Systems & Control Letters 203 (2025) 106147

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Privacy-preserving distributed adaptive estimation for non-stationary regression data☆

Shuning Chen [a,b], Die Gan [c], Siyu Xie [d], Jinhu Lü [e,*]

[a] State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[c] College of Artificial Intelligence, Nankai University, Tianjin 300350, China
[d] School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu 611731, China
[e] School of Automation Science and Electrical Engineering at Beihang University, Beijing 100191, China

ARTICLE INFO

Keywords:
Distributed adaptive estimation
Differential privacy
Stochastic regression model
Time-varying parameter

ABSTRACT

Distributed adaptive estimation techniques allow agents in multi-agent networks to cooperatively estimate system parameters, but directly sharing information among agents increases the risk of privacy breaches. In this paper, we consider the problem of estimating unknown time-varying parameters in a discrete-time stochastic regression model over multi-agent networks, with a focus on protecting data privacy. We propose a privacy-preserving distributed consensus-based normalized least mean square algorithm that protects the local information of agents by obfuscating the information exchanged. The proposed algorithm achieves rigorous differential privacy for sensitive information by incorporating persistent additive noise to the exchanged estimates. Furthermore, we analyze the stability of the proposed algorithm and establish the upper bound of the estimation error without assuming the independency or stationarity of the regression data. Some simulation results are presented to validate the effectiveness of our theoretical findings.

1. Introduction

1.1. Background

With the development of sensing and communication technologies, distributed estimation or filtering algorithms based on multi-agent networks have received widespread attention [1–4]. The agents within networks are often limited in processing and computing resources, but they can cooperatively accomplish global tasks through information interaction. Distributed parameter estimation not only offers a robust alternative to centralized methods that rely on fusion centers, but also makes efficient use of network resources by distributing computational and communication burdens among agents. In numerous theoretical studies concerning distributed parameter estimation, a widely recognized classical model is the discrete linear regression model given by

$$y_{k,i} = x_{k,i}^{\top}\xi_k + d_{k,i},\ k \geq 0, i \in \{1, 2, \dots, n\}, \tag{1}$$

where $y_{k,i} \in \mathbb{R}$ is the local output, $x_{k,i} \in \mathbb{R}^m$ is the regression vector, $d_{k,i} \in \mathbb{R}$ is the random system disturbance, all associated with agent i at time k, and $\xi_k \in \mathbb{R}^m$ is the unknown parameter to be estimated. Agents are designed to be interconnected to acquire and process the local information from neighbors to finish a common estimation task.

1.2. Related works

A lot of distributed algorithms have been designed for estimating time-invariant parameters [1,5–8] or time-varying parameters [9–12]. In order to investigate the theoretical performance of the algorithms, many contributions require the conditions about the regression data. Some works focus on the performance analysis of distributed parameter estimation algorithms using deterministic or even time-invariant regressors [1,5,9,13], while the stability of some other distributed adaptive filtering algorithms is established based on the independence or stationarity and ergodicity of stochastic regressors [6,10,14,15]. However, as noted in [16], complex dynamic systems with uncertainty often contain various feedback loops, and the properties of observed data are usually determined by complex dynamic equations. Therefore, these systems are far from satisfying the traditional statistical assumptions

☆ This work was supported in part by the Natural Science Foundation of Tianjin under Grant 24JCQNJC01930, the National Key Research and Development Program of China under Grant 2022YFB3305600 and Sichuan Science and Technology Program under Grant 2025ZNSFSC1511.

* Corresponding author.
E-mail addresses: chenshuning@amss.ac.cn (S. Chen), gandie@nankai.edu.cn (D. Gan), syxie@uestc.edu.cn (S. Xie), jhlu@iss.ac.cn (J. Lü).

https://doi.org/10.1016/j.sysconle.2025.106147
Received 22 January 2025; Received in revised form 2 April 2025; Accepted 13 May 2025
Available online 4 June 2025

of independence, stationarity, and ergodicity. To relax the stringent conditions on random regression vectors to non-stationary scenarios, some progress has also been made in distributed adaptive estimation and filtering algorithms [7,8,11,12].

It is worth noting that there are often attackers in the multi-agent network who use various available information in the network to infer underlying private data. However, in the aforementioned works, agents directly transmit their local estimates to neighbors in the public network, which can lead to significant privacy breaches when the target parameter vector contains sensitive information. For instance, as shown in [17], private ratings and transactions of individuals on commercial websites can be successfully deduced by leveraging information from public recommendation systems. This highlights the necessity for developing distributed estimation mechanisms that preserve privacy. A number of privacy-preserving algorithms have already been proposed in fields such as distributed optimization [18–20], distributed consensus [21,22], and federated learning [23–26], among which commonly used privacy protection techniques are homomorphic encryption and adding artificial noise, etc. Homomorphic encryption [18,21,25,27] of privacy data in the local computation process can highly protect privacy, but it introduces significant computational overhead, which is impractical for mobile devices [28]. In contrast, adding noise to the data transmitted between agents to protect privacy is simpler and more feasible [19,20,23,24,26]. Differential privacy (DP) is a mathematical concept that guarantees statistical indistinguishability for individual inputs by adding noises [29].

Integrating DP with distributed parameter estimation may provide strong privacy guarantees while maintaining high estimation performance. As far as we know, there are few results studying the theoretical performance analysis of the privacy-preserving distributed estimation algorithms [26,30–33]. In [26], the least-square procedure was applied for the federated estimation problem. However, this work was only concerned with the analysis of static databases, failing to work with dynamic, time-varying data streams. Le Ny and Papps in [30] integrated DP mechanisms with distributed Kalman filtering to ensure privacy protection performance and further considered the scenario with continual observation data, but the results were derived on deterministic regressors. Wang et al. [33] imposed persistent excitation conditions on regressors and utilized DP theory to perform privacy-preserving distributed estimation of time-invariant parameters. However, agents may need to perform estimation in a constantly changing environment, these works [26,32,33] will lose efficacy when dealing with dynamic systems where parameters change over time. Overall, the existing methods all have limitations on estimating time-varying parameters or assuming strong conditions for the regression data, and therefore cannot be widely applied to practical systems.

1.3. Challenges and contributions

In this paper, we investigate unknown time-varying parameter estimation for a discrete-time stochastic regression model over multi-agent networks. We add noise to the local estimates exchanged between agents to prevent potential attackers from recovering local sensitive data through the publicly exchanged estimates. Then, a privacy-preserving distributed adaptive estimation algorithm is proposed and the privacy analysis is guaranteed by DP theory. We introduce a cooperative excitation condition for non-stationary regressors and theoretically establish the upper bound of the tracking error in the case of time-varying parameters.

Notably, the impact of the additional noise introduced by privacy protection will accumulate over time, and coupled with the randomness brought by non-independent and non-stationary regressors, this brings significant challenges to theoretical analysis. We address these challenges by employing algebraic graph theory and stochastic stability theory to establish the stable properties of some auxiliary matrices (based on our previous work [11]), thereby further overcoming the technical difficulties. It is worth noting that our results are applicable to real-world systems, unlike the theoretical findings in [26,30,32,33], which either rely on assumptions of independence or stationarity of the regression signals, or focus solely on estimating time-invariant parameters.

The remainder of this paper is organized as follows: Section 2 details the design of our proposed algorithm, whose privacy analysis is presented in Section 3. Section 4 introduces essential definitions and assumptions. Section 5 presents the stability analysis of proposed algorithm. Experimental results are displayed in Section 6, followed by some concluding remarks in Section 7.

2. Problem formulation

2.1. Some preliminaries

2.1.1. Notations

For an $n \times m$-dimensional real matrix A, we use $\|A\|$ to represent the Euclidean norm, i.e., $\|A\| \triangleq \{\lambda_{\max}(AA^\top)\}^{\frac{1}{2}}$, where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of the matrix, and $(\cdot)^\top$ denotes the transpose of the matrix. Correspondingly, the smallest eigenvalue of the matrix is denoted as $\lambda_{\min}(\cdot)$. For an m-dimensional real vector x, the p-norm of x is defined as $\|x\|_p = (\sum_{i=1}^m |x_i|^p)^{1/p}$, with x_i being the ith element of x and $1 \le p < \infty$. If there is no special indication, $\|\cdot\|$ refers to 2-norm (also the Euclidean norm). The symbol $\mathbf{1}_n$ represents an n-dimensional column vector with all elements equal to 1, and I_m denotes the m-dimensional identity matrix.

