Received: 29 April 2020 | Revised: 10 November 2021 | Accepted: 3 July 2022

DOI: 10.1002/rnc.6303

RESEARCH ARTICLE

WILEY

Differentially private resilient distributed cooperative online estimation over digraphs

Jimin Wang[1] | **Ji-Feng Zhang**[2,3] | **Xiao-Kang Liu**[4]

[1]School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing, China

[2]Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

[3]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, China

[4]School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan, China

Correspondence
Ji-Feng Zhang, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.
Email: jif@iss.ac.cn

Funding information
National Key R&D Program of China, Grant/Award Number: 2018YFA0703800; National Natural Science Foundation of China, Grant/Award Numbers: 61877057, 62103156

Abstract

This article investigates resilient distributed online estimation (DOE) in unreliable directed networks with differential privacy requirements. In the network considered, some agents are subject to Byzantine attacks and thus could send arbitrary incorrect messages to their neighbors. The remaining agents aim to collaboratively estimate the value of an unknown vector parameter while protecting their private data. In this article, by adding private noises to mask the estimate, a stochastic approximation-type resilient differentially private DOE algorithm is proposed to protect the privacy of sensitive information. A time-varying step size is introduced to attenuate the divergence caused by the private noise, and furthermore, guarantees the convergence of the algorithm. When the directed graph is $(2F+1)$-robust, the algorithm is shown to be both mean square and almost sure convergence in the sense of ϵ-differential privacy. A simulation example is given to verify the effectiveness and superiority of the algorithm.

KEYWORDS

differential privacy, distributed online estimation, resilient estimation, stochastic approximation

1 | INTRODUCTION

Distributed online estimation (DOE) is a fundamental problem in many practical complex distributed systems including cyber-physical systems, sensor networks, social and economic networks. In this problem, each distributed unit cooperatively estimates or learns the unknown parameter of the system based on its own noisy measurements and the interaction information from its neighbor. Compared with the centralized estimation, DOE has two advantages: *scalability* and *reliability*. Specifically, DOE does not require to transmit all the data to a fusion center. This not only avoids the potential computation and communication burden with the increasing scale of the network but also enhances the robustness against attackers. Due to parallel information processing, DOE algorithms have been developed and widely applied in industrial monitoring, multi-agent systems, smart grid.[1-6] Among the existing DOE algorithms, consensus-based distributed estimation algorithms are important and popular, since it contains a consensus scheme to implement the coordination among sensors. In fact, due to its wide applications in distributed control and distributed estimation, distributed consensus problem of multi-agent systems has obtained deep and comprehensive results.[7-11]

But in practical application, DOE algorithms are vulnerable to the attack from an adversary, which can easily result in a failure of estimation. According to its ability, the adversary can be divided into the active adversary and the passive adversary. The active adversary acts arbitrarily and causes the original algorithm to no longer be applicable. Recently, DOE algorithms under the active adversary have been studied. For example, when the attackers can directly manipulate the sensor measurement, resilient distributed parameter estimation algorithms have been given in the work of Chen et al.[12-14] When the attackers are Byzantine, resilient distributed parameter estimation algorithms have been given in the work of Chen et al[15] and LeBlanc and Hassan.[16] As a special case of the active adversary, Byzantine attackers hijack a subset of the agents and drive them to spread the incorrect information to the network, which have been intensively studied in recent papers,[15-22] including distributed consensus, distributed optimization and distributed estimation. Unlike the active adversary, the passive adversary is only curious about the sensitive information of agents. If there is a passive adversary in the network, the information interaction among agents brings a privacy crisis. For example, by collecting interactive information, adversarial agents can obtain the sensitive information from their neighbors.[23] In medical prediction systems, each hospital wants to infer a regression model for medical treatment prediction. A good model requires the training data of patients as much as possible, however, data publication and sharing among hospitals may leak the medical records of individuals. In cooperative guidance systems, interactions may expose the positions of other missiles and the launch stations. In biological networks, animals are interested in moving toward a target (such as a nutrition source). The "home" of the animals is sensitive for hunters/predators. Therefore, there is a great need of privacy preserving algorithm to protect the sensitive information in control systems.[24] In the realm of distributed algorithms, several privacy preserving approaches have been recently proposed, such as adding artificial noises,[25] homomorphic encryption,[23,26] state decomposition,[27] and so forth.

Fortunately, differential privacy is widely used to protect the privacy[28] and has applications in many domains, such as social networks,[29] machine learning,[30-33] distributed optimization,[34,35] and so on. Differential privacy is to add noises to the message before delivering. However, noise addition disrupts the learning process and heavily degrades the performance of the trained model in the existing studies,[31-33] especially when large amounts of noises are considered to guarantee a small loss of privacy.

Differential privacy also has some results in the area of distributed consensus. For example, the differentially private iterative synchronous consensus problem has been studied in the work of Huang et al.,[36] where agents are required to achieve convergence to the average of the initial values while protecting the privacy of their initial values from honest-but-curious adversaries. In the work of Nozari et al.,[37] by properly designing the private noise, a privacy preserving algorithm has been proposed to guarantee the almost sure convergence to the average of agents' initial values while allowing agents to choose their own privacy level. When there are faulty agents and the graph topology is directed, an algorithm has been proposed to guarantee the consensus of non-faulty agents with differential privacy requirements in the work of Fiore and Russo.[38] However, the private noise used in the above literature[37,38] is exponentially decaying. By using noise with time-invariant variances, a differentially private consensus algorithm for continuous-time heterogeneous high-order multi-agent systems has been proposed in the work of Liu et al.[39] Recently, differentially private estimation has been developed via centralized methods[40,41] and distributed methods.[42,43] By adding noise to the private data (input perturbation), a differentially private distributed stochastic gradient algorithm under gossiping communication has been given by Liu et al.[42] However, the convergence for the proposed algorithm cannot be guaranteed, and the convergence error is inversely proportional to the privacy level. The differentially private distributed estimation problem under an undirected graph topology has been investigated by Liu et al.[43] Up to now, the design and analysis for DOE with the unknown subset of faulty agents and privacy requirement remains open. This is a challenging issue to discuss because it is much more practical and useful.

In this article, the differentially private DOE problem is considered over a directed graph under Byzantine attacks. The agents under attack are unknown, and the sensitive data of every non-faulty agent are required to be protected. First, we propose a stochastic approximation-type Byzantine-resilient differentially private DOE algorithm, which can achieve distributed parameter estimation under differential privacy requirement and Byzantine attack. Then, we prove that the algorithm is ϵ-differentially private and guarantees both mean square and almost sure convergence. Finally, a simulation example is given to verify the effectiveness and superiority of the algorithm. The contributions of this article are summarized as follows.

We apply the stochastic approximation-type algorithm to study the resilient differentially private distributed parameter estimation problem for the first time. We consider a certain number of faulty agents and privacy protection of non-faulty agents' sensitive information simultaneously while guaranteeing the convergence of estimation states. Compared with the previous work on distributed parameter estimation,[1-6,12-16] the algorithm in this article is differentially private. Compared with the work of Katewa et al.[41] and Liu et al.,[42] the convergence of the algorithm both in mean square and almost sure is established. Compared with the work of Ny and Pappas[40] and Katewa et al.,[41] the algorithm in this article is distributed, robust and has less computation and communication resources.

Instead of using adaptive sequential composition theorem, we explicitly give a method to determine the noise variance for ϵ-differential privacy. This method is more general than the exponential decay Laplace noise given in advance.[36-38] Compared with the work of Han et al.,[34] the private noise σ_k in this article can be $O(\frac{1}{k^\gamma})$, $\frac{1}{2} < \gamma \leq 1$. Moreover, different from the work of Fiore and Russo,[38] the faulty agents studied in this article do not affect the differential privacy of the non-faulty agents.

Compared with the existing differentially private distributed consensus algorithm,[36-39] the estimation errors of this article converge to zero. Different from that only directly manipulates sensor measurements[12-14] and some agents know the parameter value in advance,[16] the attacker in this article has stronger attack ability and all the agents do not know the parameter value a priori.

The rest of this article is organized as follows. In Section 2, graph theory is introduced, and the DOE problem is formulated. In Section 3, a DOE algorithm is proposed and shown to be ϵ-differentially private. Besides, both mean square and almost sure convergence are established. In Section 4, a numerical example is given to demonstrate the theoretical results. In Section 5, concluding remarks and further research topics are drawn.

The notations throughout this article are standard. $X \geq 0$ ($X > 0$) means that X is symmetric and semi-positive definite (positive definite). $\mathbf{1}_N$ stands for the N-dimensional vector with all elements being one. $\mathbb{R}^n$ and $\mathbb{R}^{m\times n}$ denote the n-dimensional Euclidean space and the set of all $m \times n$ real matrices, respectively. $\|x\|_2$ refers to Euclidean norm of the vector x. $\|x\|_1$ denotes the 1-norm of the vector $x \in \mathbb{R}^n$, that is, $\|x\|_1 = \sum_{i=1}^{n} |x_i|$. I, 0 are identity matrix and zero matrix with appropriate dimensions, respectively. In addition, diag $\{A_1, \ldots, A_n\}$ stands for a (block) diagonal matrix with $A_1, \ldots, A_n$ in order on the diagonal. $|\mathcal{A}|$ denotes the number of elements in the set $\mathcal{A}$. The probability, expectation, and variance of a random variable X are denoted by $\mathbb{P}[X]$, $\mathbb{E}[X]$, and $\mathbb{V}[X]$, respectively.

