2, \dots, N, \\ J_{0,i}(t_0; p, q, \theta(p, q)) &= 0, \quad i = 1, 2, \dots, N, \end{aligned}$$

where $0 < a_{ij} \leq 1$ is the i th row and j th column of the weighted adjacency matrix \mathcal{A} of the sensor network, and $y_i(s), \phi_i(s; p, q)$ are considered as the outcomes, thus, the integral is Riemann integral. By the above recursive form,

we have for $t \in [t_n, t_{n+1})$

$$\begin{aligned}
 & J_{n,i}(t; p, q, \theta(p, q)) \\
 &= \sum_{j=1}^N \sum_{k=0}^{n-1} a_{ij}^{(n-k+1)} \int_{t_k}^{t_{k+1}} \left\{ \frac{dy_j(s)}{ds} \right. \\
 &\quad \left. - \theta^T(p, q) \phi_j(s; p, q) \right\}^2 ds \\
 &\quad + \int_{t_{n+1}}^t \left\{ \frac{dy_i(s)}{ds} - \theta^T(p, q) \phi_i(s; p, q) \right\}^2 ds,
 \end{aligned} \tag{4}$$

and for $t = t_{n+1}$

$$\begin{aligned}
 & J_{n+1,i}(t_{n+1}; p, q, \theta(p, q)) \\
 &= \sum_{j=1}^N \sum_{k=0}^n a_{ij}^{(n-k+1)} \int_{t_k}^{t_{k+1}} \left\{ \frac{dy_j(s)}{ds} \right. \\
 &\quad \left. - \theta^T(p, q) \phi_j(s; p, q) \right\}^2 ds.
 \end{aligned} \tag{5}$$

By minimizing the objective functions (4) and (5), we can derive the following diffusion least squares algorithm. The details can be found in [40].

Algorithm 1 Diffusion least squares algorithm

For $i \in \{1, \dots, n\}$ and $t > 0$, given any initial estimates $\theta_{0,i}(p, q; 0)$, and any initial positive definite matrices $P_{0,i}(p, q; 0)$.

Step 1 (Adaptation): For $t \in [t_n, t_{n+1})$, $\theta_{n,i}(t; p, q)$ and $P_{n,i}(t; p, q)$ are generated according to the following adaptation process

$$\begin{aligned}
 d\theta_{n,i}(t; p, q) &= P_{n,i}(t; p, q) \phi_{n,i}(t; p, q) \{dy_i(t) \\
 &\quad - \phi_{n,i}^T(t; p, q) \theta_{n,i}(t; p, q) dt\},
 \end{aligned} \tag{6}$$

$$\begin{aligned}
 dP_{n,i}(t; p, q) &= -P_{n,i}(t; p, q) \phi_{n,i}(t; p, q) \\
 &\quad \cdot \phi_{n,i}^T(t; p, q) P_{n,i}(t; p, q) dt.
 \end{aligned} \tag{7}$$

Step 2 (Diffusion): For $t = t_{n+1}$, $P_{n+1,i}^{-1}(t_{n+1}; p, q)$ and $\theta_{n+1,i}(t_{n+1}; p, q)$ are updated by using the information of neighbors $\{P_{n,j}^{-1}(t_{n+1}; p, q), \theta_{k,j}(t_{n+1}; p, q)\}_{j \in \mathcal{N}_i}$,

$$P_{n+1,i}^{-1}(t_{n+1}; p, q) = \sum_{j \in \mathcal{N}_i} a_{ij} P_{n,j}^{-1}(t_{n+1}; p, q), \tag{8}$$

$$\begin{aligned}
 \theta_{n+1,i}(t_{n+1}; p, q) &= P_{n+1,i}(t_{n+1}; p, q) \\
 &\quad \cdot \sum_{j \in \mathcal{N}_i} a_{ij} P_{n,j}^{-1}(t_{n+1}; p, q) \theta_{n,j}(t_{n+1}; p, q).
 \end{aligned} \tag{9}$$

Output: $\{\theta_{n,i}(t; p, q), t \in [t_n, t_{n+1}), n \in \mathbb{N}\}$.

For each sensor $i \in \{1, \dots, N\}$ and any $t \geq 0$, we introduce the following local information criterion to estimate the order \bar{p}, \bar{q} of the dynamic system (2),

$$\begin{aligned}
 & F_{n,i}(t; p, q) \\
 &= J_{n,i}(t; p, q, \theta_{n,i}(t; p, q)) \\
 &\quad + (p + q) \xi(t), \quad t \in [t_n, t_{n+1}), n \in \mathbb{N},
 \end{aligned} \tag{10}$$

where the definition of $J_{n,i}(t; p, q, \theta_{n,i}(t; p, q))$ is given in (3), $\theta_{n,i}(t; p, q)$ is the estimate of θ given in Algorithm 1. In (10), the first term is the accumulative error between the system output and the prediction, while the second term $(p + q)\xi(t)$ is the penalty term, and can be regarded as the L_0 regularization of the unknown parameter vector. The function $\xi(t)$ reflects the balance between the accumulative prediction error and the penalty term, and we will provide a principle to select $\xi(t)$ in Sect. 3. The distributed order estimation algorithm based on the information criterion (10) is given in the following Algorithm 2.

Algorithm 2 Distributed order estimation algorithm

For each sensor $i \in \{1, \dots, n\}$, $k > 0$, given any initial estimates $\theta_{0,i}(0, p, q)$ and any initial positive definite matrix $P_{0,i}(0, p, q)$.

Step 1: For $t \in [t_k, t_{k+1})$ and for all $(p, q) \in I$, Algorithm 1 is conducted to obtain the parameter estimates $\{\theta_{k+1,i}(t_{k+1}; p, q)\}_{(p,q) \in I}$.

Step 2 (Order estimation): For $t = t_{k+1}$, based on the parameter estimates $\{\theta_{k+1,i}(t_{k+1}; p, q)\}_{(p,q) \in I}$, the estimate $(p_i(t_{k+1}), q_i(t_{k+1}))$ of the system order (\bar{p}, \bar{q}) can be obtained by minimizing the local information criterion (10), i.e.,

$$(p_i(t_{k+1}), q_i(t_{k+1})) = \arg \min_{(p,q) \in I} F_{k+1,i}(t_{k+1}; p, q). \tag{11}$$

Step 3 (Parameter estimation): For $t \in [t_{k+1}, t_{k+2})$, the parameter estimate $\theta_{k+1,i}(t; p_i(t_{k+1}), q_i(t_{k+1}))$ is obtained by replacing (p, q) in Step 2 by $(p_i(t_{k+1}), q_i(t_{k+1}))$.

Output: The order estimate $(p_i(t_{k+1}), q_i(t_{k+1}))$ and the parameter estimate $\theta_{k+1,i}(t; p_i(t_{k+1}), q_i(t_{k+1}))$.

3 Convergence analysis of Algorithm 2

In this section, we will establish theoretical results for convergence of the distributed order estimation algorithm (Algorithm 2) proposed in Sect. 2.2. For this purpose, we introduce some assumptions on the network topology and regressors.

Assumption 1 The graph \mathcal{G} is undirected and connected.

Remark 1 Denote the i th row and j th column element of the matrix \mathcal{A}^m as $a_{ij}^{(m)}$. By Lemma 8.1.2 in [37], we know that under Assumption 1, $a_{ij}^{(m)} \geq a_{\min} > 0$ holds for all $m > D_{\mathcal{G}}$, where $a_{\min} = \min_{i,j \in \mathcal{V}} a_{ij}^{(D_{\mathcal{G}})} > 0$.