For a matrix sequence $\{A_k, k \ge 0\}$ and a positive scalar sequence $\{a_k, k \ge 0\}$, if there exists a positive constant C, such that $\|A_k\| \le Ca_k$ holds for all $k \ge 0$, then we say $A_k = O(a_k)$.

2.1.2. Graph theory

For a multi-agent network, we can construct a corresponding topology $\mathcal{G} = \{\mathcal{V}, \mathcal{E}, \mathcal{A}\}$ to show the information interaction between agents. Take an n-agent network for instance, let the node set $\mathcal{V} = \{1, 2, \dots, n\}$. The elements in the adjacency matrix $\mathcal{A} = [a_{ij}]_{1 \le i,j \le n}$ represent the weight of information interaction between agents. If there exists communication from agent i to agent j, then the edge (i, j) belongs to the edge set $\mathcal{E}$ and $a_{ij} > 0$, otherwise $a_{ij} = 0$. The neighbor set of agent i is denoted as $\mathcal{N}_i = \{j | (j, i) \in \mathcal{E}\}$. In this paper, we suppose the adjacency matrix is symmetric and stochastic, i.e., $\sum_{j=1}^n a_{ij} = \sum_{j=1}^n a_{ji} = 1$ holds for any $i \in \{1, 2, \dots, n\}$.

2.2. Algorithm design

In this paper, we consider using the discrete-time stochastic regression model (1) to estimate the time-varying parameter ξ_k, whose variation at time k can be denoted as:

$$\gamma\omega_k \triangleq \xi_k - \xi_{k-1}, k \ge 1, \tag{2}$$

where γ is a non-negative scalar that characterizes the rate of parameter variation, and ω_k is an $m \times 1$-dimensional vector. In particular, when $\gamma = 0$, it degenerates to the time-invariant parameter case.

For the linear regression model (1), traditional distributed adaptive filtering algorithms [5,15], including Kalman filtering, least squares, and least mean squares methods, pose privacy risks by potentially exposing sensitive data $\{y_{k,i}\}$ during the process of estimating parameters $\{\xi_k\}$. This vulnerability arises because these methods typically require sharing true intermediate results or updated parameter estimates across network nodes, which may inadvertently reveal information about the underlying private data to potential attackers [30]. For example, in smart grids, $y_{k,i}$ may represent the real-time electricity consumption of households or enterprises; in cooperative guidance systems, $y_{k,i}$ may indicate the position of missiles or launch sites, both of which are sensitive data that should not be disclosed to attackers. To tackle this

Algorithm 1 K-step Distributed Privacy-Preserving Normalized Least Mean Squares Algorithm

Data: $\{y_{k,i}, x_{k,i}\}_{i=1}^{n}$, $k = 0, 1, ..., K$
Result: $\{\hat{\xi}_{k+1,i}\}_{i=1}^{n}$, $k = 0, 1, ..., K$
Set instant $k = 0$;
Initialize estimates $\hat{\xi}_{0,i} \in \mathbb{R}^m$ for each agent $i \in \{1, 2, \cdots, n\}$.
while $k \le K$ **do**
  for agent $i = 1$ to n **do**
    Step 1. Add Laplacian noise to local estimates:

$$\hat{\xi}_{k,i}^{\sharp} = \hat{\xi}_{k,i} + \eta_{k,i}, \tag{3}$$

    with $\eta_{k,i} \sim \mathrm{Lap}(0, \sigma, m, 1)$;
    Step 2. Transmit $\hat{\xi}_{k+1,i}^{\sharp}$ to all neighbors in $\mathcal{N}_i$.
    Step 3. Update estimates with local and neighbors' noised information $\hat{\xi}_{k,l}^{\sharp}$:

$$\hat{\xi}_{k+1,i} = \hat{\xi}_{k,i}^{\sharp} + \mu \left\{ \frac{x_{k,i}}{1 + \|x_{k,i}\|^2}(y_{k,i} - x_{k,i}^{\top}\hat{\xi}_{k,i}^{\sharp}) - \nu \sum_{l \in \mathcal{N}_i} a_{li}(\hat{\xi}_{k,i}^{\sharp} - \hat{\xi}_{k,l}^{\sharp}) \right\}, \tag{4}$$

    where $\mu \in (0, 1)$ is the step-size, $\nu \in (0, 1)$ is a weighting constant, and a_{li} is the lth row ith column element of adjacency matrix $\mathcal{A}$;
  $k = k + 1$;

problem, we propose a distributed privacy-preserving adaptive filtering algorithm (see Algorithm 1).

To safeguard the local privacy-sensitive data from potential leakage due to the exchange of estimates with neighbors, a Laplacian noise term is introduced to the estimates. Specifically, the noisy estimate $\hat{\xi}_{k,i}^{\sharp}$ is generated as:

$$\hat{\xi}_{k,i}^{\sharp} = \hat{\xi}_{k,i} + \mathrm{Lap}\,(0, \sigma, m, 1)\,,$$

with the notation $\mathrm{Lap}(0, \sigma, m, 1)$ represents an $m \times 1$-dimensional matrix with each element i.i.d. to Laplacian distribution with the mean value 0 and the scale parameter σ. Then, each agent transmits the perturbed estimates to its neighboring nodes. By sharing these inaccurate (noisy) estimates rather than the true values, the attack will fail to infer the sensitive data, as shown in Section 3. After that, each agent updates the local estimates for the next instant using local and neighbors' noisy estimates. Here we adopt a consensus-based normalized least mean squares algorithm (4), whose right-hand side can be regarded as being composed of two components. The first part is the usual (normalized) LMS update mechanism which aims to reduce the prediction error, while the second part tries to minimize the weighted distance between estimates of the agent i and its neighboring agents as explained in [10]. This collaborative approach leverages the perturbed information to refine future predictions while maintaining privacy protection.

2.3. Recursive error equation

Our distributed Privacy-Preserving (PP)-NLMS algorithm is designed to protect the privacy of the output data $\{Y_k\}_{k=0}^{K}$ as well as estimate the unknown time-varying parameter vector $\{\xi_k\}$. Consequently, balancing the level of privacy protection and the performance of parameter estimation is a critical issue.

In order to centrally measure the estimation error of each agent, we further introduce the following series of notations.

$$\begin{aligned} Y_k &= [y_{k,1}, y_{k,2}, \dots, y_{k,n}]^{\top} \in \mathbb{R}^n \\ D_k &= [d_{k,1}, \dots, d_{k,n}]^{\top} \in \mathbb{R}^n, \end{aligned} \tag{5}$$

$$\begin{aligned} X_k &= \mathrm{diag}\{x_{k,1}, \dots, x_{k,n}\} \in \mathbb{R}^{mn \times n}, \\ H_k &= \mathrm{col}\{\eta_{k,1}, \dots, \eta_{k,n}\} \in \mathbb{R}^{mn}, \\ \Omega_k &= \mathbf{1}_n \otimes \omega_k \in \mathbb{R}^{mn}, \\ L_k &= \mathrm{diag}\left\{ \frac{x_{k,1}}{1 + \|x_{k,1}\|^2}, \dots, \frac{x_{k,i}}{1 + \|x_{k,n}\|^2} \right\} \in \mathbb{R}^{mn \times n}, \\ \hat{\Xi}_k &= \mathrm{col}\{\hat{\xi}_{k,1}, \dots, \hat{\xi}_{k,n}\} \in \mathbb{R}^{mn}, \\ \hat{\Xi}_k^{\sharp} &= \mathrm{col}\{\hat{\xi}_{k,1}^{\sharp}, \dots, \hat{\xi}_{k,n}^{\sharp}\} \in \mathbb{R}^{mn}, \\ \Xi_k &= \mathbf{1}_n \otimes \xi_k \in \mathbb{R}^{mn}, \\ \widetilde{\Xi}_k &= \hat{\Xi}_k - \Xi_k \in \mathbb{R}^{mn}, \quad \widetilde{\Xi}_k^{\sharp} = \hat{\Xi}_k^{\sharp} - \Xi_k \in \mathbb{R}^{mn}, \\ \mathscr{L} &= (I_n - \mathcal{A}) \otimes I_m \in \mathbb{R}^{mn \times mn}, \\ G_k &= L_k X_k^{\top} + \nu \mathscr{L} \in \mathbb{R}^{mn \times mn}. \end{aligned} \tag{6}$$