2 | PRELIMINARIES

2.1 | Graph theory

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consists of a set of *agents* $\mathcal{V} = \{1, 2, \ldots, N\}$, and a set of *edges* $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The graph is said to be *undirected* if $(i,j) \in \mathcal{E} \Leftrightarrow (j,i) \in \mathcal{E}$, and *directed*, otherwise. The set of *in-neighbors* and *out-neighbors* of agent i is denoted by $\mathcal{N}_i^{\text{in}} = \{j \in \mathcal{V} | (j,i) \in \mathcal{E}\}$ and $\mathcal{N}_i^{\text{out}} = \{j \in \mathcal{V} | (i,j) \in \mathcal{E}\}$, respectively. $\mathcal{G}$ is called a balanced digraph if $\sum_{j=1}^{N} a_{ij}(k) = \sum_{j=1}^{N} a_{ji}(k), \forall i \in \mathcal{V}$. A directed tree is a directed graph where every agent has exactly one parent except the root, and the root has directed paths to every other agent. A directed spanning tree of a directed graph is a directed tree that contains all agents of the directed graph. A directed graph contains a directed spanning tree if there exists a directed spanning tree as a subset of the directed graph. $A = [a_{ij}(k)]$ is the adjacency matrix of $\mathcal{G}$, where $a_{ij}(k) > 0$ if $(i,j) \in \mathcal{E}$ and $a_{ij}(k) = 0$, otherwise. $\mathcal{G}$ is called strongly connected if for any pair agents (i_1, i_l), there exists a path from i_1 to i_l consisting of edges $(i_1, i_2), (i_2, i_3), \ldots, (i_{l-1}, i_l)$. For a digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a nonempty subset $\mathcal{S} \subset \mathcal{V}$ is r-reachable set if there exists an agent $i \in \mathcal{S}$ such that $|\mathcal{N}_i^{\text{in}} \setminus \mathcal{S}| \geq r$. That is, the subset $\mathcal{S}$ has at least r in-neighbors in $\mathcal{N}_i^{\text{in}} \setminus \mathcal{S}$.

First, the concept of r-robustness for the digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, as well as a useful lemma, is given as follows.

Definition 1 (38). A digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is r-robust if for every pair of non-empty, disjoint subsets of $\mathcal{V}$, at least one of the subsets is r-reachable.

Lemma 1 (21). *Suppose a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is r-robust. Let $\mathcal{G}'$ be a graph obtained by removing $r - 1$ or fewer incoming edges from each agent in $\mathcal{G}$. Then, $\mathcal{G}'$ contains a directed spanning tree.*

2.2 | Observation model

The observation model of each agent i is governed by:

$$y_i(k) = H_i\theta^* + \omega_i(k),$$

where $y_i(k) \in \mathbb{R}^{m_i}$ is the measurement vector, $\omega_i(k) \in \mathbb{R}^{m_i}$ is the zero-mean independent and identically distributed measurement noise with covariance R_i, and $H_i \in \mathbb{R}^{m_i}$ is the constant measurement matrix, $\theta^* \in \mathbb{R}$ is an unknown parameter to be estimated.

To identify the system parameter θ^*, a consensus-based distributed estimation algorithm[1,3-6] can be designed in the form of

$$x_i(k+1) = x_i(k) - \alpha(k) \sum_{j \in \mathcal{N}_i} a_{ij}(k)(x_i(k) - x_j(k)) + \alpha(k) H_i^T (y_i(k) - H_i x_i(k)), \quad i \in \mathcal{V}, \tag{1}$$

where x_i is the estimate state of agent i, and $\alpha(k)$ is a positive step size. However, in the traditional distributed estimation algorithm (1), each agent transmits its estimate to its neighbors, which may expose the sensitive information. Moreover, the algorithm is vulnerable to attacks and may lead to failure if adversarial agents deliver incorrect information to the network.

2.3 | Fault/adversary model

Under the Byzantine attacks, agents can be divided into two subsets:[17-19,21,22,38] a set of adversarial (faulty) agents, denoted as $\mathcal{A}$, and a set of non-faulty agents, denoted as $\mathcal{V}/\mathcal{A}$. The faulty agents can send different falsified massages to different neighbors at each time-step. Mathematically, denote $\zeta_i^j(k) \in \mathbb{R}$ as the message sent from a faulty agent $i \in \mathcal{A}$ to a non-faulty agent $j \in \mathcal{V}/\mathcal{A}$ at time k. Then, faulty agent allows that $\zeta_i^j(k) \neq \zeta_i^{j'}(k)$ for $j \neq j'$ and $j, j' \in \mathcal{V}/\mathcal{A}$. Although faulty agents can deliver completely arbitrary, malicious, and possibly conflicting messages into the network, the non-faulty agents aim to estimate the unknown parameter collaboratively regardless of the incorrect information.

Remark 1. Different from the previous work,[12-14] instead of manipulating the sensor measurement, the Byzantine attacker hijacks a subset of agents to spread the incorrect information in the network, which is more general and realistic.

Definition 2 (F-total vs. F-local[17]). For $F \in \mathbb{N}$, we say that the set of faulty agents $\mathcal{A}$ is an F-total set if $|\mathcal{A}| \leq F$, and an F-local set if $|\mathcal{N}_i^{\text{in}} \cap \mathcal{A}| \leq F$, for all $i \in \mathcal{V}/\mathcal{A}$.

F-total model indicates that there are no more than F faulty agents in the entire network, whereas the F-local model indicates that there are no more than F faulty agents in the in-neighbors of any non-faulty agents.

Assumption A1. The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is directed and $(2F+1)$-robust.

Assumption A1 is fairly standard as seen in the work of Fiore and Russo.[38] Note that F-total model is a special case of F-local model under Assumption A1. Hence, we will focus on the F-local model in this article.

2.4 | Differential privacy

Besides the Byzantine attacks, there is potential eavesdropper that has access to the information in the communication network. In order to protect the privacy of non-faulty agents, the definition of ϵ-differential privacy is introduced. Inspired by the work of Ny and Pappas,[40] the definition of δ-adjacent is given first.

Definition 3 (δ-Adjacent). Given $\delta > 0$ and $k \in \mathbb{N}$, two signal sets $D_k = \{y_i(k), i \in \mathcal{V}/\mathcal{A}\}$ and $D'_k = \{y'_i(k), i \in \mathcal{V}/\mathcal{A}\}$, D_k and D'_k are δ-adjacent if there exists $i_0 \in \mathcal{V}/\mathcal{A}$ such that

$$y_i(k) = y'_i(k), \quad \forall i \neq i_0, \quad \|y_{i_0}(k) - y'_{i_0}(k)\|_1 \leq \delta. \tag{2}$$

Definition 4 (Differential privacy[33]). Given $\epsilon > 0$, a randomized algorithm $\mathcal{Q}$ is ϵ-differentially private if for all *δ-adjacent* datasets $D = \{D_k, k \in \mathbb{N}\}$ and $D' = \{D'_k, k \in \mathbb{N}\}$, and for any set of outputs $\Upsilon \subseteq \text{Range}(\mathcal{Q})$,

$$\mathbb{P}\{\mathcal{Q}(D) \in \Upsilon\} \leq \mathbb{P}\{\mathcal{Q}(D') \in \Upsilon\} \times \exp(\epsilon \times |D \oplus D'|),$$

where $\oplus$ denotes symmetric difference, and $|D \oplus D'|$ denotes the number of different elements in datasets D and D'.

Remark 2. By actively adding noise, the estimate is randomized so that the eavesdropper cannot tell from the estimate with high probability whether the measurement corresponding to the agent in the network has changed. Thus, the sensitive information $D_k = \{y_i(k), i \in \mathcal{V}/\mathcal{A}\}$ is protected. The index ϵ measures the privacy level of the randomized algorithm $\mathcal{Q}$, and a small ϵ implies a high privacy level. As pointed out in the work of Ny and Pappas,[40] ϵ is generally settled by a small constant, for example, $\epsilon \approx 0.1$.

In the following, the sensitivity of the output map is introduced, which is a key quantity determines how much noises should be added in each iteration for achieving differential privacy.

Definition 5 (Sensitivity[40]). For the adjacency relation (2), the sensitivity of an output map m at time k is defined as

$$\Delta(k) = \sup_{D_k, D'_k : \mathrm{Adj}(D_k, D'_k)} \|m(D_k) - m(D'_k)\|_1,$$

where D_k and D'_k are input datasets at time k.

Remark 3. The sensitivity of an output map m captures a magnitude by which a single agent's data can change the output map m in the worst case. For our proposed algorithm, m refers to the output map from $y_i(k)$ to $x_i(k)$, $i \in \mathcal{V}/\mathcal{A}$.

In this article, we will design an ϵ-differentially private resilient distributed algorithm to achieve DOE with an unknown subset of faulty agents over a digraph.

3 | MAIN RESULT

3.1 | Algorithm design

A Byzantine-resilient differentially private distributed parameter estimation algorithm is proposed in Algorithm 1. The key steps of the algorithm are summarized as: (i) each agent actively adds random noise into the message before sharing it with its neighbors; (ii) each agent discards part of messages from neighbors before updating; (iii) each agent updates its estimate state based on the reserved messages.