In fact, the above assumption for the communication graph can be generalized to the case where \mathcal{G} is a strongly connected balanced graph, and the corresponding analysis is similar to the undirected graph case. Thus, we just provide the analysis of Algorithm 2 under Assumption 1.

The properties of regressors are important for the convergence of the identification algorithms. In this paper, we introduce the following cooperative excitation condition on the regressors.

Assumption 2 (Cooperative excitation condition) There exists a real function $\xi(t)$ that satisfies

- (1) $\xi(t) > 0$, for $t > 0$;
- (2) $\xi(t) \rightarrow \infty$, as $t \rightarrow \infty$;
- (3) $\lim_{t \rightarrow \infty} \frac{\log(R(t; U_p, U_q))}{\xi(t)} = 0$, $\lim_{t \rightarrow \infty} \frac{\xi(t)}{\lambda_{\min}(t; p, q)} = 0$, $(p, q) \in I^*$, a.s. where $I^* = \{(\bar{p}, U_q), (U_p, \bar{q})\}$,

$$R(t; U_p, U_q) = \lambda_{\max} \left\{ P_0^{-1}(t_0; U_p, U_q) + \sum_{i=1}^N \int_{t_0}^t \|\phi_i(s; U_p, U_q)\|^2 ds, \lambda_{\min}(n; p, q) \right\}$$

$$= \lambda_{\min} \left\{ \sum_{j=1}^N \int_{t_0}^{t_{n-D_G+1}} \phi_j(s; p, q) \phi_j^T(s; p, q) ds + \sum_{j=1}^N P_{0,j}^{-1}(t_0; p, q) \right\}.$$

Remark 2 The choice of $\xi(t)$ depends on the regressors, and here we show how to choose it for a special case. Assume that for $i \in \{1, \dots, N\}$, the regression vector $\phi_i(s; \bar{p}, \bar{q})$ is bounded, and there exists a matrix Υ_i such that $\frac{1}{t} \int_0^t \phi_i(s; \bar{p}, \bar{q}) \phi_i^T(s; \bar{p}, \bar{q}) ds \xrightarrow{t \rightarrow \infty} \Upsilon_i$, and $\sum_{i=1}^N \Upsilon_i$ is positive definite [41]. Then $\xi(t)$ can be taken as $\xi(t) = t^\alpha$, $0 < \alpha < 1$.

Before giving convergence results for the estimation of the system order and parameters in Algorithm 2, we give some preliminary lemmas. We first provide the upper bound on the accumulative noise, which is an important step for the convergence of the algorithm.

Lemma 3 For fixed order (p, q) , let

$$H_{n+1,i}(t_{n+1}; p, q) \triangleq \sum_{j=1}^N \sum_{k=0}^n a_{ij}^{(n+1-k)} \int_{t_k}^{t_{k+1}} \phi_j(s; p, q) dv_j(s),$$

$$H_{0,i}(t_0; p, q) = 0.$$

Then the following inequality holds:

$$H_{n+1,i}^T(t_{n+1}; p, q) P_{n+1,i}(t_{n+1}; p, q) H_{n+1,i}(t_{n+1}; p, q) = O(\log R(n; p, q)),$$

where $P_{n+1,i}(t_{n+1}; p, q)$ is defined in (8). □

Proof Denote

$$\begin{cases} \Psi(t; p, q) \triangleq \text{diag}\{\phi_1(t; p, q), \dots, \phi_N(t; p, q)\}, \\ V(t) \triangleq (v_1(t), \dots, v_N(t))^T, \\ P_n(t) \triangleq \text{diag}\{P_{n,1}(t; p, q), \dots, P_{n,N}(t; p, q)\}, \\ H_n(t) \triangleq \text{col}\{H_{n,1}(t; p, q), \dots, H_{n,N}(t; p, q)\}. \end{cases} \quad (12)$$

For simplicity of expression, we omit (p, q) in the equations without confusion. By the definition of $H_{n+1,i}(t_{n+1})$ and (12), we have for $t = t_{n+1}$,

$$H_{n+1}(t_{n+1}) = \sum_{k=0}^n \mathcal{A}^{(n+1-k)} \int_{t_k}^{t_{k+1}} \Psi(s) dV(s) = \mathcal{A}(H_n(t_n) + \int_{t_n}^{t_{n+1}} \Psi(s) dV(s)), \quad (13)$$

and for $t \in [t_n, t_{n+1})$, we have

$$H_n(t) = \sum_{k=0}^{n-1} \mathcal{A}^{(n-k)} \int_{t_k}^{t_{k+1}} \Psi(s) dV(s) + \int_{t_n}^t \Psi(s) dV(s), \quad (14)$$

where \mathcal{A} is the adjacency matrix of graph \mathcal{G} . Differentiating both sides of (14), we have $dH_n(t) = \Psi(t) dV(t)$. Integrating both sides of this equation over $[t_n, t_{n+1})$ and substituting it into (13), we can derive the following equation:

$$H_{n+1}(t_{n+1}) = \mathcal{A} H_n(t_{n+1}). \quad (15)$$

Consider the stochastic Lyapunov function $H_n^T(t) P_n(t) H_n(t)$. By the Itô formula in Lemma 2, we have

$$\begin{aligned} d(H_n^T(t) P_n(t) H_n(t)) &= H_n^T(t) dP_n(t) H_n(t) + 2H_n^T(t) P_n(t) dH_n(t) \\ &\quad + \text{tr}(P_n(t) \Psi(t) \Psi^T(t)) dt \\ &= -H_n^T(t) (P_n(t) \Psi(t) \Psi^T(t) P_n(t)) H_n(t) dt \\ &\quad + 2H_n^T(t) P_n(t) \Psi(t) dV(t) \\ &\quad + \text{tr}(P_n(t) \Psi(t) \Psi^T(t)) dt. \end{aligned}$$

Integrating both sides of the equation over $[t_n, t_{n+1})$ yields

$$\begin{aligned} H_n^T(t_{n+1}) P_n(t_{n+1}) H_n(t_{n+1}) &= H_n^T(t_n) P_n(t_n) H_n(t_n) \\ &\quad - \int_{t_n}^{t_{n+1}} H_n^T(t) P_n(t) \Psi(t) \Psi^T(t) P_n(t) H_n(t) dt \\ &\quad + 2 \int_{t_n}^{t_{n+1}} H_n^T(t) P_n(t) \Psi(t) dV(t) \\ &\quad + \int_{t_n}^{t_{n+1}} \text{tr}(P_n(t) dP_n^{-1}(t)). \end{aligned} \quad (16)$$

By (15), (16) and lemma A.1 in [10], we have

$$\begin{aligned}
 & H_{n+1}^T(t_{n+1})P_{n+1}(t_{n+1})H_{n+1}(t_{n+1}) \\
 &= H_n^T(t_{n+1})\mathcal{A}P_{n+1}(t_{n+1})\mathcal{A}H_n(t_{n+1}) \\
 &\leq H_n^T(t_{n+1})P_n(t_{n+1})H_n(t_{n+1}) \\
 &\leq H_n^T(t_n)P_n(t_n)H_n(t_n) \\
 &\quad - \int_{t_n}^{t_{n+1}} H_n^T(t)P_n(t)\Psi(t)\Psi^T(t)P_n(t)H_n(t)dt \\
 &\quad + 2 \int_{t_n}^{t_{n+1}} H_n^T(t)P_n(t)\Psi(t)dV(t) \\
 &\quad + \int_{t_n}^{t_{n+1}} \text{tr}(P_n(t)dP_n^{-1}(t)) \\
 &\leq H_0^T(t_0)P_0(t_0)H_0(t_0) \\
 &\quad - \sum_{k=0}^n \int_{t_k}^{t_{k+1}} H_k^T(t)P_k(t)\Psi(t)\Psi^T(t)P_k(t)H_k(t)dt \\
 &\quad + 2 \sum_{k=0}^n \int_{t_k}^{t_{k+1}} H_k^T(t)P_k(t)\Psi(t)dV(t) \\
 &\quad + \sum_{k=0}^n \int_{t_k}^{t_{k+1}} \text{tr}(P_k(t)dP_k^{-1}(t)).
 \end{aligned}$$