Then, the following compact iteration can be derived by (3) and (4) that

$$\begin{cases} \widetilde{\Xi}_k^{\sharp} = \widetilde{\Xi}_k + H_k, \\ \hat{\Xi}_{k+1} = \hat{\Xi}_k^{\sharp} + \mu L_k (Y_k - X_k^{\top}\hat{\Xi}_k^{\sharp}) - \mu\nu\mathscr{L}\hat{\Xi}_k^{\sharp}. \end{cases}$$

Since $\mathscr{L}\Xi_k \equiv 0$, the estimation error $\widetilde{\Xi}_{k+1}$ can be obtained further by (1) and (2) that

$$\begin{aligned} \widetilde{\Xi}_{k+1} &= \widetilde{\Xi}_{k+1}^{\sharp} - H_{k+1} \\ &= \left(I_{mn} - \mu G_k\right) \widetilde{\Xi}_k^{\sharp} + \mu L_k D_k - \gamma \Omega_{k+1} \\ &= \left(I_{mn} - \mu G_k\right) (\widetilde{\Xi}_k + H_k) + \mu L_k D_k - \gamma \Omega_{k+1}. \end{aligned} \tag{7}$$

In the next section, we will first analyze the privacy protection performance of the distributed PP-NLMS algorithm.

3. Privacy analysis

Before evaluating the privacy protection level of our algorithm, it is necessary to introduce some key concepts of differential privacy. Differential privacy provides a mathematical framework designed to protect individual privacy during data analysis and sharing processes. It ensures that the output of an algorithm remains statistically similar whether or not any specific individual's data is included.

Definition 1 (δ-*adjacency*). For any given $\delta > 0$, two vectors $Y_k = [y_{k,1}, y_{k,2}, \dots, y_{k,n}]^{\top}$ and $Y_k' = [y_{k,1}', y_{k,2}', \dots, y_{k,n}']^{\top}$ are called δ-adjacent if there exists some $i_0 \in \mathcal{V}$ such that for any $k \ge 0$,

$$y_{k,i} = y_{k,i}', \forall i \ne i_0, \quad \|y_{k,i_0} - y_{k,i_0}'\|_1 \le \delta. \tag{8}$$

It suggests that two signal sequences are considered "adjacent" if the measurement of only one agent is changed.

Definition 2 (*Sensitivity [30]*). For the given δ-adjacent relation (8), define the sensitivity of an output map $g(\cdot)$ at kth iteration as

$$S(k) = \sup_{Y_k, Y_k' \text{ are } \delta-\text{adjacent}} \left\| g\left(Y_k\right) - g\left(Y_k'\right) \right\|_1.$$

Remark 1. The sensitivity of an output map $g(\cdot)$ measures the maximum possible change in the output of $g(\cdot)$ when the data of a single agent is altered in Y_k. In this paper, $g(\cdot)$ corresponds to the recursive estimation Eq. (4), which can be represented as:

$$g(Y_k) = \hat{\Xi}_{k+1} = (I - \mu G_k)\hat{\Xi}_k^{\sharp} + \mu L_k Y_k.$$

Lemma 1. *For given δ-adjacency relation (8), the sensitivity of the distributed NLMS algorithm at each iteration k satisfies*

$$S(k) \le \mu\delta.$$

Proof. By Remark 1, we can get $g\left(Y_k\right)-g\left(Y_k'\right)=\mu L_k\left(Y_k-Y_k'\right)$. Notice the fact that $\|L_k\|\leq 1$, thus for two δ-adjacency vectors, it follows

$$\left\|\mu L_k\left(Y_k-Y_k'\right)\right\|_1\leq\mu\left\|Y_k-Y_k'\right\|_1\leq\mu\delta,$$

which completes the proof. ■

Assume the attackers can observe all communication information between nodes, i.e. $\{\hat{\Xi}_k^\sharp\}_{k=1}^K$, and based on this, they aim to infer the privacy data $\{Y_k\}_{k=0}^{K-1}$. From Eqs. (3) and (4), this antagonistic mechanism at instant k can essentially be described by the mapping $\mathcal{M}^{(k)}:\mathbb{R}^n\times\mathbb{R}^{nm}\mapsto\mathbb{R}^{nm}$ with

$$\begin{aligned}\mathcal{M}^{(k)}\left(Y_k,H_{k+1}\right)&\triangleq g(Y_k)+H_{k+1}\\&=(I-\mu G_k)\hat{\Xi}_k^\sharp+\mu L_kY_k+H_{k+1}.\end{aligned}\tag{9}$$

The efficacy of this privacy protection is critically dependent on the characteristics of the added Laplace noise H_k, particularly its scale parameter σ. Then, the differential privacy of Algorithm 1 is defined as follows.

Definition 3 (*Differential Privacy [33]*). Given $\varepsilon>0$, a randomized mechanism $\mathcal{M}^{(k)}$ is called ε-differential privacy if for any δ-adjacent vectors Y_k and Y_k', any stochastic noise H_{k+1}, and any set of outputs $\mathcal{Y}\subseteq\text{Range}(\mathcal{M}^{(k)})$,

$$\Pr\left\{\mathcal{M}^{(k)}\left(Y_k,H_{k+1}\right)\in\mathcal{Y}\right\}\leq e^\varepsilon\Pr\left\{\mathcal{M}^{(k)}\left(Y_k',H_{k+1}\right)\in\mathcal{Y}\right\}$$

holds. Here, Range$(\mathcal{M}^{(k)})$ denotes the range of the mapping $\mathcal{M}^{(k)}$, and $\Pr\{\cdot\}$ denotes the probability.

Remark 2. Note that the constant ε measures the privacy level of the randomized mapping $\mathcal{M}^{(k)}$; specifically, a smaller ε indicates a higher privacy level. As mentioned in [30], ε is typically chosen to be a small constant, such as $\varepsilon=0.1$ or perhaps even $\ln 2$ or $\ln 3$.

Theorem 1. *For given $\varepsilon>0$, each iteration of Algorithm 1 satisfies ε-differential privacy if*

$$\sigma\geq\frac{\mu\delta}{\varepsilon}.\tag{10}$$

Proof. Here, we analyze the impact of H_{k+1} on the output distribution when the sensitive data Y_k changes to a δ-adjacent Y_k', while keeping all other random variables fixed. For two δ-adjacent vectors Y_k and Y_k', note that $\mathcal{M}^{(k)}\left(Y_k,H_{k+1}\right)$ and $\mathcal{M}^{(k)}\left(Y_k',H_{k+1}\right)$ share the same range, where each element of H_{k+1} is an independent random variable following a Laplacian distribution with mean 0 and scale parameter σ. Considering fixed Y_k, Y_k', and the terms constructed from the regressors X_k, and denoting the probability density function of H_{k+1} as $f(\cdot)$, we have:

$$\begin{aligned}&\frac{f\left(\mathcal{M}^{(k)}\left(Y_k,H_{k+1}\right)=Z\right)}{f\left(\mathcal{M}^{(k)}\left(Y_k',H_{k+1}\right)=Z\right)}\\=&\frac{f\left(H_{k+1}=Z-(I-\mu G_k)\hat{\Xi}_k^\sharp-\mu L_kY_k\right)}{f\left(H_{k+1}=Z-(I-\mu G_k)\hat{\Xi}_k^\sharp-\mu L_kY_k'\right)}\\=&\frac{\exp\left(-\frac{1}{\sigma}\left\|Z-(I-\mu G_k)\hat{\Xi}_k^\sharp-\mu L_kY_k\right\|_1\right)}{\exp\left(-\frac{1}{\sigma}\left\|Z-(I-\mu G_k)\hat{\Xi}_k^\sharp-\mu L_kY_k'\right\|_1\right)}\\\leq&\exp\left(\frac{1}{\sigma}S(k)\right)\leq\exp\left(\frac{\mu\delta}{\sigma}\right).\end{aligned}$$