Algorithm 1. Byzantine-resilient differentially private distributed parameter estimation algorithm

```
Input: $F$
Initialization: Set $x_i(0)$ randomly for each agent $i \in \mathcal{V}$
for $k = 1, \dots,$ do
Obtain a new measurement $y_i(k)$;
Transmit phase
Generate noise $n_i(k)$; and send $\tilde{x}_i(k) = x_i(k) + n_i(k)$ to each out-neighbor.
Receive phase
Get $\tilde{x}_j(k)$ from in-neighbors, and store the values in the vector $r_i(k)$ of size $|\mathcal{N}_i^{\text{in}}|$.
Update phase
If there are less than $F$ values strictly greater than $\tilde{x}_i(k)$ in $r_i(k)$, then agent $i$ removes all values strictly greater than $\tilde{x}_i(k)$ from $r_i(k)$. Otherwise, agent $i$ removes the $F$ largest values from $r_i(k)$. In addition, if there are less than $F$ values strictly less than $\tilde{x}_i(k)$ in $r_i(k)$, then agent $i$ removes all values strictly less than $\tilde{x}_i(k)$ from $r_i(k)$. Otherwise, agent $i$ removes the $F$ smallest values from $r_i(k)$.
Let $\mathcal{R}_i$ be the set stores the agent indexes whose state valueswere removed from $r_i(k)$, and denote the remaining agent index set as $\mathcal{N}_i^* = \mathcal{N}_i^{\text{in}}/\mathcal{R}_i$. Then, each non-faulty agent updates its estimate as follows.
$$x_i(k+1) = \tilde{x}_i(k) - \alpha(k) \sum_{j \in \mathcal{N}_i^*} a_{ij}(k)(\tilde{x}_i(k) - \tilde{x}_j(k)) + \alpha(k) H_i^T (y_i(k) - H_i \tilde{x}_i(k)).$$
end
```

Transmit phase: Each agent $i \in \mathcal{V}/\mathcal{A}$ transmits its current noisy estimate $\tilde{x}_i(k)$ to its neighbors $j \in \mathcal{N}_i^{\text{out}}$. Specifically, each agent broadcasts

$$\tilde{x}_i(k) = x_i(k) + n_i(k), \tag{3}$$

where $x_i(k)$ is the estimate of agent i at time $k \geq 1$ for the unknown parameter θ^*, and $n_i(k) \in \mathbb{R}$ is a zero-mean Laplacian noise with the variance of $2\sigma_k^2$, that is, $n_i(k) \sim \text{Lap}(0, \sigma_k)$.

Update phase: At each time step k, each agent i gets $\tilde{x}_j(k)$ from the in-neighbors and stores the values in the vector $r_i(k)$. If there are less than F values strictly greater than $\tilde{x}_i(k)$ in $r_i(k)$, then agent i removes all values strictly greater than $\tilde{x}_i(k)$ from $r_i(k)$. Otherwise, agent i removes the F largest values from $r_i(k)$. In addition, if there are less than F values strictly less than $\tilde{x}_i(k)$ in $r_i(k)$, then agent i removes all values strictly less than $\tilde{x}_i(k)$ from $r_i(k)$. Otherwise, agent i removes the F smallest values from $r_i(k)$. Then, each non-faulty agent updates its estimate as follows.

$$x_i(k+1) = \tilde{x}_i(k) - \alpha(k) \sum_{j \in \mathcal{N}_i^*} a_{ij}(k)(\tilde{x}_i(k) - \tilde{x}_j(k)) + \alpha(k) H_i^T (y_i(k) - H_i \tilde{x}_i(k)), \tag{4}$$

where $\alpha(k)$ is the step size of the algorithm. The selection of $\alpha(k)$ to ensure the convergence and ϵ-differential privacy of the algorithm simultaneously will be discussed in more details later.

Remark 4. Different from the existing distributed parameter estimation algorithms,[1-6,12-16] to ensure ϵ-differential privacy requirement, we add a Laplacian noise to the estimate $x_i(k)$ before broadcasting. Moreover, different from the previous work,[1-6,40-43] the algorithm addresses distributed parameter estimation when a subset of agents are Byzantine.

The algorithm cannot definitely eliminate the values generated by faulty agents. However, by the fact that the network topology is $(2F+1)$-robust, it is always possible to express the value $x_j(k)$ of any faulty agent in $\mathcal{N}_i^*$ by a convex combination of the non-faulty agents in $\mathcal{N}_i^{\text{in}} \cup \{i\}$.[21,38] In the following, a mathematically equivalent representation of (4) is presented.

Lemma 2. *For each agent $i \in \mathcal{V}/\mathcal{A}$, if Assumption A1 holds, then the update rule (4) for the multi-agent system (3) is mathematically equivalent to the following form*

$$\begin{aligned} x_i(k+1) = x_i(k) - \alpha(k) \sum_{j \in \mathcal{N}_i^{\text{in}} \cap (\mathcal{V}/\mathcal{A})} \overline{a}_{ij}(k)(x_i(k) - x_j(k)) + n_i(k) - \alpha(k) \sum_{j \in \mathcal{N}_i^{\text{in}} \cap (\mathcal{V}/\mathcal{A})} \overline{a}_{ij}(k)(n_i(k) - n_j(k)) \\ + \alpha(k) H_i^T (y_i(k) - H_i \tilde{x}_i(k)), \end{aligned} \tag{5}$$

where the nonnegative weights $\bar{a}_{ij}(k)$ satisfy the following property at each time-step k,

$$\overline{a}_{ii}(k) + \alpha(k) \sum_{j \in \mathcal{N}_i^{\text{in}} \cap (\mathcal{V}/\mathcal{A})} \overline{a}_{ij}(k) = 1.$$

Proof. Without loss of generality, we suppose the first $N - |\mathcal{A}|$ agents are non-faulty. This is because we can reindex the agents, and the communication topology is still $(2F+1)$-robust. This lemma can be proved by using arguments similar to the work of Sundaram and Gharesifard,[21] where an analogous result with the noise-free case (i.e., $n_i(k) = 0, \forall i \in \mathcal{V}$) was proved.

From (4), it follows that

$$\begin{aligned} x_i(k+1) &= (1 - \alpha(k) \sum_{j \in \mathcal{N}_i^*} a_{ij}(k)) \tilde{x}_i(k) + \alpha(k) \sum_{j \in \mathcal{N}_i^*} a_{ij}(k) \tilde{x}_j(k) + \alpha(k) H_i^T (y_i(k) - H_i \tilde{x}_i(k)) \\ &= a_{ii}(k) \tilde{x}_i(k) + \alpha(k) \sum_{j \in \mathcal{N}_i^*} a_{ij}(k) \tilde{x}_j(k) + \alpha(k) H_i^T (y_i(k) - H_i \tilde{x}_i(k)). \end{aligned} \tag{6}$$

Note that the set $\mathcal{N}_i^*$ in (6) may contain values coming from faulty agents. For example, faulty agents might generate values between the F largest and the F smallest values received by a non-faulty agent, and hence, those values would not be eliminated by the algorithm. In this case, as shown in the work of Sundaram and Gharesifard,[21] under the assumption that the network topology is $(2F+1)$-robust, it is always possible to express the value $\tilde{x}_j(k)$ from any faulty agent in $\mathcal{N}_i^*$ as the convex combination of the non-faulty agents in $\{i\} \cup \mathcal{N}_i^{\text{in}}$. By combining (3) and (6), we have

$$\begin{aligned} x_i(k+1) &= a_{ii}(k)\tilde{x}_i(k) + \alpha(k)\sum_{j\in\mathcal{N}_i^*} a_{ij}(k)\tilde{x}_j(k) + \alpha(k)H_i^T(y_i(k) - H_i\tilde{x}_i(k)) \\ &= \overline{a}_{ii}(k)\tilde{x}_i(k) + \alpha(k)\sum_{j\in\mathcal{N}_i^{\text{in}}\cap(\mathcal{V}/\mathcal{A})} \overline{a}_{ij}(k)\tilde{x}_j(k) + \alpha(k)H_i^T(y_i(k) - H_i\tilde{x}_i(k)) \\ &= \tilde{x}_i(k) - \alpha(k)\sum_{j\in\mathcal{N}_i^{\text{in}}\cap(\mathcal{V}/\mathcal{A})} \overline{a}_{ij}(k)(\tilde{x}_i(k) - \tilde{x}_j(k)) + \alpha(k)H_i^T(y_i(k) - H_i\tilde{x}_i(k)), \end{aligned}$$

where the last equality holds since

$$\overline{a}_{ii}(k) = 1 - \alpha(k)\sum_{j\in\mathcal{N}_i^{\text{in}}\cap(\mathcal{V}/\mathcal{A})} \overline{a}_{ij}(k).$$

The proof is completed. ∎

3.2 | Privacy analysis

In this subsection, we investigate ϵ-differential privacy of the algorithm. As explained previously, to protect the privacy, each non-faulty agent generates a noisy signal and adds the noise to the estimate $x_i(k)$, that is, $\tilde{x}_i(k) = x_i(k) + n_i(k)$. In order to satisfy ϵ-differential privacy, the output of the mechanism should satisfy Definition 4. In the context of differential privacy, the corresponding mechanism of the algorithm maps $D = \{y_i(k), i \in \mathcal{V}/\mathcal{A}\}_{k\in\mathbb{N}}$ to a sequence of messages $\Upsilon = \{\tilde{x}_i(k), i \in \mathcal{V}/\mathcal{A}\}_{k\in\mathbb{N}}$. This method is known as output perturbation.[28,40]

Definition 6 (L_1-sensitivity). For any given *δ-adjacent* datasets D_k and D'_k, the L_1-sensitivity of the algorithm at k is defined as

$$\Delta(k) = \sup_{D_k, D'_k : \text{Adj}(D_k, D'_k)} \|x_i(k) - x'_i(k)\|_1,$$

where $x'_i(k)$ corresponds to the estimate of agent $i \in \mathcal{V}/\mathcal{A}$ of the *adjacent* datasets.