This leads to the following equivalent form:

$$\begin{aligned}
 & H_{n+1}^T(t_{n+1})P_{n+1}(t_{n+1})H_{n+1}(t_{n+1}) \\
 &\quad + \sum_{k=0}^n \int_{t_k}^{t_{k+1}} \|\Psi^T(t)P_k(t)H_k(t)\|^2 dt \\
 &\leq 2 \sum_{k=0}^n \int_{t_k}^{t_{k+1}} H_k^T(t)P_k(t)\Psi(t)dV(t) \\
 &\quad + \sum_{k=0}^n \int_{t_k}^{t_{k+1}} \text{tr}(P_k(t)dP_k^{-1}(t)).
 \end{aligned} \tag{17}$$

By Theorem 1, the first term on the right side of (17) satisfies

$$\begin{aligned}
 & \sum_{k=0}^n \int_{t_k}^{t_{k+1}} H_k^T(t)P_k(t)\Psi(t)dV(t) \\
 &= O(1) + o\left(\sum_{k=0}^n \int_{t_k}^{t_{k+1}} \|\Psi^T(t)P_k(t)H_k(t)\|^2 dt\right),
 \end{aligned} \tag{18}$$

and a tedious derivation leads to

$$\sum_{k=0}^n \int_{t_k}^{t_{k+1}} \text{tr}(P_k(t)dP_k^{-1}(t)) = O(\log R(n)). \tag{19}$$

Substituting (18) and (19) into (17), we have

$$H_{n+1}^T(t_{n+1})P_{n+1}(t_{n+1})H_{n+1}(t_{n+1}) = O(\log R(n)).$$

This completes the proof of the lemma. \square

As we know, in the estimation process of the algorithm, it is possible for the estimated order to be greater than the true

order. In the following lemma, we provide an analysis of the parameter estimation error for this case, which is helpful for the subsequent theoretical analysis.

Lemma 4 For any $p \geq \bar{p}$ and $q \geq \bar{q}$, the estimation error $\tilde{\theta}_{k,i}(t; p, q)$ of Algorithm 1 satisfies the following equation,

$$\begin{aligned}
 & P_{k+1,i}^{-1}(t_{k+1}; p, q)\tilde{\theta}_{k+1,i}(t_{k+1}; p, q) \\
 &= \sum_{j \in \mathcal{N}_i} a_{ij} \left[P_{k,j}^{-1}(t_k; p, q)\tilde{\theta}_{k,j}(t_k; p, q) \right. \\
 &\quad \left. - \int_{t_k}^{t_{k+1}} \phi_j(s; p, q)dv_j(s) \right],
 \end{aligned}$$

where $\tilde{\theta}_{k,i}(t; p, q) = [a_1 - a_{1,i}(t), \dots, a_p - a_{p,i}(t), b_1 - b_{1,i}(t), \dots, b_q - b_{q,i}(t)]^T$, $\{a_{\tau,i}(t)\}_{\tau=1}^p$ and $\{b_{r,i}(t)\}_{r=1}^q$ are components of $\theta_{k,i}(t; p, q)$ generated by Algorithm 1.

Proof By (7) we have

$$dP_{k,i}^{-1}(t; p, q) = \phi_i(t; p, q)\phi_i^T(t; p, q)dt. \tag{20}$$

By (8) and (20), we have

$$\begin{aligned}
 & d(P_{k,i}^{-1}(t; p, q)\tilde{\theta}_{k,i}(t; p, q)) \\
 &= \phi_i(t; p, q)\phi_i^T(t; p, q)\tilde{\theta}_{k,i}(t; p, q)dt \\
 &\quad - \phi_i(t; p, q)[\phi_i^T(t; p, q)\tilde{\theta}_{k,i}(t; p, q)dt + dv_i(t)] \\
 &= -\phi_i(t; p, q)dv_i(t).
 \end{aligned}$$

Integrating the above equation over $[t_k, t_{k+1})$ leads to

$$\begin{aligned}
 & P_{k,i}^{-1}(t_{k+1}; p, q)\tilde{\theta}_{k,i}(t_{k+1}; p, q) \\
 &= P_{k,i}^{-1}(t_k; p, q)\tilde{\theta}_{k,i}(t_k; p, q) - \int_{t_k}^{t_{k+1}} \phi_i(s; p, q)dv_i(s).
 \end{aligned} \tag{21}$$

Thus, we have

$$\begin{aligned}
 & P_{k+1,i}^{-1}(t_{k+1}; p, q)\tilde{\theta}_{k+1,i}(t_{k+1}; p, q) \\
 &= P_{k+1,i}^{-1}(t_{k+1}; p, q)\theta_{k+1,i}(t_{k+1}; p, q) \\
 &\quad - P_{k+1,i}^{-1}(t_{k+1}; p, q)\theta(p, q) \\
 &= \sum_{j \in \mathcal{N}_i} a_{ij} [P_{k,j}^{-1}(t_{k+1}; p, q)\tilde{\theta}_{k,j}(t_{k+1}; p, q)] \\
 &= \sum_{j \in \mathcal{N}_i} a_{ij} [P_{k,j}^{-1}(t_k; p, q)\tilde{\theta}_{k,j}(t_k; p, q) \\
 &\quad - \int_{t_k}^{t_{k+1}} \phi_j(s; p, q)dv_j(s)],
 \end{aligned}$$

where the last equation is obtained from (21). We complete the proof. \square

Now, we give the convergence result for the order estimation of Algorithm 2.

Theorem 5 Under Assumptions 1 and 2, for any $i \in \{1, \dots, n\}$, the order $(p_i(t), q_i(t))$ estimated by Algorithm 2 converges to the true order (\bar{p}, \bar{q}) almost surely, i.e.,

$$\lim_{t \rightarrow \infty} (p_i(t), q_i(t)) = (\bar{p}, \bar{q}) \text{ a.s., } i \in \{1, \dots, n\}.$$

Proof We will show that for any $i \in \{1, \dots, n\}$, $(p_i(t), q_i(t))$ has one and only one limit point (\bar{p}, \bar{q}) by reduction to absurdity. Suppose $(p'_i, q'_i) \in I$ is another limit point of $(p_i(t), q_i(t))$ and $(p'_i, q'_i) \neq (\bar{p}, \bar{q})$. Then there exists a subsequence $\{t_{m_k}\}$ such that

$$\lim_{k \rightarrow \infty} (p_i(t_{m_k}), q_i(t_{m_k})) = (p'_i, q'_i). \tag{22}$$

To prove that $\lim_{t \rightarrow \infty} (p_i(t), q_i(t)) = (\bar{p}, \bar{q})$, we need to show that neither of the following two cases holds,

- (1) $p'_i \geq \bar{p}, q'_i \geq \bar{q}$ and $p'_i + q'_i > \bar{p} + \bar{q}$;
- (2) $p'_i < \bar{p}$ or $q'_i < \bar{q}$.