Thus, for any measurable set of $\mathcal{Y}\subseteq Range(\mathcal{M}^{(k)}(Y_k,H_{k+1}))$, it holds

$$\begin{aligned}&\Pr\{\mathcal{M}^{(k)}(Y_k,H_{k+1})\in\mathcal{Y}\}=\int_{\mathcal{Y}}f\left(\mathcal{M}^{(k)}(Y_k,H_{k+1})=Z\right)dZ\\&\leq\exp\left(\frac{\mu\delta}{\sigma}\right)\int_{\mathcal{Y}}f\left(\mathcal{M}^{(k)}(Y_k',H_{k+1})=Z\right)dZ\\&\leq e^\varepsilon\Pr\{\mathcal{M}^{(k)}(Y_k',H_{k+1})\in\mathcal{Y}\},\end{aligned}$$

which means each iteration of Algorithm 1 satisfies ε-differential privacy. ■

After obtaining the differential privacy of each iteration, we introduce the following lemma to get the differential privacy of the entire K-step algorithm.

Lemma 2 (*See Theorem 3.14 in [34]*). *Consider a sequence of mechanisms $\{\mathcal{M}^{(k)}\}_{k=1}^K$ that all preserve ε-differential privacy. If $\mathcal{M}_K(Y)$ is designed to be $\mathcal{M}_K(Y)=(\mathcal{M}^{(0)}(Y),\ldots,\mathcal{M}^{(K-1)}(Y))$, then $\mathcal{M}_K(Y)$ is $K\varepsilon$-differentially private.*

As a result of the above lemma combining Theorem 1, the following corollary can be easily derived.

Corollary 1. *For given $\varepsilon>0$, the K-step Algorithm 1 satisfies $K\varepsilon$-differential privacy if $\sigma\geq\frac{\mu\delta}{\varepsilon}$.*

In Theorem 1 and Corollary 1, we give the relationship between the scale parameter σ of the added noise, the step-size μ and the privacy index ε. From (10) it follows that the scalar parameter σ of the added noise is inversely proportional to privacy index ε. In other words, each agent preserves stronger privacy when the added noise is more dispersive. However, this can have a detrimental effect on parameter estimation, thus necessitating further stability analysis of the proposed algorithm.

4. Definitions and assumptions

Due to the randomness of regression vector $\{x_{t,i},t\geq 0\}_{i=1}^n$, we first give some necessary definitions and assumptions on random matrices before going straight to the performance discussion.

4.1. Definitions

Definition 4 (*[35]*). A random matrix sequence $\{A_t,t\geq 0\}$ defined on the basic probability space $(\Omega,\mathcal{F},P)$ is called L_p-stable ($p>0$) if $\sup_{t\geq 0}\mathbb{E}(\|A_t\|^p)<\infty$. We define $\|A_t\|_{L_p}\triangleq[\mathbb{E}(\|A_t\|^p)]^{\frac{1}{p}}$ as the L_p-norm of the random matrix A_t, where $\mathbb{E}(\cdot)$ denotes the expectation operator.

Definition 5 (*[11]*). For a sequence of $n\times n$ random matrices $A=\{A_t,t\geq 0\}$, if it belongs to the following set

$$\begin{aligned}S_p(\alpha)=\Big\{A:\Big\|\prod_{j=k+1}^t(I_n-A_j)\Big\|_{L_p}&\leq C\alpha^{t-k},\\\forall t\geq k,\forall k\geq 0,&\text{ for some }C>0\Big\},\end{aligned}$$

then $\{I-A_t,t\geq 0\}$ is called L_p-exponentially stable ($p\geq 0$) with parameter $\alpha\in[0,1)$.

For convenience, we introduce the following subclass of $S_1(\alpha)$ for a scalar sequence $b=\{b_t,t\geq 0\}$:

$$\begin{aligned}S^0(\alpha)=\Big\{b:b_t\in[0,1),\mathbb{E}\left(\prod_{j=k+1}^t(1-b_j)\right)&\leq C\alpha^{t-k},\\\forall t\geq k,\forall k\geq 0,&\text{ for some }C>0\Big\}.\end{aligned}$$

Definition 6 (*[35]*). A sequence $\zeta=\{\zeta_k\}$ is considered an element of the weakly dependent set $\mathcal{W}_p$ (where $p\geq 1$) if there exists a constant C_p^ζ, which depends sorely on p and the distribution of $\{\zeta_k\}$, satisfying the condition that for any $k\geq 0$ and $l\geq 1$,

$$\left\|\sum_{i=k+1}^{k+l}\zeta_i\right\|_{L_p}\leq C_p^\zeta\sqrt{l}.$$

Remark 3. It is known that the martingale differences, sequences with zero mean that are ϕ- or α-mixing, and the linear process driven by white noises, are all belong to the set $\mathcal{W}_p$ [35].

4.2. Assumptions

For the stability and performance analysis, we make the following assumptions about the regression vector, the agent network topology, the measurement noise, and the variation of parameters to be estimated.

Assumption 1 (*Cooperative Excitation Condition*)**.** For the adapted sequences $\{x_{t,i}, \mathcal{F}_t, t \geq 0\}$, where $\mathcal{F}_t$ is a sequence of non-decreasing σ-algebras, there exists an integer $l > 0$ such that $\{\alpha_t, t \geq 0\} \in S^0(\alpha)$ for some $\alpha \in (0,1)$, where α_t is defined by

$$\alpha_t \triangleq \lambda_{\min}\left[\mathbb{E}\left(\frac{1}{n(1+l)}\sum_{i=1}^{n}\sum_{k=t+1}^{t+l}\frac{x_{k,i}x_{k,i}^{\top}}{1+\|x_{k,i}\|^2}\middle|\mathcal{F}_t\right)\right]. \tag{11}$$

with $\mathbb{E}(\cdot|\cdot)$ being the conditional expectation operator.

Remark 4. Assumption 1 is a spatial–temporal joint excitation condition, which is less stringent than conditions imposed solely on individual agents or merely on the regression vectors of agents at a single point in time. Intuitively, this assumption implies that over a period of time, the aggregate information from all agents must not be "too close to zero", or it will fail to achieve the estimation task.

Assumption 2. The graph $\mathcal{G}$ of the multi-agent network is undirected and connected.

Assumption 3. For some $p \geq 1$, $\|\widetilde{\Xi}_0\|_{L_{2p}} < \infty$, $\{L_k D_k\} \in \mathcal{W}_{2p}$ and $\{\Omega_k\} \in \mathcal{W}_{2p}$.

Remark 5. Assumption 3 posits that the noise and parameter variations are weakly correlated in some bounded moments sense, enabling the establishment of a finer bound in Theorem 3 compared to Theorem 2. In contrast, weaker assumptions (specifically, moment conditions) are imposed on D_k and Ω_k in Theorem 2.

We are now ready to present the main estimation performance findings regarding the proposed algorithm.

5. Stability analysis

In this section, we aim to investigate the magnitude of the estimation error after introducing Laplacian noise for privacy protection. Before presenting our theorem, we first introduce two lemmas that will assist in the proof.