Lemma 3. *Let $H_{\max} = \max_{i\in\mathcal{V}/\mathcal{A}} \|H_i\|_1$. Then, for the adjacency relation (2), the L_1 sensitivity of the algorithm at k satisfies*

$$\Delta(k) \le \alpha(k-1)\delta H_{\max}.$$

Proof. For two *adjacent* datasets D_k and D'_k, from (2) and (5), it follows that

$$\begin{aligned} \Delta(k) &= \max_{i\in\mathcal{V}/\mathcal{A}} \|x_i(k) - x'_i(k)\|_1 \\ &= \max_{i\in\mathcal{V}/\mathcal{A}} \|\alpha(k-1)H_i^T(y_i(k-1) - y'_i(k-1))\|_1 \\ &\le \alpha(k-1)\delta H_{\max}. \end{aligned}$$

Recall that σ_k is the scale parameter of the noise $n_i(k)$. We derive the following theorem on the scale parameter with respect to ϵ-differential privacy requirement. ∎

Theorem 1 (Differential privacy). *For a given $\epsilon > 0$, Algorithm 1 is ϵ-differentially private for i $\in \mathcal{V}/\mathcal{A}$ if σ_k satisfies*

$$\sigma_k \ge \frac{\Delta(k)}{\epsilon}. \tag{7}$$

Proof. Consider any pair of δ-adjacent datasets D_k and D'_k and any observation sequence $\Upsilon = \{\tilde{x}_i(k), i \in \mathcal{V}/\mathcal{A}\}_{k\in\mathbb{N}}$. From (5), the observation sequence depends on D_k, D'_k, and the noise sequence $\{n_i(k), i \in \mathcal{V}/\mathcal{A}\}_{k\in\mathbb{N}}$. We introduce $\mathcal{Q}(D_k) = \{\tilde{x}_i(k), i \in \mathcal{V}/\mathcal{A}\}$, $\mathcal{Q}(D'_k) = \{\tilde{x}'_i(k), i \in \mathcal{V}/\mathcal{A}\}$ to represent the injective map from the sensitive information to the observation sequence. In order to preserve ϵ-differential privacy, D_k and D'_k generate identical observation, that is, for any $k \in \mathbb{N}$, $\tilde{x}_i(k) = \tilde{x}'_i(k), i \in \mathcal{V}/\mathcal{A}$. Note that $\mathcal{P} = \{\chi(D_k, \mathcal{Q}(D_k))\}$ and $\mathcal{P}' = \{\chi(D'_k, \mathcal{Q}(D'_k))\}$ are the set of possible state evolution sequences under the adjacent datasets D_k and D'_k in the observation set Υ, with the probability density functions $f(D_k, \chi(D_k, \mathcal{Q}(D_k)))$ and $f(D'_k, \chi(D'_k, \mathcal{Q}(D'_k)))$, respectively. Then, it can be derived that

$$\frac{\mathbb{P}[\mathcal{Q}(D) \in \Upsilon]}{\mathbb{P}[\mathcal{Q}(D') \in \Upsilon]} = \lim_{k\to\infty} \frac{\int_{\chi(D_k,\mathcal{Q}(D_k))\in\mathcal{P}} f\,(D_k, \chi(D_k, \mathcal{Q}(D_k)))\, d\tau}{\int_{\chi(D'_k,\mathcal{Q}(D'_k))\in\mathcal{P}'} f\left(D'_k, \chi(D'_k, \mathcal{Q}(D'_k))\right) d\tau'}. \tag{8}$$

In the following, we prove that there exists a bijection $g(\cdot) : \mathcal{P} \to \mathcal{P}'$, such that for any pair of $\chi(D_k, \mathcal{Q}(D_k)) \in \mathcal{P}$ and $\chi(D'_k, \mathcal{Q}(D'_k)) \in \mathcal{P}'$, it has $g(\chi(D_k, \mathcal{Q}(D_k))) = \chi(D'_k, \mathcal{Q}(D'_k))$. From (5), when we use the sensitivity given in Lemma 3, we have the same observation over time horizon T, and there exists a bijection $g(\cdot) : \mathcal{P} \to \mathcal{P}'$, such that for any pair of $\chi(D_k, \mathcal{Q}(D_k)) \in \mathcal{P}$ and $\chi(D'_k, \mathcal{Q}(D'_k)) \in \mathcal{P}'$, it has $g(\chi(D_k, \mathcal{Q}(D_k))) = \chi(D'_k, \mathcal{Q}(D'_k))$. Then, from (8) one can get

$$\begin{aligned}\frac{\mathbb{P}[\mathcal{Q}(D) \in \Upsilon]}{\mathbb{P}[\mathcal{Q}(D') \in \Upsilon]} &= \lim_{k\to\infty} \frac{\int_{\chi(D_k,\mathcal{Q}(D_k))\in\mathcal{P}} f\,(D_k, \chi(D_k, \mathcal{Q}(D_k)))\, d\tau}{\int_{g(\chi(D_k,\mathcal{Q}(D_k)))\in\mathcal{P}'} f\left(D'_k, g(\chi(D_k, \mathcal{Q}(D_k)))\right) d\tau} \\ &= \lim_{k\to\infty} \frac{\int_{\chi(D_k,\mathcal{Q}(D_k))\in\mathcal{P}} f\,(D_k, \chi(D_k, \mathcal{Q}(D_k)))\, d\tau}{\int_{\chi(D_k,\mathcal{Q}(D_k))\in\mathcal{P}} f\left(D'_k, g(\chi(D_k, \mathcal{Q}(D_k)))\right) d\tau}.\end{aligned} \tag{9}$$

Note that we can write the probability density function of $\mathcal{Q}(D)$ as follows

$$\begin{aligned}&f(D_k, \chi(D_k, \mathcal{Q}(D_k))) = f(\tilde{x}(0), \tilde{x}(1), \dots)\\ &= \prod_{k\in\mathbb{N}} f(\tilde{x}(k)|\tilde{x}(0), \tilde{x}(1), \dots, \tilde{x}(k-1))\\ &= \prod_{k\in\mathbb{N}} f(\tilde{x}(k)|\tilde{x}(k-1))\\ &= \prod_{k\in\mathbb{N}, i\in\mathcal{V}/\mathcal{A}} f(\tilde{x}_i(k)|\tilde{x}(k-1))\\ &= \prod_{k\in\mathbb{N}, i\in\mathcal{V}/\mathcal{A}} f_{\sigma_k}(\tilde{x}_i(k) - x_i(k)),\end{aligned}$$

where the third equation holds since the randomness comes from the additive noise (3) and $\tilde{x}(k)$ conditioned on $\tilde{x}(k-1)$ is independent of all the previous $(\tilde{x}(0), \dots, \tilde{x}(k-2))$ of the algorithm. The added noise of each agent is independent of each other, which implies the fourth equation. In the fifth equation, $f_{\sigma_k}(x)$ is the probability density function at x in term of $\mathrm{Lap}(0, \sigma_k)$.

In the context of this problem, adjacent datasets differ at most one agent, denoted by $i_0 \in \mathcal{V}/\mathcal{A}$. Besides, since we do not know how many faulty messages are removed by the algorithm at each step, the worst case where the noise coming from all faulty agents has effect on each non-faulty agent is considered. Then, for single one iteration $k \in \mathbb{N}$, from the property of Laplace distribution, we have

$$\begin{aligned}\frac{f\{\mathcal{Q}(D_k) = \tilde{x}(k)\}}{f\{\mathcal{Q}(D'_k) = \tilde{x}(k)\}} &= \prod_{i\in\mathcal{V}/\mathcal{A}} \frac{f_{\sigma_k}(\tilde{x}_i(k) - x_i(k))}{f_{\sigma_k}(\tilde{x}_i(k) - x'_i(k))}\\ &= \prod_{i\in\mathcal{V}/\mathcal{A}} \frac{\exp\left(-\frac{|\tilde{x}_i(k) - x_i(k)|}{\sigma_k}\right)}{\exp\left(-\frac{|\tilde{x}_i(k) - x'_i(k)|}{\sigma_k}\right)}\\ &\leq \prod_{i\in\mathcal{V}/\mathcal{A}} \exp\left(\frac{|\tilde{x}_i(k) - x_i(k) - \tilde{x}_i(k) + x'_i(k)|}{\sigma_k}\right)\end{aligned}$$

$$
\begin{aligned}
&= \exp\left(\frac{|x_{i_0}(k) - x'_{i_0}(k)|}{\sigma_k}\right)\\
&\leq e^{\frac{\Delta(k)}{\sigma_k}},
\end{aligned}
$$

and hence,

$$
\prod_{k\in\mathbb{N}} f\{\mathcal{Q}(D_k) \in \Upsilon\} = \prod_{k\in\mathbb{N}} f\{\mathcal{Q}(D'_k) \in \Upsilon\} \times \prod_{k\in\mathbb{N}} \exp\left(\left(\frac{\Delta(k)}{\sigma_k}\right) \times |D_k \oplus D'_k|\right),
$$

which together with (7) gives

$$
\begin{aligned}
\prod_{k\in\mathbb{N}} f\{\mathcal{Q}(D_k) \in \Upsilon\} &\leq \prod_{k\in\mathbb{N}} f\{\mathcal{Q}(D'_k) \in \Upsilon\} \times \prod_{k\in\mathbb{N}} \exp\left(\left(\frac{\Delta(k)}{\sigma_k}\right) \times |D_k \oplus D'_k\right)\\
&\leq \prod_{k\in\mathbb{N}} f\{\mathcal{Q}(D'_k) \in \Upsilon\} \times \exp(\epsilon \times |D \oplus D'|). \qquad (10)
\end{aligned}
$$

From (9) and (10), it follows that

$$
\mathbb{P}\{\mathcal{Q}(D) \in \Upsilon\} \leq \mathbb{P}\{\mathcal{Q}(D') \in \Upsilon\} \times \exp(\epsilon \times |D \oplus D'|).
$$

Therefore, according to Definition 4, the statement of Theorem 1 is obtained. ■

Remark 5. Compared with the previous work on differentially private distributed algorithm,[34,36,37,43] we consider differentially private distributed algorithm with unknown subset of faulty agents, which is closer to the reality of distributed network. Furthermore, different from the work of Fiore and Russo,[38] the faulty agents studied in Theorem 1 do not affect the differential privacy of non-faulty agents.