We first consider the case (1). By (1) and (5), we have

$$\begin{aligned} & J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) \\ &= \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \left[\tilde{\theta}_{m_k,i}^T(t_{m_k}; p'_i, q'_i) \right. \\ & \quad \cdot \phi_j(s; p'_i, q'_i) + \left. \frac{dv_j(s)}{ds} \right]^2 ds \\ &= \tilde{\theta}_{m_k,i}^T(t_{m_k}; p'_i, q'_i) \\ & \quad \cdot \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; p'_i, q'_i) \right. \\ & \quad \cdot \phi_j^T(s; p'_i, q'_i) ds \left. \right] \tilde{\theta}_{m_k,i}(t_{m_k}; p'_i, q'_i) \\ & \quad + 2\tilde{\theta}_{m_k,i}^T(t_{m_k}; p'_i, q'_i) \\ & \quad \cdot \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; p'_i, q'_i) dv_j(s) \right] \\ & \quad + \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \left[\frac{dv_j(s)}{ds} \right]^2 ds \\ &\triangleq \text{I} + \text{II} + \text{III}. \tag{23} \end{aligned}$$

We first estimate the term I. For any p, q , integrating both sides of the equation (20) over $[t_k, t_{k+1}]$, we have

$$\begin{aligned} & P_{k,i}^{-1}(t_{k+1}; p, q) \\ &= P_{k,i}^{-1}(t_k; p, q) + \int_{t_k}^{t_{k+1}} \phi_i(s; p, q) \phi_i^T(s; p, q) ds. \tag{24} \end{aligned}$$

Substituting the above equation into (8), we obtain that

$$\begin{aligned} & P_{k+1,i}^{-1}(t_{k+1}; p, q) = \sum_{j \in \mathcal{N}_i} a_{ij} \left\{ P_{k,j}^{-1}(t_k; p, q) \right. \\ & \quad \left. + \int_{t_k}^{t_{k+1}} \phi_j(s; p, q) \phi_j^T(s; p, q) ds \right\}. \tag{25} \end{aligned}$$

Moreover we have by (24) and (25)

$$\begin{aligned} & P_{m_k,i}^{-1}(t_{m_k}; p, q) \\ &= \sum_{j \in \mathcal{N}_i} a_{ij} \left\{ P_{m_k-1,j}^{-1}(t_{m_k}; p, q) \right. \\ & \quad \left. + \int_{t_{m_k-1}}^{t_{m_k}} \phi_j(s; p, q) \phi_j^T(s; p, q) ds \right\} \\ &= \sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; p, q) \\ & \quad + \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; p, q) \phi_j^T(s; p, q) ds. \tag{26} \end{aligned}$$

By (26) and Theorem 3.8 in [42] it follows that

$$\text{I} = O(\log R(t_{m_k}; p'_i, q'_i)) = O(\log R(t_{m_k}; U_p, U_q)). \tag{27}$$

For the second term II in (23), it is clear that

$$\begin{aligned} \|\text{II}\| &\leq 2 \left\| \tilde{\theta}_{m_k,i}^T(t_{m_k}; p'_i, q'_i) P_{m_k,i}^{-\frac{1}{2}}(t_{m_k}; p'_i, q'_i) \right\| \\ & \quad \cdot \left\| P_{m_k,i}^{\frac{1}{2}}(t_{m_k}; p'_i, q'_i) \right. \\ & \quad \cdot \left. \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; p'_i, q'_i) dv_j(s) \right] \right\|. \tag{28} \end{aligned}$$

By applying Theorem 3.8 in [42] to the first term in (28), we can get the following equation,

$$\begin{aligned} & \left\| \tilde{\theta}_{m_k,i}^T(t_{m_k}; p'_i, q'_i) P_{m_k,i}^{-\frac{1}{2}}(t_{m_k}; p'_i, q'_i) \right\| \\ &= O(\log^{\frac{1}{2}}(R(t_{m_k}; p'_i, q'_i))) \\ &= O(\log^{\frac{1}{2}}(R(t_{m_k}; U_p, U_q))). \tag{29} \end{aligned}$$

From Lemma 3, we see that the second term in (28) satisfies

$$\begin{aligned} & \left\| P_{m_k,i}^{\frac{1}{2}}(t_{m_k}; p'_i, q'_i) \right. \\ & \quad \cdot \left. \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; p'_i, q'_i) dv_j(s) \right] \right\| \end{aligned}$$

$$\begin{aligned}
 &= O(\log^{\frac{1}{2}}(R(t_{m_k}; p'_i, q'_i))) \\
 &= O(\log^{\frac{1}{2}}(R(t_{m_k}; U_p, U_q))). \tag{30}
 \end{aligned}$$

By(27)–(29), we see that there exists a constant $M_1 > 0$ such that the following inequality holds,

$$\begin{aligned}
 &J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) - \text{III} \\
 &\geq -M_1 \log(R(t_{m_k}; U_p, U_q)). \tag{31}
 \end{aligned}$$

Next we consider $J_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}, \theta_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}))$. By a similar derivation as that of (23), we have

$$\begin{aligned}
 &J_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}, \theta_{m_k,i}(t_{m_k}; \bar{p}, \bar{q})) \\
 &= \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \bar{p}, \bar{q}) \right. \\
 &\quad \cdot \phi_j^T(s; \bar{p}, \bar{q}) ds \Big] \tilde{\theta}_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) + 2\tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \\
 &\quad \cdot \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \bar{p}, \bar{q}) dv_j(s) \right] \\
 &\quad + \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \left[\frac{dv_j(s)}{ds} \right]^2 ds \\
 &\triangleq \text{I}' + \text{II}' + \text{III}. \tag{32}
 \end{aligned}$$

By (26) we have

$$\begin{aligned}
 \text{I}' &= \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) P_{m_k,i}^{-1}(t_{m_k}; \bar{p}, \bar{q}) \tilde{\theta}_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\
 &\quad - \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \left(\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \right) \\
 &\quad \cdot \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}). \tag{33}
 \end{aligned}$$

For the second term II' , we have by Lemma 4,

$$\begin{aligned}
 &P_{m_k,i}^{-1}(t_{m_k}; \bar{p}, \bar{q}) \tilde{\theta}_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\
 &= \sum_{j=1}^N a_{ij} P_{m_k-1,j}^{-1}(t_{m_k-1}; \bar{p}, \bar{q}) \tilde{\theta}_{m_k-1,j}(t_{m_k-1}; \bar{p}, \bar{q}) \\
 &\quad - \sum_{j=1}^N a_{ij} \int_{t_{m_k-1}}^{t_{m_k}} \phi_j(s; \bar{p}, \bar{q}) dv_j(s) \\
 &= \sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \tilde{\theta}_{0,j}(t_0; \bar{p}, \bar{q}) \\
 &\quad - \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \bar{p}, \bar{q}) dv_j(s). \tag{34}
 \end{aligned}$$

Substituting (34) into II' yields the following equation,

$$\begin{aligned}
 \text{II}' &= 2\tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \\
 &\quad \cdot \left[\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \tilde{\theta}_{0,j}(t_0; \bar{p}, \bar{q}) \right. \\
 &\quad \left. - P_{m_k,i}^{-1}(t_{m_k}; \bar{p}, \bar{q}) \tilde{\theta}_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \right]. \tag{35}
 \end{aligned}$$