Lemma 3 (*See Lemma 4.1 in [36]*)**.** *Let $\{b_{ki}, k \geq i \geq 0\}$, $\{c_{ki}, k \geq i \geq 0\}$, and $\{\zeta_k, k \geq 0\}$ be three nonnegative processes satisfying:*

(i) *$b_{nk} \in [0,1]$, $\mathbb{E}b_{nk} \leq c_1\alpha^{n-k}$, for all $n \geq k \geq 0$, for some $c_1 > 0$ and $\alpha \in [0,1)$;*

(ii) *There exist some constants $c_2 > 0$ and $c_3 > 0$ such that $\sup_{n\geq k\geq 0}\mathbb{E}[exp(c_2 c_{nk}^{1/c_3})] < \infty$;*

(iii) *$\chi_p \triangleq \sup_k \|\zeta_k \log^{\beta}(e+\zeta_k)\|_{L_p} < \infty$, for some $p \geq 1, \beta > 1$.*

Then, there exists a constant $c_4 > 0$ which is independent of χ_p such that

$$\sum_{k=0}^{n}\|b_{nk}c_{nk}\zeta_k\|_{L_p} \leq c_4\chi_p\log(e+\chi_p^{-1}), \forall n \geq 0$$

holds if $\{c_{nk}\}$ is deterministic.

Lemma 4 (*By Corollary 5.5 and Lemma 5.6 in [11]*)**.** *Consider the estimation error equation (7). Suppose that Assumptions 1 and 2 are satisfied. For any $\mu \in (0,1)$ and $\nu \in (0,1)$ satisfying $\mu(1+2\nu) \leq 1$, then $\{I - \mu G_k, k \geq 1\}$ is L_p-exponentially stable ($\forall p \geq 1$). Specifically, $\{\mu G_k\} \in S_p(\rho^{\lambda(p)})$, where $\rho = \alpha^{\frac{\lambda_2\nu\mu}{(1+l)(2+\nu\lambda_2)}}$, with α being defined in Assumption 1, λ_2 is the second smallest eigenvalue of matrix $I - \mathcal{A}$ and*

$$\lambda(p) = \begin{cases} \dfrac{1}{8l(l+1)^2}, & 1 \leq p \leq 2, \\ \dfrac{1}{4l(l+1)^2 p}, & p > 2. \end{cases}$$

Recalling the estimation error recursion equation (7) in Section 2.3, the exponential stability of its homogeneous portion is guaranteed by Lemma 4. With this foundation, we can derive a preliminary upper bound on the estimation error of the proposed algorithm.

Theorem 2. *Consider the estimation error equation (7). Under Assumptions 1 and 2, if for some $p \geq 1$ and $\beta > 1$, $\chi_p \triangleq \sup_k \|\zeta_k \log^{\beta}(e+\zeta_k)\|_{L_p} < \infty$, $\|\widetilde{\Xi}_0\|_{L_p} < \infty$ hold, where $\zeta_k \triangleq \|D_k\| + \|\Omega_{k+1}\| + \|H_k\|$. Then for any $\mu \in (0,1)$ and $\nu \in (0,1)$ satisfying $\mu(1+2\nu) \leq 1$, $\{\widetilde{\Xi}_k, k \geq 1\}$ is L_p-stable and*

$$\|\widetilde{\Xi}_k\|_{L_p} = O\left(\left(\rho^{\lambda(p)}\right)^k + \chi_p\log(e+\chi_p^{-1})\right), \forall k \geq 1,$$

where $\rho, \lambda(p)$ are defined in Lemma 4.

Proof. From (7) we can recursively obtain that

$$\begin{aligned}\widetilde{\Xi}_{k+1} &= \left(I_{mn} - \mu G_k\right)\left(\widetilde{\Xi}_k + H_k\right) + \mu L_k D_k - \gamma\Omega_{k+1} \\ &= \prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\widetilde{\Xi}_0 + \sum_{i=0}^{k}\left(\prod_{j=i+1}^{k}\left(I_{mn} - \mu G_j\right)\right) \\ &\quad \cdot\left[\left(I_{mn} - \mu G_i\right)H_i + \mu L_i D_i - \gamma\Omega_{i+1}\right].\end{aligned}$$

Then, we have

$$\begin{aligned}\left\|\widetilde{\Xi}_{k+1}\right\|_{L_p} &\leq \left\|\prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\widetilde{\Xi}_0\right\|_{L_p} \\ &+ \sum_{i=0}^{k}\left\|\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\left[\left(I - \mu G_i\right)H_i + \mu L_i D_i - \gamma\Omega_{i+1}\right]\right\|_{L_p} \\ &\leq \left\|\prod_{i=0}^{k}\left(I - \mu G_i\right)\widetilde{\Xi}_0\right\|_{L_p} + \sum_{i=0}^{k}\left\{\mathbb{E}\left(\left\|\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\right\|^p\right.\right. \\ &\quad \left.\left.\cdot\left(\|(I - \mu G_i)H_i\| + \mu\|L_i D_i\| + \gamma\|\Omega_{i+1}\|\right)^p\right)\right\}^{1/p}.\end{aligned} \tag{12}$$

Let $b_{ki} \triangleq \left\|\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\right\|$, then from Lemma 4 we can get that there exists a constant $C > 0$ such that

$$\left(\mathbb{E}\|b_{ki}\|^p\right)^{1/p} \leq C\left(\rho^{\lambda(p)}\right)^{k-i}, \tag{13}$$

where constants $\rho \in [0,1)$ and $\lambda(p) > 0$ are the same as in Lemma 4.

Combining the facts $\|L_i\| \leq 1$ and $\|I - \mu G_i\| \leq 1$ with (13), it can be finally obtained by Lemmas 3 and 4 that

$$\begin{aligned}\left\|\widetilde{\Xi}_{k+1}\right\|_{L_p} &\leq \left\|\prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\widetilde{\Xi}_0\right\|_{L_p} + b_0\sum_{i=0}^{k}\|b_{ki}\zeta_k\|_{L_p} \\ &\leq \left\|\prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\widetilde{\Xi}_0\right\|_{L_p} + b[\chi_p\log(e+\chi_p^{-1})] \\ &= O\left(\left(\rho^{\lambda(p)}\right)^k + \chi_p\log(e+\chi_p^{-1})\right),\end{aligned}$$

which completes the proof. ■

Remark 6. Intuitively, by Theorem 2 we know that when the parameter variation Ω_{k+1}, the measurement noise D_k and the Laplacian noise H_k are all small, the process ζ_k and χ_p will also be small, and hence the parameter tracking error $\widetilde{\Xi}_k$ will be small too.

However, to ensure that the algorithm possesses privacy-preserving capabilities, the Laplacian noise cannot be made arbitrarily small,

necessitating a trade-off between estimation performance and privacy protection.

By leveraging additional Assumption 3 on the system noise D_k and the parameter variation Ω_k, we can further construct a more precise upper bound on the tracking error $\widetilde{\Xi}_{k+1}$, that explicitly contains parameters such as the update step-size μ, parameter change rate γ, and Laplacian noise scale parameter σ. We also begin by presenting a few auxiliary lemmas.

Lemma 5. *For any given $p \geq 2$, the augmented Laplacian noise sequence $\{H_k\}$ in (6) is L_p-stable.*

Proof. Notice each element of H_k is i.i.d. as a Laplace random variable with mean 0 and variance $2\sigma^2$, whose probability density function is given by

$$f(x) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right).$$

Denote the ith element of H_k as $h_{k,i} (1 \leq i \leq mn)$. Then, we can obtain the pth moment of $|h_{k,i}|$ as

$$\begin{aligned}\mathbb{E}|h_{k,i}|^p &= \int_{-\infty}^{+\infty} |x|^p \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right) \mathrm{d}x \\ &= \frac{1}{\sigma} \int_0^{+\infty} x^p \exp\left(-\frac{x}{\sigma}\right) \mathrm{d}x \\ &\overset{y=x/\sigma}{=\!=\!=} \sigma^p \int_0^{+\infty} y^p \exp(-y)\, \mathrm{d}y \\ &= \sigma^p \Gamma(p+1),\end{aligned}$$

where $\Gamma(\cdot)$ refers to the Gamma function.