Remark 6. Note that when $|D \oplus D'| = T$, $T > 0$, the differentially private distributed parameter estimation algorithm satisfies $T\epsilon$-differential privacy, which reduces to the general *adaptive sequential composition*.[34] According to *adaptive sequential composition*, we know that the privacy level of the algorithm (3) and (5) will degrade after T-iteration adaptive composition. From Theorem 1, it follows that the privacy level after T iterations is the sum of that in single iteration.

Based on Theorem 1 and Lemma 3, we give the following corollary directly.

Corollary 1. *For given $\epsilon > 0$, if σ_k satisfies*

$$
\sigma_k \geq \frac{\alpha(k-1)}{\epsilon} \delta H_{\max}, \qquad (11)
$$

then the algorithm (3) and (5) is ϵ-differentially private for $\forall i \in \mathcal{V}/\mathcal{A}$.

Remark 7. In Corollary 1, a relationship between scale parameter σ_k and the step size $\alpha(k)$ is derived. From (11), it follows that the scale parameter σ_k of the added noise is inversely proportional to privacy level ϵ. In other words, each agent protects stronger privacy when the added noise is more dispersive. In order to preserve ϵ-differential privacy and meanwhile make the parameter estimation error as accurate as possible, we will take $\sigma_k = \frac{\alpha(k-1)}{\epsilon} \delta H_{\max}$ in the following.

3.3 | Convergence analysis

Privacy comes with a price, the noise term in Algorithm 1 makes the convergence rate slower than the noise-free case. However, in this subsection, both mean square and almost sure convergence of the algorithm are proved. Without loss of generality, we suppose the indexes $\{1, 2, \ldots, N - |\mathcal{A}|\}$ are non-faulty agents. This is because we can reindex the agents, and the communication topology is still $(2F + 1)$-robust. The stacked vectors of the dynamics (5) are given as follows.

$$
\Theta^* = \mathbf{1}_{N-|\mathcal{A}|} \otimes \theta^*,
$$

$$\begin{aligned} X(k) &= \left[x_1^T(k), \ \dots \ , x_{N-|\mathcal{A}|}^T(k)\right]^T, \\ Y(k) &= \left[y_1^T(k), \ \dots \ , y_{N-|\mathcal{A}|}^T(k)\right]^T, \\ H &= \text{diag}\{H_1^T, \ \dots \ , H_{N-|\mathcal{A}|}^T\}, \\ \omega(k) &= \left[\omega_1^T(k), \ \dots \ , \omega_{N-|\mathcal{A}|}^T(k)\right]^T, \\ n(k) &= \left[n_1^T(k), \ \dots \ , n_{N-|\mathcal{A}|}^T(k)\right]^T. \end{aligned} \tag{12}$$

From (5) and (12), the closed-loop system can be described in the following compact form:

$$\begin{aligned} X(k+1) = X(k) - \alpha(k)\mathcal{L}_{\mathcal{G}'}X(k) + \alpha(k)HH^T(\Theta^* - X(k)) + \alpha(k)H\omega(k) + n(k) \\ - \alpha(k)(\mathcal{L}_{\mathcal{G}'} + HH^T)n(k), \end{aligned} \tag{13}$$

where $\mathcal{L}_{\mathcal{G}'}$ is the Laplacian matrix of the reduced graph.

There are some commonly used assumptions, which we encapsulate below.

A1′. The graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is directed and $(2F+1)$-robust. The reduced graph $\mathcal{G}'$ is balanced.
A2. The matrix $G = \sum_{i\in\mathcal{V}/\mathcal{A}} H_i^T H_i$ is of full rank.
A3. $\{\omega(k), k \geq 0\}$ is with bounded second moments $\sigma_\omega^2 \triangleq \sup_{k\geq 0}\mathbb{E}[\|\omega(k)\|^2] < \infty$.
A4. The step size $\{\alpha(k), k \geq 0\}$ satisfies

$$\alpha(k) > 0, \ \sum_{k\geq 0}\alpha(k) = \infty, \ \sum_{k\geq 0}\alpha^2(k) < \infty, \ \alpha(k) = O(\alpha(k-1)).$$

Assumptions A1′–A4 are standard in practice, as seen in existing literature.[1,3,21] Before giving the main result of this article, the following definition and lemmas are introduced.

Definition 7 (44,45). Let $Z_m \subset \mathbb{R}^{m\times m}$ denote the set of all square matrices of dimension m with non-positive off-diagonal entries. A matrix $U \in \mathbb{R}^{m\times m}$ is a nonsingular M-matrix if $U \in Z_m$ and all eigenvalues of U have positive real parts.

Lemma 4 (44,45). *For a matrix $U \in Z_m$, U is a nonsingular M-matrix if and only if there exists a diagonal matrix $D =$ diag$\{d_1, \ \dots \ , d_m\}$ with $d_i > 0, i = 1, \ \dots \ , m$ such that $Q = DU + U^TD$ is a symmetric positive definite matrix.*

Lemma 5. *Under Assumptions A1′ and A2, $(\Phi\mathcal{L}_{\mathcal{G}'} + \mathcal{L}_{\mathcal{G}'}^T\Phi) + 2\Phi HH^T$ is a positive definite matrix. Furthermore, there exists a positive definite matrix $M \in \mathbb{R}^{(N-|\mathcal{A}|)\times(N-|\mathcal{A}|)}$ and a sufficiently large integer T, such that for any $k > T$,*

$$\alpha(k)M < \alpha(k)((\Phi\mathcal{L}_{\mathcal{G}'} + \mathcal{L}_{\mathcal{G}'}^T\Phi) + 2\Phi HH^T),$$

where $\Phi = \text{diag}\{\xi_1, \ \dots \ , \xi_{N-|\mathcal{A}|}\}$, with $\sum_{i=1}^{N-|\mathcal{A}|}\xi_i = 1$, $\xi_i > 0, i = 1, \ \dots \ , N - |\mathcal{A}|$.

Proof. When $\mathcal{G}$ is $(2F+1)$-robust, from Lemma 1, it follows that the reduced graph $\mathcal{G}'$ consisting of non-faulty agents contains a directed spanning tree after removing $2F$ or fewer edges from each non-faulty agent. Note that a balanced digraph is strongly connected if and only if it has a directed spanning tree. Then, we have $\mathcal{G}'$ is strongly connected. Furthermore, similar to the proof of Theorem 3.1 in the work of Mei and Ren,[6] the result can be obtained. ■

Lemma 6 (46). *Let $\{V(k), k = 0, 1, \ \dots \ \}$, $\{p(k), k = 0, 1, \ \dots \ \}$, and $\{q(k), k = 0, 1, \ \dots \ \}$ be real sequences, satisfying $0 < q(k) \leq 1$, $p(k) \geq 0$, $k = 0, 1, \ \dots \ $, $\sum_{k=0}^{\infty} q(k) = \infty$, $\dfrac{p(k)}{q(k)} \to 0$, as $k \to \infty$, and $V(k+1) \leq (1 - q(k))V(k) + p(k)$. Then, $\limsup_{k\to\infty} V(k) \leq 0$. Particularly, if $V(k) \geq 0$, $k = 0, 1, \ \dots$, then $\lim_{k\to\infty} V(k) = 0$.*

For convenience analysis, when private noises exist, we need to properly choose the step size. In the following, we establish the mean square and almost sure convergence of Algorithm 1 by using the stochastic approximation-type conditions under two different graph topology conditions, that is, Assumption A1′ and A1, respectively.

Theorem 2. *If Assumptions A1′–A4 hold, then for any $\epsilon > 0$ and $\sigma_k = \frac{\alpha(k-1)}{\epsilon}\delta H_{\max}$, the estimate $x_i(k)$ given by (5) is convergent to θ^* in mean square, that is,*

$$\lim_{k\to\infty}\mathbb{E}\left[\|\delta(k)\|^2\right] = 0.$$

Proof. Let $\delta(k) = X(k) - \Theta^*$ and $V(k) = \delta^T(k)\Phi\delta(k)$. Then, from (13), it follows that

$$V(k+1) = \delta^T(k)I_0^T\Phi I_0\delta(k) + 2\delta^T(k)I_0^T\Phi W(k) + W^T(k)\Phi W(k), \tag{14}$$

where $I_0 = I \;-\; \alpha(k)(\;\mathcal{L}_{\mathcal{G}'} + HH^T)$, $W(k) = n(k) + \alpha(k)H\omega(k) - \alpha(k)(\mathcal{L}_{\mathcal{G}'} + HH^T)n(k)$.