Substituting (33) and (35) into (32) yields

$$\begin{aligned}
 &J_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}, \theta_{m_k,i}(t_{m_k}; \bar{p}, \bar{q})) - \text{III} \\
 &= \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) P_{m_k,i}^{-1}(t_{m_k}; \bar{p}, \bar{q}) \tilde{\theta}_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\
 &\quad - \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \left(\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \right) \\
 &\quad \cdot \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) + 2\tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \\
 &\quad \cdot \left[\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \tilde{\theta}_{0,j}(t_0; \bar{p}, \bar{q}) \right. \\
 &\quad \left. - P_{m_k,i}^{-1}(t_{m_k}; \bar{p}, \bar{q}) \tilde{\theta}_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \right] \\
 &\leq -\tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \left(\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \right) \\
 &\quad \cdot \tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) + 2\tilde{\theta}_{m_k,i}^T(t_{m_k}; \bar{p}, \bar{q}) \\
 &\quad \cdot \left(\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \tilde{\theta}_{0,j}(t_0; \bar{p}, \bar{q}) \right) \\
 &\leq \sum_{j=1}^N a_{ij}^{(m_k)} \tilde{\theta}_{0,j}^T(t_0; \bar{p}, \bar{q}) P_{0,j}^{-1}(t_0; \bar{p}, \bar{q}) \tilde{\theta}_{0,j}(t_0; \bar{p}, \bar{q}) \\
 &= O(1). \tag{36}
 \end{aligned}$$

where the last inequality uses the inequality $2x^T Ay \leq x^T Ax + y^T Ay$ for $A \geq 0$. From (31) and (36), there exists a constant $M_2 > 0$ such that the following inequality holds,

$$\begin{aligned}
 &J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) \\
 &\quad - J_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}, \theta_{m_k,i}(t_{m_k}; \bar{p}, \bar{q})) \\
 &\geq -M_1 \log(R(t_{m_k}; U_p, U_q)) - M_2. \tag{37}
 \end{aligned}$$

Note that $p_i(t_{m_k}), q_i(t_{m_k})$ in (22) are positive integers. Then, there exists a $K \in \mathbb{N}^+$ such that for $k > K$, $(p_i(t_{m_k}), q_i(t_{m_k})) \equiv (p'_i, q'_i)$. By

$$(p_i(t_{m_k}), q_i(t_{m_k})) = \arg \min_{(p,q) \in I} F_{m_k,i}(t_{m_k}; p, q),$$

we know that

$$F_{m_k,i}(t_{m_k}; p_i(t_{m_k}), p_i(t_{m_k})) - F_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \leq 0. \tag{38}$$

On the other hand, by (37) and Assumption 2, we have for sufficiently large k ,

$$\begin{aligned}
 & F_{m_k,i}(t_{m_k}; p_i(t_{m_k}), q_i(t_{m_k})) - F_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\
 &= F_{m_k,i}(t_{m_k}; p'_i, q'_i) - F_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\
 &= J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) \\
 &\quad - J_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}, \theta_{m_k,i}(t_{m_k}; \bar{p}, \bar{q})) \\
 &\quad + (p'_i + q'_i - \bar{p} - \bar{q}) \xi(t_{m_k}) \\
 &\geq -M_1 \log(R(t_{m_k}; U_p, U_q)) - M_2 \\
 &\quad + (p'_i + q'_i - \bar{p} - \bar{q}) \xi(t_{m_k}) \\
 &= -M_2 + \xi(t_{m_k}) \left(\frac{-M_1 \log(R(t_{m_k}; U_p, U_q))}{\xi(t_{m_k})} \right) \\
 &\quad + (p'_i + q'_i - \bar{p} - \bar{q}) \rightarrow \infty, \quad k \rightarrow \infty.
 \end{aligned}$$

This contradicts (38), so case (1) does not hold.

We next show that the case (2) will not happen. If case (2) holds, i.e., $p'_i < \bar{p}$ or $q'_i < \bar{q}$ holds. We construct the following $(\kappa_i + \mu_i)$ dimensional vector,

$$\theta_{m_k,i}(t_{m_k}; \kappa_i, \mu_i) = [a_{1,i}(t_{m_k}), \dots, a_{\kappa_i,i}(t_{m_k}), b_{1,i}(t_{m_k}), \dots, b_{\mu_i,i}(t_{m_k})]^T,$$

where $\kappa_i = \max\{p'_i, \bar{p}\}$, $\mu_i = \max\{q'_i, \bar{q}\}$, $a_{m,i}(t_{m_k}) \triangleq 0$ ($p'_i < m \leq \bar{p}$) if $p'_i < \bar{p}$; $b_{m,i}(t_{m_k}) \triangleq 0$ ($q'_i < m \leq \bar{q}$) if $q'_i < \bar{q}$.

Denote $\tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i) \triangleq \theta(\kappa_i, \mu_i) - \theta_{m_k,i}(t_{m_k}; \kappa_i, \mu_i)$. Then we have

$$\|\tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i)\|^2 \geq \min\left\{|a_{\bar{p}}|^2, |b_{\bar{q}}|^2\right\} \triangleq \gamma_0 > 0. \tag{39}$$

Similar to the analysis of (23) and (26), we have

$$\begin{aligned}
 & J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) - \text{III} \\
 &= \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \left[\frac{dy_j(s)}{ds} - \theta_{m_k,i}^T(t_{m_k}; p'_i, q'_i) \phi_j(s; p'_i, q'_i) \right]^2 ds \\
 &= \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) P_{m_k,i}^{-1}(t_{m_k}; \kappa_i, \mu_i) \cdot \tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i) - \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) \\
 &\quad \cdot \left(\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \kappa_i, \mu_i) \right) \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) \\
 &\quad + 2\tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) \cdot \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \kappa_i, \mu_i) dv_j(s) \right]. \tag{40}
 \end{aligned}$$

The following is a term-by-term analysis of the right-hand side of (40). From (26) and Remark 1 it follows that for any p and q the following inequality holds,

$$\lambda_{\min}(P_{m_k,i}^{-1}(t_{m_k}; p, q)) \geq a_{\min} \lambda_{\min}(t_{m_k}; p, q), \tag{41}$$

where $a_{\min} = \min_{i,j \in \mathcal{V}} a_{ij}^{(D_G)} > 0$. From (39), (41) and Theorem 2.6 in [13], we obtain the following estimate of the first term at the right-hand side of the equation (40),

$$\begin{aligned}
 & \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) P_{m_k,i}^{-1}(t_{m_k}; \kappa_i, \mu_i) \tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i) \\
 &\geq \gamma_0 a_{\min} \lambda_{\min}(t_{m_k}; \kappa_i, \mu_i) \\
 &\geq \gamma_0 a_{\min} \min\{\lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q})\}. \tag{42}
 \end{aligned}$$

By (41) and Theorem 3.8 in [42], we have for any $p \geq \bar{p}$ and $q \geq \bar{q}$,

$$\begin{aligned}
 & \sum_{i=1}^n \|\tilde{\theta}_{n,i}(t; p, q)\|^2 \\
 &= O\left(\frac{\log R(n; p, q)}{\lambda_{\min}(n; p, q)}\right), \quad t \in [t_n, t_{n+1}]. \tag{43}
 \end{aligned}$$