It can be derived by C_r-inequality that for any $p \geq 2$,

$$\begin{aligned}&\sup_k \|H_k\|_{L_p} = \sup_k \left(\mathbb{E}\|H_k\|^p\right)^{\frac{1}{p}} \\ =&\sup_k \left[\mathbb{E}\left(h_{k,1}^2 + \cdots + h_{k,mn}^2\right)^{\frac{p}{2}}\right]^{\frac{1}{p}} \\ \leq&\sup_k \left\{\mathbb{E}\left[(mn)^{\frac{p}{2}-1}\left(|h_{k,1}|^p + \cdots + |h_{k,mn}|^p\right)\right]\right\}^{\frac{1}{p}} \\ =&(mn)^{\frac{1}{2}} \sup_k \left(\mathbb{E}|h_{k,1}|^p\right)^{\frac{1}{p}} = \sqrt{mn}\sigma\left(\Gamma(p+1)\right)^{\frac{1}{p}} < \infty.\end{aligned}$$

Furthermore, we can know from Stirling's Approximation as

$$\Gamma(p+1) \approx \sqrt{2\pi}(p+1)^{p+\frac{1}{2}} e^{-p-1} \quad (p \to \infty),$$

so we can get an approximate bound of $\|H_k\|_{L_p}$ that

$$\begin{aligned}\|H_k\|_{L_p} &\approx \sqrt{mn}\sigma(2\pi(p+1)e^{-2})^{\frac{1}{2p}} e^{-1}(p+1) \\ &\approx \sqrt{mn}\sigma e^{-1}(p+1) \quad (p \to \infty),\end{aligned}$$

which completes the proof. ■

Lemma 5 analyzes the moment properties of Laplacian noise. Once p is determined, its L_p norm is proportional to σ. Next, we introduce two lemmas for dealing with the product of random matrices and the summation of these products.

Lemma 6 (*See Lemma 5.7 in [11]*). *Suppose that Assumptions 1 and 2 are satisfied. Then for any $p \geq 2$, any $\mu \in (0,1)$ and $\nu \in (0,1)$ satisfying $\mu(1+2\nu) \leq 1$ and $\forall k \geq i+1 > 0$, there exists positive constants C_p and ρ_p depending on $\{G_j, j > 0\}$ and p such that*

$$\left\|\prod_{j=i+1}^{k} (I - \mu G_j)\right\|_{L_p} \leq C_p (1 - \mu\rho_p)^{k-i}.$$

I.e., $\{\mu G_k\} \in S_p(1 - \mu\rho_p)$. In fact, we can take ρ_p as $1 - \rho^{\frac{\lambda(p)}{\mu}}$ corresponding to Lemma 4 (cf. [11]).

Lemma 7 (*See Lemma A.2 in [35]*). *Let $\{e_k\} \in \mathcal{W}_{2p}, p \geq 1$, if $\{\mu G_k\} \in S_{4p}(1 - \mu\rho_{4p})$ and $\sup_k \|G_k\|_{L_{4p}} < \infty$, then*

$$\left\|\sum_{i=0}^{k}\prod_{j=i+1}^{k}(I - \mu G_j)e_i\right\|_{L_p} = O(\mu^{-1/2}), \quad \forall \mu \in (0, 1 - \rho_{4p}].$$

Then, we give a more detailed estimate error bound explicitly with specific parameters.

Theorem 3. *Assume that Assumptions 1–3 are satisfied, and let the scale parameter of Laplacian noise $\sigma \geq \mu\delta/\varepsilon$, then for any $\mu \in (0,1)$ and $\nu \in (0,1)$ satisfying $\mu(1+2\nu) \leq 1$, we have for some p and $\forall k \geq 0$,*

$$\|\widetilde{\Xi}_{k+1}\|_{L_p} = O\left(\sqrt{\mu} + \frac{\gamma}{\sqrt{\mu}} + (1 - \rho_{2p}\mu)\frac{\sigma}{\mu} + (1 - \rho_{2p}\mu)^{k+1}\right),$$

where $\rho_{2p} \in (0,1)$ is a constant which is defined in the proof.

Proof. Using recursive relationship in (12), we first get

$$\begin{aligned}\left\|\widetilde{\Xi}_{k+1}\right\|_{L_p} \leq& \left\|\prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\widetilde{\Xi}_0\right\|_{L_p} \\ &+ \left\|\sum_{i=0}^{k}\prod_{j=i}^{k}\left(I - \mu G_j\right)H_i\right\|_{L_p} + \left\|\sum_{i=0}^{k}\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\mu L_i D_i\right\|_{L_p} \\ &+ \left\|\sum_{i=0}^{k}\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\gamma\Omega_{i+1}\right\|_{L_p}.\end{aligned}$$

Then, we will analyze each component on the right-hand side of the above inequality in turn.

For the first term, it can be naturally obtained by Assumption 3 that

$$\begin{aligned}\left\|\prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\widetilde{\Xi}_0\right\|_{L_p} &\leq \left\|\prod_{i=0}^{k}\left(I_{mn} - \mu G_i\right)\right\|_{L_{2p}} \cdot \left\|\widetilde{\Xi}_0\right\|_{L_{2p}} \\ &\leq O\left((1 - \mu\rho_{2p})^{k+1}\right). \qquad (14)\end{aligned}$$

From Lemma 5 we have

$$\begin{aligned}&\left\|\sum_{i=0}^{k}\prod_{j=i}^{k}\left(I - \mu G_j\right)H_i\right\|_{L_p} \leq \sum_{i=0}^{k}\left\|\prod_{j=i}^{k}\left(I - \mu G_j\right)H_i\right\|_{L_p} \\ \leq&\sum_{i=0}^{k}\left\|\prod_{j=i}^{k}\left(I - \mu G_j\right)\right\|_{L_{2p}} \cdot \|H_i\|_{L_{2p}} \\ \leq&\sqrt{mn}\sigma(\Gamma(2p+1))^{\frac{1}{2p}}\sum_{i=0}^{k} C_{2p}(1 - \mu\rho_{2p})^{k-i+1} \\ =&\sqrt{mn}\sigma(\Gamma(2p+1))^{\frac{1}{2p}} C_{2p}\frac{(1 - \mu\rho_{2p})[1 - (1 - \mu\rho_{2p})^{k+1}]}{\mu\rho_{2p}} \\ =&O\left(\frac{(1 - \mu\rho_{2p})\sigma}{\mu}\right). \qquad (15)\end{aligned}$$

From Lemma 6 we can see $\{\mu G_k\} \in S_{4p}(1 - \mu\rho_{4p})$ holds. Further combining Assumption 3 and the fact that $\|G_k\| \leq 1 + 2\nu < \infty$ with Lemma 7, it can be obtained that

$$\left\|\sum_{i=0}^{k}\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\mu L_i D_i\right\|_{L_p} = O(\mu^{1/2}), \qquad (16)$$

$$\left\|\sum_{i=0}^{k}\prod_{j=i+1}^{k}\left(I - \mu G_j\right)\gamma\Omega_{i+1}\right\|_{L_p} = O(\gamma\mu^{-1/2}). \qquad (17)$$

Combining (14) to (17), the result can be concluded. ■

Remark 7. Given that σ can be taken as $\mu\delta/\varepsilon$, the bound in Theorem 3 can be rewritten as $\sqrt{\mu} + \frac{\gamma}{\sqrt{\mu}} + (1 - \rho_{2p}\mu)\frac{\delta}{\varepsilon} + (1 - \rho_{2p}\mu)^{k+1}$. From this, it is easy to see that as ε decreases, the privacy protection performance

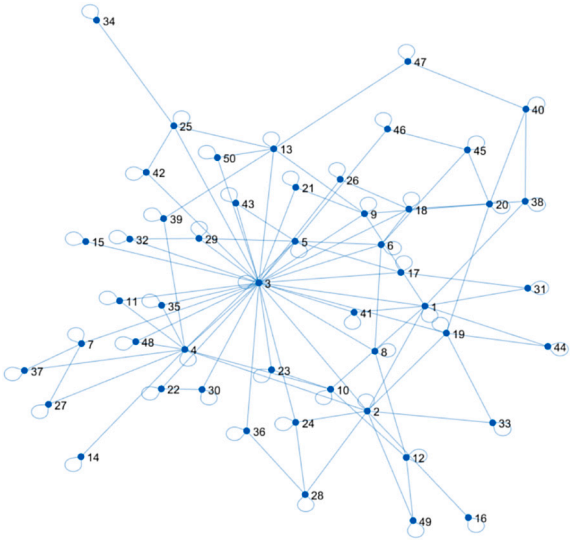


Fig. 1. The scale-free network topology of 50-agent network.

improves, but this leads to an increase in the estimation bound. On the other hand, to reduce the error caused by the time-varying nature of the parameters to be estimated, the update step size μ can be increased. However, this will amplify the effect of the inherent system noise. Therefore, the trade-offs between privacy protection and estimation accuracy, as well as between tracking ability and noise sensitivity, require careful consideration.