From Assumption A1′, Assumption A2, and Lemma 5, it follows that there exists a constant $b_0 > 0$ such that $\frac{1}{2}(\Phi\mathcal{L}_{\mathcal{G}'} + \mathcal{L}_{\mathcal{G}'}^T\Phi) + \Phi HH^T \geq b_0 I$. Thus, we have

$$\delta^T(k)\left[\frac{1}{2}(\Phi\mathcal{L}_{\mathcal{G}'} + \mathcal{L}_{\mathcal{G}'}^T\Phi) + \Phi HH^T\right]\delta(k) \geq b_0\delta^T(k)\delta(k) \geq \frac{b_0}{\xi_{\max}}V(k), \tag{15}$$

where $\xi_{\max} = \max_{i=1,\dots,N-|\mathcal{A}|}\xi_i$. Note that

$$(\Phi\mathcal{L}_{\mathcal{G}'} + \mathcal{L}_{\mathcal{G}'}^T\Phi) + 2\Phi HH^T = (\mathcal{L}_{\mathcal{G}'} + HH^T)^T\Phi + \Phi(\mathcal{L}_{\mathcal{G}'} + HH^T) > 0. \tag{16}$$

Then, $\mathcal{L}_{\mathcal{G}'} + HH^T$ is nonsingular. Otherwise, there exists a nonzero vector $\psi \in \mathbb{R}^{N-|\mathcal{A}|}$ such that $\psi^T\left[(\Phi\mathcal{L}_{\mathcal{G}'} + \mathcal{L}_{\mathcal{G}'}^T\Phi) + 2\Phi HH^T\right]\psi = 0$, which is contrary to (16). $\mathcal{L}_{\mathcal{G}'} + HH^T$ is positive semi-definite, thus it is positive definite. Let λ be the maximum eigenvalue of $\mathcal{L}_{\mathcal{G}'} + HH^T$ and $\mathcal{F}_k = \sigma\{X(0), \omega(l), n(l), 0 \leq l \leq k-1\}$. Then, we have

$$\mathbb{E}\left[\delta^T(k)I_0^T\Phi I_0\delta(k)|\mathcal{F}_k\right] \leq \left[1 - 2\frac{b_0}{\xi_{\max}}\alpha(k) + \lambda^2\alpha^2(k)\right]V(k). \tag{17}$$

From $\mathbb{E}[\omega(k)] = 0$ and $\mathbb{E}[n(k)] = 0$, we have $\mathbb{E}\left[\omega(k)|\mathcal{F}_k\right] = \mathbb{E}\left[n(k)|\mathcal{F}_k\right] = 0$, which further implies that

$$\mathbb{E}\left[\delta^T(k)I_0^T\Phi W(k)|\mathcal{F}_k\right] = 0. \tag{18}$$

In addition, from the independence of $\omega(k)$ and $n(k)$, it follows that

$$\begin{aligned}\mathbb{E}[W^T(k)\Phi W(k)|\mathcal{F}_k] \leq \xi_{\max}(&\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k] + \alpha^2(t)\|H\|^2\mathbb{E}[\|\omega(k)\|^2|\mathcal{F}_k]\\ &+ \alpha^2(k)\|\mathcal{L}_{\mathcal{G}'} + HH^T\|^2\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k]).\end{aligned}$$

This together with (14)–(18) gives

$$\begin{aligned}\mathbb{E}\left[V(k+1)|\mathcal{F}_k\right] \leq &\left[1 - 2\frac{b_0}{\xi_{\max}}\alpha(k) + \lambda^2\alpha^2(k)\right]V(k) + \xi_{\max}\left(\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k] + \alpha^2(k)\|H\|^2\mathbb{E}[\|\omega(k)\|^2|\mathcal{F}_k]\right.\\ &\left.+\alpha^2(k)\|\mathcal{L}_{\mathcal{G}'} + HH^T\|^2\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k]\right)\end{aligned} \tag{19}$$

and furthermore, by Assumption A3 and $\sigma_k = \frac{\alpha(k-1)}{\epsilon}\delta H_{\max}$, implies that

$$\begin{aligned}\mathbb{E}\left[V(k+1)|\mathcal{F}_k\right] \leq &\left[1 - 2\frac{b_0}{\xi_{\max}}\alpha(k) + \lambda^2\alpha^2(k)\right]V(k) + \xi_{\max}\left(2(N - |\mathcal{A}|)\sigma_k^2\right.\\ &\left.+\alpha^2(k)(\|H\|^2\sigma_\omega^2 + 2\|\mathcal{L}_{\mathcal{G}'} + HH^T\|^2(N - |\mathcal{A}|)\sigma_k^2)\right)\\ \leq &\left[1 - 2\frac{b_0}{\xi_{\max}}\alpha(k) + \lambda^2\alpha^2(k)\right]V(k) + \xi_{\max}\Bigg(2(N - |\mathcal{A}|)\sigma_k^2\\ &+\alpha^2(k)(\|H\|^2\sigma_\omega^2 + 2\|\mathcal{L}_{\mathcal{G}'} + HH^T\|^2(N - |\mathcal{A}|)\sigma_k^2)\Big)\end{aligned}$$

$$\leq \left[1 - 2\frac{b_0}{\xi_{\max}}\alpha(k) + \lambda^2\alpha^2(k)\right] V(k) + \alpha^2(k)\xi_{\max}\left(\|H\|^2\sigma_\omega^2 \right.$$
$$\left. + \frac{2(N-|\mathcal{A}|)\delta^2 H_{\max}^2}{\epsilon^2} + \alpha^2(k)\|\mathcal{L}_{\mathcal{G}'} + HH^T\|^2 \frac{2(N-|\mathcal{A}|)\delta^2 H_{\max}^2}{\epsilon^2}\right). \tag{20}$$

Note that $b_0 > 0$, $\xi_{\max} > 0$, $\lim_{k\to\infty}\alpha(k) = 0$. Then, there exists $k_0 > 0$ such that $\lambda^2\alpha(k) \leq \frac{b_0}{\xi_{\max}}$ and $2\frac{b_0}{\xi_{\max}}\alpha(k) \leq 1$, $\forall k > k_0$. Thus, from Assumption A4, it follows that

$$0 \leq 1 - 2\frac{b_0}{\xi_{\max}}\alpha(k) + \lambda^2\alpha^2(k) < 1, \quad \forall k > k_0,$$
$$\sum_{k=k_0}^{\infty}\left[2\frac{b_0}{\xi_{\max}}\alpha(k) - \lambda^2\alpha^2(k)\right] \geq \frac{b_0}{\xi_{\max}}\sum_{k=k_0}^{\infty}\alpha(k) = \infty,$$
$$\lim_{k\to\infty}\frac{\alpha(k)\xi_{\max}\Pi}{2\frac{b_0}{\xi_{\max}} - \lambda^2\alpha(k)} = 0,$$

where $\Pi = \frac{2(N-|\mathcal{A}|)\delta^2 H_{\max}^2}{\epsilon^2} + \|H\|^2\sigma_\omega^2 + \alpha^2(k)\|\mathcal{L}_{\mathcal{G}'} + HH^T\|^2\frac{2(N-|\mathcal{A}|)\delta^2 H_{\max}^2}{\epsilon^2}$. By (20) and Lemma 6, it has $\lim_{k\to\infty}\mathbb{E}\left[V(k)\right] = 0$. From $0 \leq \delta^T(k)\delta(k) \leq \frac{V(k)}{\min_{i=1,\dots,N-|\mathcal{A}|}\xi_i}$, it is further obtained that $\lim_{k\to\infty}\mathbb{E}\left[\|\delta(k)\|^2\right] = 0$. ■

Most existing literature on multi-agent systems with noises/attacks only focuses on the mean square convergence. However, almost sure convergence and mean square convergence does not imply each other. Generally, the analysis of mean square convergence is easier than that of almost sure convergence since taking mean square yields a deterministic system. Besides, in many applications, the analysis of almost sure convergence is much more reasonable since people can only observe the trajectory of the network in one random experiment. In the following, based on nonnegative supermartingale convergence theorem, we establish the almost sure convergence of Algorithm 1.

Theorem 3. *If Assumptions A1′–A4 hold, then for any $\epsilon > 0$ and $\sigma_k = \frac{\alpha(k-1)}{\epsilon}\delta H_{\max}$, the estimate $x_i(k)$ given by (5) is convergent to θ^* almost surely, that is,*

$$\lim_{k\to\infty} x_i(k) = \theta^* \text{ a.s.}, \quad i \in \mathcal{V}/\mathcal{A}. \tag{21}$$

In addition, if $\alpha(k) \downarrow 0, k \to \infty$, then

$$\frac{1}{k}\sum_{\rho=0}^{k}\|\delta(\rho)\| = o((\alpha(k)k)^{-1/2}) \text{ a.s.}, \quad k \to \infty.$$

Proof. From Assumption A3 and the monotone convergence theorem,[47] we have

$$\mathbb{E}\sum_{k=0}^{\infty}\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k] = \sum_{k=0}^{\infty}\mathbb{E}[\|n(k)\|^2] < \infty,$$
$$\mathbb{E}\sum_{k=0}^{\infty}\alpha^2(k)\mathbb{E}[\|\omega(k)\|^2|\mathcal{F}_k] = \sum_{k=0}^{\infty}\alpha^2(k)\mathbb{E}[\|\omega(k)\|^2] < \infty,$$
$$\mathbb{E}\sum_{k=0}^{\infty}\alpha^2(k)\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k] = \sum_{k=0}^{\infty}\alpha^2(k)\mathbb{E}[\|n(k)\|^2] < \infty,$$

which implies that $\sum_{k=0}^{\infty}\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k] < \infty$, a.s., $\sum_{k=0}^{\infty}\alpha^2(k)\mathbb{E}[\|\omega(k)\|^2|\mathcal{F}_k]$, a.s., $\sum_{k=0}^{\infty}\alpha^2(k)\mathbb{E}[\|n(k)\|^2|\mathcal{F}_k]$, a.s. Thus, by (19) in the proof of Theorem 3 and nonnegative supermartingale convergence theorem,[46] $V(k)$ converges almost surely as $k \to \infty$, and

$$\sum_{k=0}^{\infty}\alpha(k)V(k) < \infty, \quad \text{a.s.}, \tag{22}$$

which together with Theorem 3 leads to (21). Moreover, by (22), Kronecker lemma[47] and Cauchy inequality, it follows that $\alpha(k) \to 0$ and $\frac{1}{k}\sum_{\rho=0}^{k}\|\delta(\rho)\| \leq (\alpha(k)k)^{-1/2}(\alpha(k)\sum_{\rho=0}^{k}\|\delta(\rho)\|^2)^{1/2} = o((\alpha(k)k)^{-1/2})$ a.s., $k \to \infty$. ■

Remark 8. Theorems 2 and 3 guarantee that any set of faulty agents that satisfies the F-local model has limited influence on the estimation errors of non-faulty agents whenever Algorithm 1 is used. This indicates the resilience of the algorithm to adversaries.