When $p'_i < \bar{p}$, ($q'_i < \bar{q}$ is the same), for the first p'_i components of $\tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i)$ utilizing (43), then from Theorem 2.6 in [13] and Assumption 2, we have $\|\tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i)\| = O(1)$. Thus the second term at the right-hand side of (40) satisfies the following relation,

$$\begin{aligned}
 & \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) \left(\sum_{j=1}^N a_{ij}^{(m_k)} P_{0,j}^{-1}(t_0; \kappa_i, \mu_i) \right) \cdot \tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i) \\
 &\leq \lambda_{\max} \left(\sum_{j=1}^n a_{ij}^{(t)} P_{0,j}^{-1}(\kappa_i, \mu_i) \right) \|\tilde{\theta}_{t_k,i}(\kappa_i, \mu_i)\|^2 \\
 &= O(1). \tag{44}
 \end{aligned}$$

Following an analysis similar to that of (28), the following estimate of the third term at the right-hand side of (40) can be obtained by using Lemma 3,

$$\begin{aligned}
 & \left| \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) \left[\sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \kappa_i, \mu_i) dv_j(s) \right] \right| \\
 &\leq \left\| \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) P_{m_k,i}^{-\frac{1}{2}}(t_{m_k}; \kappa_i, \mu_i) \right\| \cdot \left\| P_{m_k,i}^{\frac{1}{2}}(t_{m_k}; \kappa_i, \mu_i) \right\|
 \end{aligned}$$

$$\begin{aligned} & \cdot \left\| \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \kappa_i, \mu_i) dv_j(s) \right\| \\ &= O \left\{ \left\| \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) P_{m_k,i}^{-\frac{1}{2}}(t_{m_k}; \kappa_i, \mu_i) \right\| \right. \\ & \quad \cdot \log^{\frac{1}{2}} \left(R(t_{m_k}; U_p, U_q) \right) \left. \right\}, \end{aligned} \tag{45}$$

Continuing from (39), (41) and Assumption 2 the final estimate of the third term at the right-hand side of (40) is

$$\begin{aligned} & \left\| \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) \right. \\ & \quad \cdot \left. \sum_{j=1}^N \sum_{\tau=0}^{m_k-1} a_{ij}^{(m_k-\tau)} \int_{t_\tau}^{t_{\tau+1}} \phi_j(s; \kappa_i, \mu_i) dv_j(s) \right\| \\ &= O \left\{ \tilde{\theta}_{m_k,i}^T(t_{m_k}; \kappa_i, \mu_i) P_{m_k,i}^{-1}(t_{m_k}; \kappa_i, \mu_i) \right. \\ & \quad \cdot \tilde{\theta}_{m_k,i}(t_{m_k}; \kappa_i, \mu_i) \left. \right\}. \end{aligned} \tag{46}$$

Substituting (42), (44) and (46) into (40), it can be shown that there exists a constant $M_3 > 0$ such that the following inequality holds,

$$\begin{aligned} & J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) - \text{III} \\ & \geq \gamma_0 a_{\min} \min \{ \lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q}) \} \\ & \quad \cdot (1 + o(1)) - M_3 \\ & \geq \frac{\gamma_0 a_{\min} \min \{ \lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q}) \}}{2} \\ & \quad - M_3. \end{aligned} \tag{47}$$

From (10), (36) and (47) it follows that the following inequality holds for sufficiently large k ,

$$\begin{aligned} & 0 \geq F_{m_k,i}(t_{m_k}; p_i(t_{m_k}), q_i(t_{m_k})) - F_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\ & = F_{m_k,i}(t_{m_k}; p'_i, q'_i) - F_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}) \\ & = J_{m_k,i}(t_{m_k}; p'_i, q'_i, \theta_{m_k,i}(t_{m_k}; p'_i, q'_i)) \\ & \quad - J_{m_k,i}(t_{m_k}; \bar{p}, \bar{q}, \theta_{m_k,i}(t_{m_k}; \bar{p}, \bar{q})) \\ & \quad + (p'_i + q'_i - \bar{p} - \bar{q}) \xi(t_{m_k}) \\ & \geq \frac{\gamma_0 a_{\min} \min \{ \lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q}) \}}{2} \\ & \quad - M_3 + (p'_i + q'_i - \bar{p} - \bar{q}) \xi(t_{m_k}) \\ & \geq -M_3 \\ & \quad + \frac{\gamma_0 a_{\min} \min \{ \lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q}) \}}{2} \\ & \quad \cdot \left\{ 1 + \frac{(p'_i + q'_i - \bar{p} - \bar{q}) \xi(t_{m_k})}{\gamma_0 a_{\min} \min \{ \lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q}) \}} \right\} \\ & \geq -M_3 \\ & \quad + \frac{\gamma_0 a_{\min} \min \{ \lambda_{\min}(t_{m_k}; \bar{p}, U_q), \lambda_{\min}(t_{m_k}; U_p, \bar{q}) \}}{4} \\ & \rightarrow \infty, k \rightarrow \infty. \end{aligned}$$

There is a contradiction, which proves that case (2) is not valid either. \square

Noting that $(p_i(t), q_i(t))$ and (\bar{p}, \bar{q}) are integers, it follows from Theorem 5 that there exists a sufficiently large integer T such that, for any $t \geq T$, $p_i(t) = \bar{p}$ and $q_i(t) = \bar{q}$ hold. Thus by (43) and Assumption 2, the following convergence result on the unknown parameter vector can be obtained.

Theorem 6 Under the condition of Theorem 5, for all $i \in \{1, \dots, n\}$ the following equation holds,

$$\theta_{n,i}(t; p_i(t), q_i(t)) \xrightarrow[t \rightarrow \infty]{} \theta(\bar{p}, \bar{q}), \text{ a.s.,}$$

where $\theta_{n,i}(t; p_i(t), q_i(t))$ is given by Algorithm 2.

Theorem 6 provides the almost sure convergence result for Algorithm 2 when both the order and parameters in (2) are unknown. We see that the convergence results are established without assuming the statistical properties of regression vectors, such as independency, stationarity or ergodicity, which makes our results applicable for feedback control systems.

4 Simulation results

In this section, we will verify effectiveness of the proposed distributed order estimation algorithm (Algorithm 2) by a simulation example.

Example 1 Consider a network of ten sensors ($N = 10$). The dynamics of each sensor is described by the following stochastic differential equation

$$y_i(t) = S\phi_i^T(t)\theta + v_i(t), \quad i = 1, 2, \dots, 10,$$

where both the parameter vector $\theta \in \mathbb{R}^{\bar{q}}$ and the system order \bar{q} are to be estimated, $v_i(t)$ is the system noise following the standard Wiener process. For $q = U_q$, the regressor $\phi_i(t) \in \mathbb{R}^{U_q}$ ($t \geq 0$), $i \in \{1, \dots, 10\}$) is taken as

$$\begin{aligned} x_i(t) &= \begin{cases} 0.5t + 0.5, & \text{if } \text{mod}(i, 3) = 0, \\ t + 0.5, & \text{if } \text{mod}(i, 3) = 1, \\ t^2 + 0.5, & \text{if } \text{mod}(i, 3) = 2, \end{cases} \\ \phi_i(t) &= \underbrace{[e_1, \dots, \mathbf{0}, e_\tau, \mathbf{0}, \dots, \mathbf{0}]^T}_{U_q} X_i(t), \end{aligned} \tag{48}$$

where $X_i(t) = \underbrace{(x_i(t), \dots, x_i(t))^T}_{U_q}$, $\mathbf{0}$ denotes the U_q -dimensional column vector with all zero elements, $\tau = 2 + \text{mod}(i, U_q - 1)$, the mod is a remainder operator, and e_τ denotes the τ th column of the unit matrix I_{U_q} . For $q < U_q$, the regressor $\phi_i(t)$ is taken as the first q elements of (48). The network topology \mathcal{G} is shown in Fig. 1, whose weighted

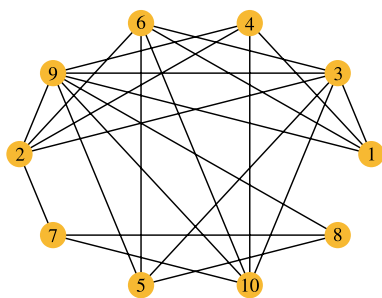


Fig. 1 Topology of the sensor network

adjacency matrix $\mathcal{A} = (a_{mi})_{10 \times 10}$ is determined by the Metropolis rule (cf. [43]).