From Theorems 2 and 3, it can be seen that the stability results of the distributed PP-NLMS algorithm do not depend on the independence or stationarity conditions of the regression vector $\{X_k\}$. Thus our algorithm can be applied to stochastic feedback systems.

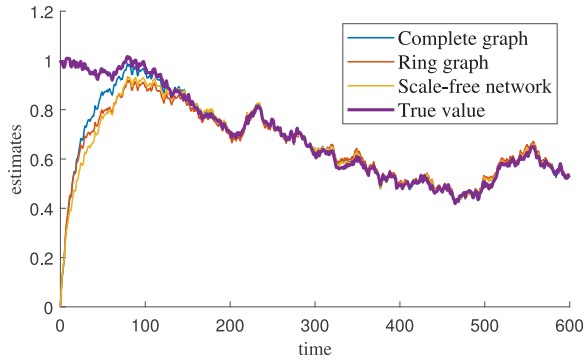
6. Simulation results

To validate the trade-off between privacy protection and parameter estimation performance of our proposed algorithm, we consider a network composed of 50 agents. The common estimation target for these agents is a time-varying 3-dimensional parameter vector ξ_k with initial value $\xi_0 = [1, 0, -1]^\top$. The variation of each element in this parameter vector is designed to follow a Gaussian distribution $N(0, 1^2)$, with the weighted parameter γ in (2) set to 0.01. The system measurement noise $\{d_{k,i}\}$ is designed to be spatially and temporally independent and identically distributed (i.i.d.) according to a Gaussian distribution $N(0, 0.1^2)$. Additionally, it is assumed that the regressors $\{x_{k,i}\}$ are generated in the following form:

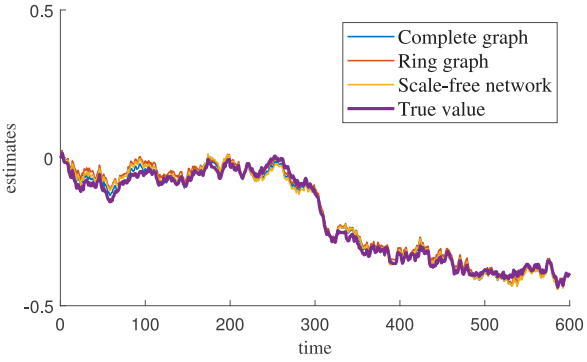
$$x_{k,i} = \begin{cases} \left[0.9^t + \sum_{j=0}^{t-1} 0.9^j \cos\left(\frac{i\pi}{n}\right) v_{k,j}, 0, 0\right]^\top, & \text{if } i \equiv 1 (\text{mod } 3) \\ \left[0, 0.9^t + \sum_{j=0}^{t-1} 0.9^j \cos\left(\frac{i\pi}{n}\right) v_{k,j}, 0\right]^\top, & \text{if } i \equiv 2 (\text{mod } 3) \\ \left[0, 0, 0.9^t + \sum_{j=0}^{t-1} 0.9^j \cos\left(\frac{i\pi}{n}\right) v_{k,j}\right]^\top, & \text{if } i \equiv 0 (\text{mod } 3) \end{cases}$$

where $\{v_{k,i}\}$ are spatially and temporally i.i.d. to a Gaussian distribution $N(0, 0.4^2)$. To demonstrate the robustness and scalability of the proposed algorithm, we examine its parameter tracking performance under different network topologies, namely the complete graph, the ring graph, and the scale-free network (as shown in Fig. 1).

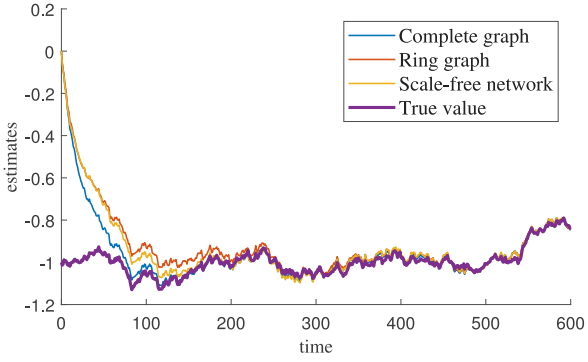
Under the given settings, it can be easily verified that Assumptions 1 to 3 are satisfied. Here, we repeat the simulation 50 times with the same initial state. Fig. 2 illustrates the average estimation performance of the proposed algorithm across 50 agents under different network topologies with the privacy parameter $\varepsilon = 0.1$. The three subplots in Fig. 2 display the estimates under complete graphs, ring graphs, and scale-free networks, and the true values for the three dimensions of the parameter to be estimated, respectively. It can be observed that all agents are able to accomplish the estimation task with low error.



(a) The first dimension of estimates versus true value



(b) The second dimension of estimates versus true value



(c) The third dimension of estimates versus true value

Fig. 2. Trajectories of 3-dimensional parameter estimation across fully connected, ring, and scale-free network topologies with privacy parameter $\varepsilon = 0.1$.

To visualize the impact of Laplacian noise on the algorithm, we conducted parameter estimations using our distributed PP-NLMS algorithm under various scale parameters of Laplace noise. We employ the metric $\frac{1}{2500}\sum_{j=1}^{50}\sum_{i=1}^{50}\|\hat{\xi}_{i,k}^{(j)} - \xi_k\|^2$ (where $i = 1, \ldots, n$ and $k = 1, \ldots, 600$) to calculate the average estimation error. Here, $\hat{\xi}_{i,k}^{(j)}$ represents the estimated parameter value of agent i at time step k in the jth experiment.

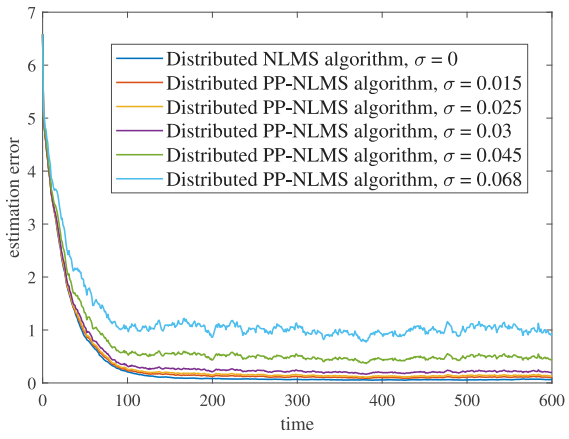


Fig. 3. The estimation error of Algorithm 1 with different Laplacian noise.

The outcomes are compared with those from the distributed NLMS algorithm without privacy protection, as illustrated in Fig. 3. Note that when the sensitivity remains unchanged, a larger σ can provide a higher degree of privacy protection. However, increased privacy protection inevitably leads to greater estimation errors. This highlights the necessity for our algorithm to balance between estimation performance and privacy protection, consistent with the previous Remark 7.

7. Concluding remarks

In summary, we have proposed a privacy-preserving distributed adaptive estimation algorithm for estimating time-varying parameters. Compared with other privacy protection algorithms, the superiority of our algorithm mainly lies in the allowance for non-stationary and non-independent regressors, as well as the effective estimation of time-varying parameters, making it applicable for feedback control systems. In this work, we establish differential privacy results for the proposed algorithm under finite iterations. In the future, we will consider designing diminishing step-sizes to ensure privacy protection under infinite iterations. Additionally, validation with real-world datasets will be conducted to further demonstrate the algorithm's practical efficacy.