Theorems 2 and 3 are established based on Assumption A1′, which requires the reduced graph is balanced. To relax the requirement, Theorems 4 and 5 are given below.

Theorem 4. *If Assumptions A1–A4 hold, for any* $\epsilon > 0$ *and* $\sigma_k = \frac{\alpha(k-1)}{\epsilon}\delta H_{\max}$, *then the estimate* $x_i(k)$ *given by (5) is convergent to* θ^* *in mean square, that is,*

$$\lim_{k\to\infty}\mathbb{E}\left[\|\delta(k)\|^2\right] = 0.$$

Proof. From Assumption A1 and Lemma 1, it follows that the reduced graph $\mathcal{G}'$ consisting of non-faulty agents contains a directed spanning tree after removing $2F$ or fewer edges from each non-faulty agent. Without loss of generality, we assume that the Laplacian matrix $\mathcal{L}_{\mathcal{G}'}$ has the following form

$$\mathcal{L}_{\mathcal{G}'} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix},$$

where $L_{ii} \in \mathbb{R}^{r_i \times r_i}, i = 1, 2$, and $r_1 + r_2 = N - |\mathcal{A}|$. Since $\mathcal{G}'$ contains a spanning tree, the agents associated with L_{11} are all the roots in the graph, which implies that the directed subgraph associated with $L_{11} \in \mathbb{R}^{r_1 \times r_1}$ is strongly connected and $L_{22} \in \mathbb{R}^{r_2 \times r_2}$ is a nonsingular M-matrix. Set $\xi_i > 0, \forall i = 1, \ldots, r_1, \xi_i = 0, \forall i = r_1 + 1, \ldots, N - |\mathcal{A}|$, and $\sum_{i=1}^{r_1}\xi_i = 1$. Let $\Delta_1(k)$ be the column stack vector of $\delta_i(k), i = 1, \ldots, r_1$. Then, from Theorem 2, it follows that $\lim_{k\to\infty}\mathbb{E}\left[\|\Delta_1(k)\|^2\right] = 0$.

Let $\Delta_2(k)$ be the column stack vector of $\delta_i(k), i = r_1 + 1, \ldots, N - |\mathcal{A}|$. Then, from (13), it follows that

$$\begin{aligned}\Delta_2(k+1) = \Delta_2(k) - \alpha(k)[(\mathcal{L}_{22} + \overline{HH}^T)\Delta_2(k) + \mathcal{L}_{21}\Delta_1(k) \\ - \overline{H}\overline{\omega}(k) + (\mathcal{L}_{22} + \overline{HH}^T)\overline{n}_2(k) + \mathcal{L}_{21}\overline{n}_1(k)] + \overline{n}_2(k),\end{aligned} \tag{23}$$

where

$$\begin{aligned}\overline{H} &= \text{diag}\{H_{r_1+1}^T, \ldots, H_{N-|\mathcal{A}|}^T\}, \\ \overline{\omega}(k) &= \left[\omega_{r_1+1}^T(k), \ldots, \omega_{N-|\mathcal{A}|}^T(k)\right]^T, \\ \overline{n}_1(k) &= \left[n_1^T(k), \ldots, n_{r_1}^T(k)\right]^T, \\ \overline{n}_2(k) &= \left[n_{r_1+1}^T(k), \ldots, n_{N-|\mathcal{A}|}^T(k)\right]^T.\end{aligned}$$

By Lemma 4 and the fact that $\mathcal{L}_{22}$ is a nonsingular M-matrix, there exists $D = \text{diag}\{d_{r_1+1}, \ldots, d_{N-|\mathcal{A}|}\}$ with $d_i > 0, i = r_1 + 1, \ldots, N - |\mathcal{A}|$, such that $Q = DL_{22} + L_{22}^T D$ is positive definite.

Let $V_2(k) = \Delta_2^T(k)D\Delta_2(k)$. Then, from (23), it follows that

$$\begin{aligned}V_2(k+1) = \Delta_2^T(k)\Psi^T D\Psi\Delta_2(k) + \overline{W}^T(k)D\overline{W}(k) + \alpha^2(k)\Delta_1^T(k)\mathcal{L}_{21}^T D\mathcal{L}_{21}\Delta_1(k) \\ - 2\alpha(k)\Delta_2^T(k)\Psi^T D\mathcal{L}_{21}\Delta_1(k) + 2\Delta_2^T(k)\Psi^T D\overline{W}(k) - 2\alpha(k)\Delta_1^T(k)\mathcal{L}_{21}^T D\overline{W}(k),\end{aligned} \tag{24}$$

where $\Psi = I - \alpha(k)(\mathcal{L}_{22} + \overline{HH}^T)$, $\overline{W}(k) = \alpha(k)\left[\overline{H}\overline{\omega}(k) - (\mathcal{L}_{22} + \overline{HH}^T)\overline{n}_2(k) - \mathcal{L}_{21}\overline{n}_1(k)\right] - \overline{n}_2(k)$.

From $\mathbb{E}[\omega(k)] = 0$ and $\mathbb{E}[n(k)] = 0$, we have $\mathbb{E}\left[\overline{\omega}(k)|\mathcal{F}_k\right] = \mathbb{E}\left[\overline{n}_1(k)|\mathcal{F}_k\right] = \mathbb{E}\left[\overline{n}_2(k)|\mathcal{F}_k\right] = 0$, which further implies that

$$\mathbb{E}\left[2\Delta_2^T(k)\Psi^T D\overline{W}(k)|\mathcal{F}_k\right] = \mathbb{E}\left[2\alpha(k)\Delta_1^T(k)\mathcal{L}_{21}^T D\overline{W}(k)|\mathcal{F}_k\right] = 0. \tag{25}$$

From $Q > 0$ and $\overline{HH}^T D \geq 0$, it follows that $Q + 2\overline{HH}^T D > 0$. By $(\mathcal{L}_{22}^T + \overline{HH}^T)D(\mathcal{L}_{22} + \overline{HH}^T) \geq 0$, we have

$$\mathbb{E}\left[\Delta_2^T(k)\Psi^T D\Psi\Delta_2(k)|\mathcal{F}_k\right] \leq \left(1 - \alpha(k)\frac{\lambda_{\min}(Q)}{d_{\min}} + \alpha^2(k)\frac{\lambda_{\max}(\mathcal{L}_{22})}{d_{\min}}\right) V_2(k), \tag{26}$$

where $d_{\min} = \min_{r_1+1\leq i\leq N-|\mathcal{A}|} d_i$, $\lambda_{\min}(Q)$ is the smallest eigenvalue of $Q + 2\overline{HH}^T D$, $\lambda_{\max}(L_{22})$ is the maximum eigenvalue of $(\mathcal{L}_{22}^T + \overline{HH}^T)D(\mathcal{L}_{22} + \overline{HH}^T)$.

Similar to Theorem 2, there exists a positive constant C, such that

$$\mathbb{E}\left[\overline{W}^T(k)D\overline{W}(k)|\mathcal{F}_k\right] \leq \alpha^2(k)C. \tag{27}$$

From $\mathcal{L}_{21}^T D\mathcal{L}_{21} \geq 0$, we have

$$\mathbb{E}\left[\alpha^2(k)\Delta_1^T(k)\mathcal{L}_{21}^T D\mathcal{L}_{21}\Delta_1(k)|\mathcal{F}_k\right] \leq \alpha^2(k)\lambda_{\max}(L_{21})\mathbb{E}\left[\|\Delta_1(k)\|^2\right], \tag{28}$$

where $\lambda_{\max}(L_{21})$ is the maximum eigenvalue of $\mathcal{L}_{21}^T D\mathcal{L}_{21}$. By using $\mathcal{L}_{21}^T D^T\Psi\Psi^T D\mathcal{L}_{21} \geq 0$ and Young's inequality, we have

$$\begin{aligned}
&\mathbb{E}\left[-2\alpha(k)\Delta_2^T(k)\Psi^T D\mathcal{L}_{21}\Delta_1(k)|\mathcal{F}_k\right] \\
&\leq \alpha(k)\left[\frac{\lambda_{\min}(Q)}{2}\Delta_2^T(k)\Delta_2(k) + \frac{2}{\lambda_{\min}(Q)}\Delta_1^T(k)\mathcal{L}_{21}^T D^T\Psi\Psi^T D\mathcal{L}_{21}\Delta_1(k)\right] \\
&\leq \alpha(k)\left[\frac{\lambda_{\min}(Q)}{2d_{\min}}V_2(k) + \frac{2\overline{\lambda}_{\max}(\mathcal{L}_{21})}{\lambda_{\min}(Q)}\|\Delta_1(k)\|^2\right],
\end{aligned} \tag{29}$$

where $\overline{\lambda}_{\max}(\mathcal{L}_{21})$ is the maximum eigenvalue of $\mathcal{L}_{21}^T D^T\Psi\Psi^T D\mathcal{L}_{21}$.