All sensors will estimate the true system order of $\bar{q} = 3$ and the unknown parameter vector $\theta = (0.3, 0.4, 1.2)^T$. It is clear that the graph \mathcal{G} is connected and the regressor $\phi_i(t)$ of all sensors $i \in \{1, \dots, 10\}$ satisfies Assumption 2 with $\xi(t) = t^{\frac{1}{2}}$.

Set the initial estimate $\theta_i(0) = \underbrace{(0, \dots, 0)^T}_q$, the fusion time instants $t_0 = 0, 0.1, 0.2, 0.3, \dots$, the upper bound of system order is taken as $U_q = 5$. Algorithm 2 is conducted 100 runs with the same initial setting, and the order estimates $\{q_i(t)\}$, $i = 1, 2, \dots, 10$ and the average estimation error of parameters can be obtained.

Figure 2 shows the simulation results for the order estimate obtained by Algorithm 2 proposed in this paper and the non-distributed algorithm (namely, $N = 1$ and $D_G = 1$). It can be seen that the sequences of order estimates $\{q_i(t)\}$ ($i = 1, \dots, 10$) by using Algorithm 2 converges to the true order, and not all sequences of order estimates can converge to the true order for the non-distributed order estimation algorithm.

Figure 3 shows the simulation results of the average estimation error of parameters by Algorithm 2 and the non-distributed algorithm. It can be seen that the estimation errors generated by Algorithm 2 for each sensor i converge to zero, while the estimation errors generated by non-distributed algorithm do not tend to zero.

From this simulation example, we can reveal the cooperative effect of multiple sensors: all sensors can cooperatively estimate unknown system order and parameters, while none of them can do it separately.

5 Concluding remarks

In this paper, we study the distributed order estimation problem of continuous-time stochastic regression models. The distributed order estimation algorithm (Algorithm 2) is designed by minimizing the piecewise continuous local information criterion based on the L_0 penalty term, and uti-

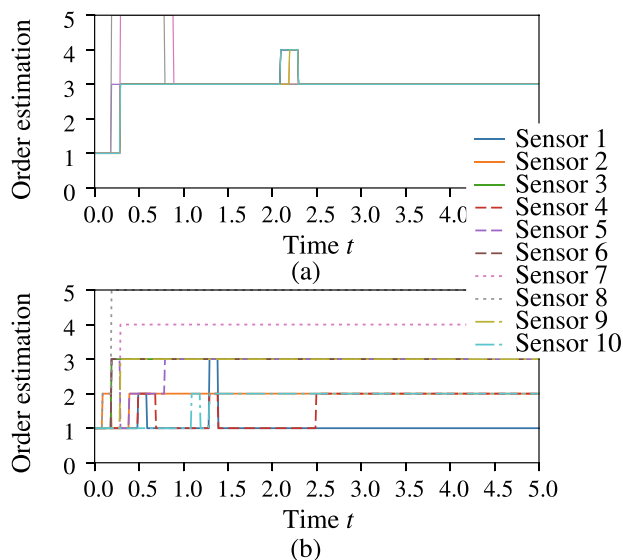


Fig. 2 Order estimates generated by (a) Algorithm 2 and (b) non-distributed algorithm

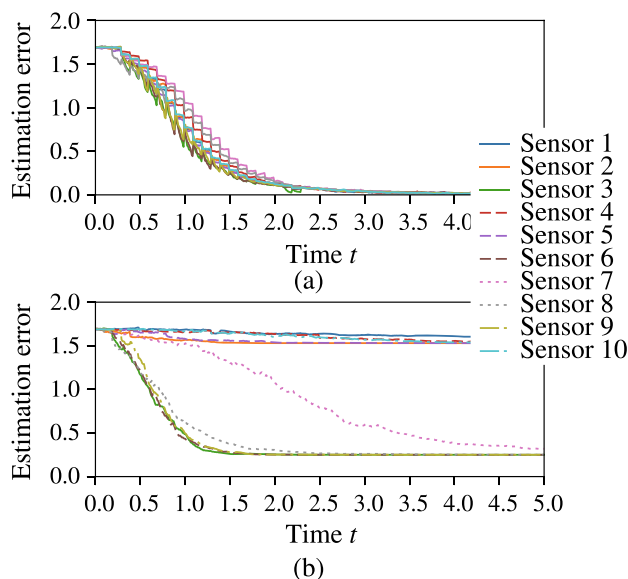


Fig. 3 Parameter estimation errors by (a) Algorithm 2 and (b) non-distributed order estimation algorithm

lizing the continuous-time distributed least squares algorithm based on the diffusion strategy. By introducing cooperative excitation condition and dealing with the complex coupling between estimates of system order and parameters, almost surely convergence results of the proposed algorithm are established. The convergence results in this paper is obtained without assuming regression vectors to be bounded or satisfy the persistent excitation conditions, which allows our results to be further used in the related studies of feedback control systems. Finally, the effectiveness of the proposed distributed algorithm for the unknown order and parameters is verified

by a simulation example. Some interesting problems such as the distributed order estimation problem with unknown upper bound, the distributed adaptive control problem, deserve to be further studied.