CRediT authorship contribution statement

Shuning Chen: Writing – original draft, Validation, Formal analysis. **Die Gan:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **Siyu Xie:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis, Conceptualization. **Jinhu Lü:** Writing – review & editing, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] S. Kar, J.M.F. Moura, H.V. Poor, Distributed linear parameter estimation: Asymptotically efficient adaptive strategies, SIAM J. Control Optim. 51 (3) (2013) 2200–2229.
[2] D. Marelli, M. Zamani, M. Fu, B. Ninness, Distributed Kalman filter in a network of linear systems, Systems Control Lett. 116 (2018) 71–77.
[3] A.S. Matveev, M. Almodarresi, R. Ortega, A. Pyrkin, S. Xie, Diffusion-based distributed parameter estimation through directed graphs with switching topology: Application of dynamic regressor extension and mixing, IEEE Trans. Autom. Control 67 (8) (2022) 4256–4263, http://dx.doi.org/10.1109/TAC.2021.3115075.
[4] E. Dobriban, Y. Sheng, Distributed linear regression by averaging, Ann. Statist. 49 (2) (2021) 918–943, URL https://doi.org/10.1214/20-AOS1984.
[5] Q. Yang, Z. Zhang, M. Fu, Q. Cai, Asymptotic convergence of a distributed weighted least squares algorithm for networked systems with vector node variables, Syst. & Control. Lett. 165 (2022) 105265.
[6] N. Takahashi, I. Yamada, A.H. Sayed, Diffusion least-mean squares with adaptive combiners: Formulation and performance analysis, IEEE Trans. Signal Process. 58 (9) (2010) 4795–4810.
[7] D. Gan, Z. Liu, Performance analysis of the compressed distributed least squares algorithm, Systems Control Lett. 164 (2022) 105228.
[8] D. Gan, Z. Liu, Convergence of the distributed SG algorithm under cooperative excitation condition, IEEE Trans. Neural Netw. Learn. Syst. 35 (5) (2024) 7087–7101.
[9] A. Glushchenko, K. Lastochkin, Robust time-varying parameters estimation based on I-DREM procedure, IFAC- Pap. 55 (12) (2022) 91–96.
[10] I.D. Schizas, G. Mateos, G.B. Giannakis, Distributed LMS for consensus-based in-network adaptive processing, IEEE Trans. Signal Process. 57 (6) (2009) 2365–2382.
[11] S. Xie, L. Guo, Analysis of normalized least mean squares-based consensus adaptive filters under a general information condition, SIAM J. Control Optim. 56 (5) (2018) 3404–3431, http://dx.doi.org/10.1137/16M1106791.
[12] D. Gan, S. Xie, Z. Liu, J. Lü, Stability of FFLS-based diffusion adaptive filter under cooperative excitation condition, IEEE Trans. Autom. Control 69 (11) (2024) 7479–7492.
[13] J.-G. Lee, Q.V. Tran, K.-H. Oh, P.-G. Park, H.-S. Ahn, Distributed object pose estimation over strongly connected networks, Syst. & Control. Lett. 175 (2023) 105505.
[14] J. Lei, H.-F. Chen, Distributed estimation for parameter in heterogeneous linear time-varying models with observations at network sensors, Commun. Inf. Syst. 15 (2015) 423–451.
[15] Y. Hua, F. Wan, B. Liao, Y. Zong, S. Zhu, X. Qing, Adaptive multitask clustering algorithm based on distributed diffusion least-mean-square estimation, Inform. Sci. 606 (2022) 628–648.
[16] L. Guo, Estimation, control, and games of dynamical systems with uncertainty, Sci. Sinica Inf. 50 (9) (2020) 1327–1344.
[17] J.A. Calandrino, A. Kilzer, A. Narayanan, E.W. Felten, V. Shmatikov, "You might also like:" privacy risks of collaborative filtering, in: 2011 IEEE Symposium on Security and Privacy, Oakland, CA, USA, 2011, pp. 231–246.
[18] Y. Lu, M. Zhu, Privacy preserving distributed optimization using homomorphic encryption, Automatica 96 (2018) 314–325.
[19] H. Gao, Y. Wang, A. Nedić, Dynamics based privacy preservation in decentralized optimization, Automatica 151 (2023) 110878.
[20] Y. Wang, T. Başar, Decentralized nonconvex optimization with guaranteed privacy and accuracy, Automatica 150 (2023) 110858.
[21] M. Ruan, H. Gao, Y. Wang, Secure and privacy-preserving consensus, IEEE Trans. Autom. Control 64 (10) (2019) 4035–4049.
[22] C. Fioravanti, L. Faramondi, G. Oliva, C. Hadjicostis, A geometrical approach for consensus security, Syst. & Control. Lett. 185 (2024) 105717.
[23] K. Wei, J. Li, M. Ding, C. Ma, H.H. Yang, F. Farokhi, S. Jin, T.Q.S. Quek, H. Vincent Poor, Federated learning with differential privacy: Algorithms and performance analysis, IEEE Trans. Inf. Forensics Secur. 15 (2020) 3454–3469, http://dx.doi.org/10.1109/TIFS.2020.2988575.
[24] M. Seif, R. Tandon, M. Li, Wireless federated learning with local differential privacy, in: 2020 IEEE International Symposium on Information Theory, ISIT, Los Angeles, CA, USA, 2020, pp. 2604–2609.
[25] J. Ma, S.-A. Naas, S. Sigg, X. Lyu, Privacy-preserving federated learning based on multi-key homomorphic encryption, Int. J. Intell. Syst. 37 (9) (2022) 5880–5901.
[26] H. Wang, L.F. Toso, J. Anderson, FedSysID: A federated approach to sample-efficient system identification, in: Proceedings of the 5th Annual Learning for Dynamics and Control Conference, Vol. 211, Riverside, CA, USA, 2023, pp. 1308–1320.
[27] X. Shen, Y. Liu, Privacy-preserving distributed estimation over multitask networks, IEEE Trans. Aerosp. Electron. Syst. 58 (3) (2022) 1953–1965, http://dx.doi.org/10.1109/TAES.2021.3124866.
[28] M. Naehrig, K. Lauter, V. Vaikuntanathan, Can homomorphic encryption be practical? in: Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop, New York, NY, USA, 2011, pp. 113–124.

[29] C. Dwork, Differential privacy: A survey of results, in: Theory and Applications of Models of Computation, Berlin, Heidelberg, 2008, pp. 1–19.
[30] J. Le Ny, G.J. Pappas, Differentially private filtering, IEEE Trans. Autom. Control 59 (2) (2014) 341–354, http://dx.doi.org/10.1109/TAC.2013.2283096.
[31] A. Moradi, N.K.D. Venkategowda, S.P. Talebi, S. Werner, Privacy-preserving distributed Kalman filtering, IEEE Trans. Signal Process. 70 (2022) 3074–3089.
[32] J. Wang, J.-F. Zhang, X.-K. Liu, Differentially private resilient distributed cooperative online estimation over digraphs, Internat. J. Robust Nonlinear Control 32 (15) (2022) 8670–8688.
[33] J. Wang, J. Tan, J.-F. Zhang, Differentially private distributed parameter estimation, J. Syst. Sci. Complex. 36 (2023) 187–204.
[34] C. Dwork, A. Roth, The algorithmic foundations of differential privacy, Found. Trends® Theor. Comput. Sci. 9 (3–4) (2014) 211–407.
[35] L. Guo, L. Ljung, Performance analysis of general tracking algorithms, IEEE Trans. Autom. Control 40 (8) (1995) 1388–1402, http://dx.doi.org/10.1109/9.402230.
[36] L. Guo, Stability of recursive stochastic tracking algorithms, SIAM J. Control Optim. 32 (5) (1994) 1195–1225.