Funding Open access funding provided by Royal Institute of Technology.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Akyildiz, I. F., & Vuran, M. C. (2010). *Wireless sensor networks*. West Sussex: Wiley.
- Edward, F., Ehrlich, L. N., Bhatta, P., Paley, D. A., Bachmayer, R., & Fratantoni, D. M. (2006). Multi-AUV control and adaptive sampling in Monterey Bay. *IEEE Journal of Oceanic Engineering*, 31(4), 935–948.
- You, J., & Wu, W. (2017). Online passive identifier for spatially distributed systems using mobile sensor networks. *IEEE Transactions on Control Systems Technology*, 25(6), 2151–2159.
- Orlov, Y., & Bentsman, J. (2000). Adaptive distributed parameter systems identification with enforceable identifiability conditions and reduced-order spatial differentiation. *IEEE Transactions on Automatic Control*, 45(2), 203–216.
- Sayed, A. H. (2014). *Chapter 9—Diffusion adaptation over networks* (Vol. 3, pp. 323–453). Amsterdam: Elsevier.
- You, J., Zhang, Z., Zhang, F., & Wu, W. (2022). Cooperative filtering and parameter identification for advection–diffusion processes using a mobile sensor network. *IEEE Transactions on Control Systems Technology*, 31, 527–542.
- Sayed, A. H., Tu, S.-Y., Chen, J., Zhao, X., & Towfic, Z. J. (2013). Diffusion strategies for adaptation and learning over networks: An examination of distributed strategies and network behavior. *IEEE Signal Processing Magazine*, 30(3), 155–171.
- Di Lorenzo, P., & Sayed, A. H. (2012). Sparse distributed learning based on diffusion adaptation. *IEEE Transactions on Signal Processing*, 61(6), 1419–1433.
- Schizas, I. D., Mateos, G., & Giannakis, G. B. (2009). Distributed LMS for consensus-based in-network adaptive processing. *IEEE Transactions on Signal Processing*, 57(6), 2365–2382.
- Xie, S., Zhang, Y., & Guo, L. (2021). Convergence of a distributed least squares. *IEEE Transactions on Automatic Control*, 66(10), 4952–4959.
- Xie, S., & Guo, L. (2018). Analysis of distributed adaptive filters based on diffusion strategies over sensor networks. *IEEE Transactions on Automatic Control*, 63(11), 3643–3658.
- Xie, S., & Guo, L. (2018). Analysis of normalized least mean squares-based consensus adaptive filters under a general information condition. *SIAM Journal on Control and Optimization*, 56(5), 3404–3431.
- Gan, D., & Liu, Z. (2022). Distributed order estimation of ARX model under cooperative excitation condition. *SIAM Journal on Control and Optimization*, 60(3), 1519–1545.
- Zill, D. G. (2012). *A first course in differential equations with modeling applications*. Boston: Cengage Learning.
- Sun, J., Shen, Y., & Rosen, J. (2021). Sensor reduction, estimation, and control of an upper-limb exoskeleton. *IEEE Robotics and Automation Letters*, 6(2), 1012–1019.
- Todorov, E. (2005). Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation*, 17(5), 1084–1108.
- Khalil, H. K. (2002). *Nonlinear Systems*. Prentice-Hall.
- Brown, C. (2007). *Differential equations: A modeling approach* (Vol. 150). California: Sage.
- Cellier, F. E., & Greifeneder, J. (2013). *Continuous system modeling*. Berlin: Springer.
- Rao, G. P., & Unbehauen, H. (2006). Identification of continuous-time systems. *IEE Proceedings-Control Theory and Applications*, 153(2), 185–220.
- Chen, W., Wen, C., Hua, S., & Sun, C. (2013). Distributed cooperative adaptive identification and control for a group of continuous-time systems with a cooperative PE condition via consensus. *IEEE Transactions on Automatic Control*, 59(1), 91–106.
- Nascimento, V. H., & Sayed, A. H. (2012). Continuous-time distributed estimation. In: *Proceedings of the 45th Asilomar conference on signals, systems and computers, CA, USA*
- Abu-Naser, M., & Williamson, G. A. (2007). Convergence of adaptive estimators of time-varying linear systems using basis functions: Continuous time results. In *2007 IEEE international conference on acoustics, speech and signal processing* (pp. 1361–1364). IEEE.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716–723.
- Weakliem, D. L. (1999). A critique of the Bayesian information criterion for model selection. *Sociological Methods & Research*, 27, 359–397.
- Guo, L., Chen, H., & Zhang, J. (1989). Consistent order estimation for linear stochastic feedback control system (CARMA model). *Automatica*, 25, 147–151.
- Hannan, E. J., & Kavalieris, L. (1984). Multivariate linear time series models. *Advances in Applied Probability*, 16, 492–561.
- Ninness, B. (2004). On the CRLB for combined model and model-order estimation of stationary stochastic processes. *IEEE Signal Processing Letters*, 11(2), 293–296.
- Chen, H., & Guo, L. (1987). Consistent estimation of the order of stochastic control systems. *IEEE Transactions on Automatic Control*, 32, 531–535.
- Huang, D., & Guo, L. (1990). Estimation of nonstationary ARMAX models based on the Hannan–Rissanen method. *The Annals of Statistics*, 18, 1729–1756.
- Zhao, W., & Chen, H. (2010). New method of order estimation for ARMA/ARMAX processes. *SIAM Journal on Control and Optimization*, 48(2), 4157–4176.
- Abo-Hammour, Z. S., Alsmadi, O. M. K., Al-Smadi, A. M., Zaqout, M. I., & Sarairh, M. S. (2012). ARMA model order and parameter estimation using genetic algorithms. *Mathematical and Computer Modelling of Dynamical Systems*, 18, 201–221.
- Moon, J., Hossain, M. B., & Chon, K. H. (2021). AR and ARMA model order selection for time-series modeling with ImageNet classification. *Signal Processing*, 183, 108026.
- Victor, S., Malti, R., Garnier, H., & Oustaloup, A. (2013). Parameter and differentiation order estimation in fractional models. *Automatica*, 49(4), 926–935.
- Belkhatir, Z., & Laleg-Kirati, T. M. (2018). Parameters and fractional differentiation orders estimation for linear continuous-time non-commensurate fractional order systems. *Systems & Control Letters*, 115, 26–33.

36. Stark, O., Pfeifer, M., & Hohmann, S. (2021). Parameter and order identification of fractional systems with application to a lithium-ion battery. *Mathematics*, 9(14), 1607.
37. Godsil, C., & Royle, G. (2001). *Algebraic graph theory* (pp. 163–192). London: Springer.
38. Chen, H. F., & Guo, L. (1991). *Identification and stochastic adaptive control* (pp. 403–419). Springer.
39. Révész, P., Liptser, R. S., Shiriyayer, A. N., & Revesz, P. (1980). Statistics of random processes. *International Statistical Review*, 48(3), 371.
40. Zhu, X., & Liu, Z. (2023). Distributed least squares algorithm for continuous-time stochastic systems under cooperative excitation condition.
41. Fang, Z., & Miu, B. (2011). *Stochastic process* (3rd ed.). Beijing: Science Press.
42. Zhu, X. (2023). *Distributed identification theory for continuous-time stochastic systems*. PhD thesis, University of Chinese Academy of Sciences.
43. Xiao, L., Boyd, S., & Lall, S. (2005). A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th international symposium on information processing in sensor networks, Boise, ID, USA* (pp. 63–70).



Xinghua Zhu received the B.Sc. degree in Mathematics from Minzu University of China, Beijing, China, in 2018, and the Ph.D. degree in systems analysis and integration from Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Science (CAS), Beijing, China, in 2023. Her research interests include continuous-time distributed adaptive filter and networked systems.



Zhixin Liu received the B.Sc. degree in Mathematics from Shandong University, Jinan, China, in 2002, and the Ph.D. degree in control theory from Academy of Mathematics and Systems Science (AMSS), Chinese Academy of Sciences (CAS), Beijing, China, in 2007. She is currently a professor of AMSS, CAS. She had visiting positions at KTH Royal Institute of Technology, University of New South Wales, Canberra and University of Maryland, College Park. Her current research interests are complex systems and multi-

agent systems. She is a coauthor of the SIGEST paper in SIREV in 2014 and also a co-recipient of the Best Theory Paper Award at The 13rd World Congress on Intelligent Control and Automation (WCICA).



Xiaoming Hu received the B.Sc. degree from the University of Science and Technology of China in 1983, and the M.Sc. and Ph.D. degrees from the Arizona State University in 1986 and in 1989, respectively. He served as a research assistant at the Institute of Automation, the Chinese Academy of Sciences, from 1983 to 1984. From 1989 to 1990 he was a Gustafsson Postdoctoral Fellow at KTH Royal Institute of Technology, Stockholm, where he is currently a professor of Optimization and Systems Theory. His main

research interests are in nonlinear dynamical systems, nonlinear observation and filtering, and multi-agent systems.