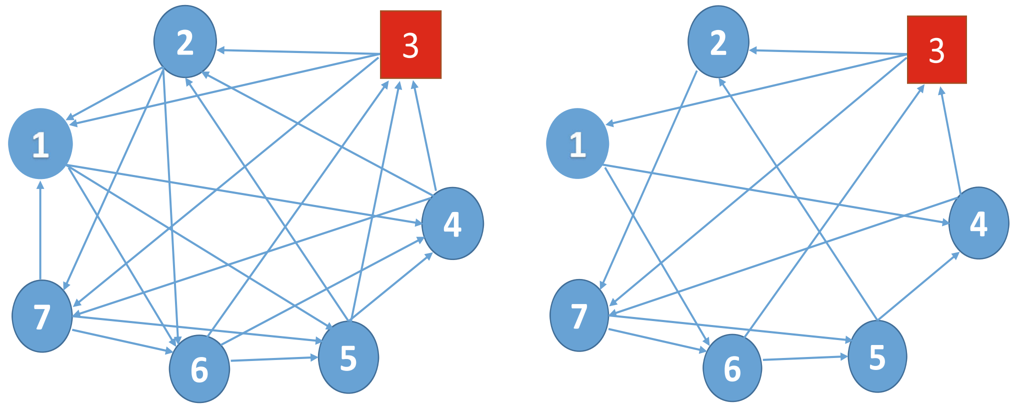
From (24)–(29), it follows that

$$\begin{aligned}
\mathbb{E}\left[V_2(k+1)|\mathcal{F}_k\right] \leq{}& \left(1 - \alpha(k)\frac{\lambda_{\min}(Q)}{2d_{\min}} + \alpha^2(k)\frac{\lambda_{\max}(\mathcal{L}_{22})}{d_{\min}}\right) V_2(k) \\
&+ \alpha^2(k)C + \alpha(k)\frac{2\overline{\lambda}_{\max}(\mathcal{L}_{21})}{\lambda_{\min}(Q)}\|\Delta_1(k)\|^2 \\
&+ \alpha^2(k)\lambda_{\max}(L_{21})\mathbb{E}\left[\|\Delta_1(k)\|^2\right].
\end{aligned}$$

Note that $\lambda_{\min}(Q) > 0$, $d_{\min} > 0$, $\lim_{k\to\infty}\alpha(k) = 0$. Then, there exists $k_0 > 0$ such that $\frac{\lambda_{\max}(\mathcal{L}_{22})}{d_{\min}}\alpha(k) \leq \frac{\lambda_{\min}(Q)}{4d_{\min}}$, and $\frac{\lambda_{\min}(Q)}{2d_{\min}}\alpha(k) \leq 1$, $\forall k > k_0$. Thus, from Assumption A4, it follows that

$$\begin{aligned}
&0 \leq 1 - \alpha(k)\frac{\lambda_{\min}(Q)}{2d_{\min}} + \alpha^2(k)\frac{\lambda_{\max}(\mathcal{L}_{22})}{d_{\min}} < 1, \quad \forall k > k_0, \\
&\sum_{k=k_0}^{\infty}\left[\alpha(k)\frac{\lambda_{\min}(Q)}{2d_{\min}} - \alpha^2(k)\frac{\lambda_{\max}(\mathcal{L}_{22})}{d_{\min}}\right] \geq \frac{\lambda_{\min}(Q)}{4d_{\min}}\sum_{k=k_0}^{\infty}\alpha(k) = \infty, \\
&\lim_{k\to\infty}\frac{\Xi}{\alpha(k)\frac{\lambda_{\min}(Q)}{2d_{\min}} - \alpha^2(k)\frac{\lambda_{\max}(\mathcal{L}_{22})}{d_{\min}}} = 0,
\end{aligned}$$

where $\Xi = \alpha(k)\frac{2\overline{\lambda}_{\max}(\mathcal{L}_{21})}{\lambda_{\min}(Q)}\|\Delta_1(k)\|^2 + \alpha^2(k)(C + \lambda_{\max}(L_{21})\mathbb{E}\left[\|\Delta_1(k)\|^2\right])$. From Lemma 6, it follows that $\lim_{k\to\infty}\mathbb{E}\left[\|\Delta_2(k)\|^2\right] = 0$. Thus, $\lim_{k\to\infty}\mathbb{E}\left[\|\delta(k)\|^2\right] = \lim_{k\to\infty}\mathbb{E}\left[\|\Delta_1(k)\|^2 + \|\Delta_2(k)\|^2\right] = 0$. ■



(A)

(B)

FIGURE 1 The topology of the sensor network. (A) $(2F+1)$-robust and (B) not $(2F+1)$-robust

Remark 9. In Theorem 4, we prove that the algorithm converges in mean square to an unbiased estimate of the unknown parameter. Different from the exponential decay Laplace noise,[36-38] the private noise in this article only needs to satisfy $\sigma_k = \frac{\alpha(k-1)}{\epsilon}\delta H_{\max}$, for example, $\sigma_k = \frac{1}{k^\gamma}, \frac{1}{2} < \gamma \leq 1$. Moreover, in contrast to the private noise σ_k being $O(\frac{1}{k})$,[34] the private noise in this article is more general.

Theorem 5. *If Assumptions A1–A4 hold, then for any $\epsilon > 0$ and $\sigma_k = \frac{\alpha(k-1)}{\epsilon}\delta H_{\max}$, the estimate $x_i(k)$ given by (5) is convergent to θ^* almost surely, that is,*

$$\lim_{k\to\infty} x_i(k) = \theta^* \text{ a.s.}, \quad i \in \mathcal{V}/\mathcal{A}.$$

In addition, if $\alpha(k) \downarrow 0, k \to \infty$, then

$$\frac{1}{k}\sum_{\rho=0}^{k}\|\delta(\rho)\| = o((\alpha(k)k)^{-1/2}) \text{ a.s.}, \quad k \to \infty. \tag{30}$$

Proof. Based on the nonnegative supermartingale convergence theorem and Theorem 4, the proof is similar to Theorem 3, and thus, is omitted here. ∎

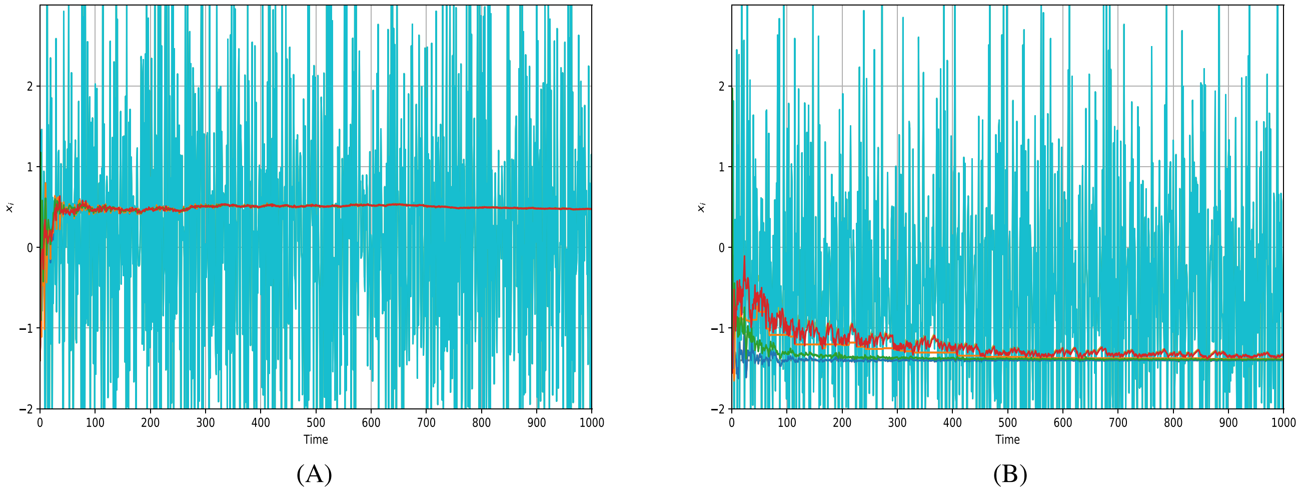
Remark 10. Theorem 5 implies that parameter estimates of agents converge asymptotically to their true values with probability one. This is different from the differentially private consensus.[36-39] In the differentially private consensus, the consensus value of agents is not deterministic but within a range with respect to the initials of agents. In contrast to the strongly connected directed graph and the row-stochastic weight matrix,[48,49] our article only needs the graph contains a directed spanning tree.

4 | SIMULATION EXAMPLE

In this section, a numerical simulation is given to verify the effectiveness of the algorithm. A multi-agent system of $N = 7$ is considered with one faulty agent over a digraph topology, as shown in Figure 1. The square(red) agent is the faulty agent, which sends the message $\hat{x}_3^j(k) = \cos(k) + \hat{n}_3^j(k)$, where $j \in \mathcal{N}_3^{\text{out}}$ and $\hat{n}_3^j(k) \sim \text{Lap}(\hat{\sigma}_{3,j}(k))$ with $\hat{\sigma}_{3,j}(k) = 1$. Different from the simulation cases in existing literature,[38] $\hat{x}_3^j(k)$ is not required to be exponentially decaying.

The true parameter is set as $\theta^* = 0.5$, the observation matrices and the initial parameter estimates of agents are given in the following forms:

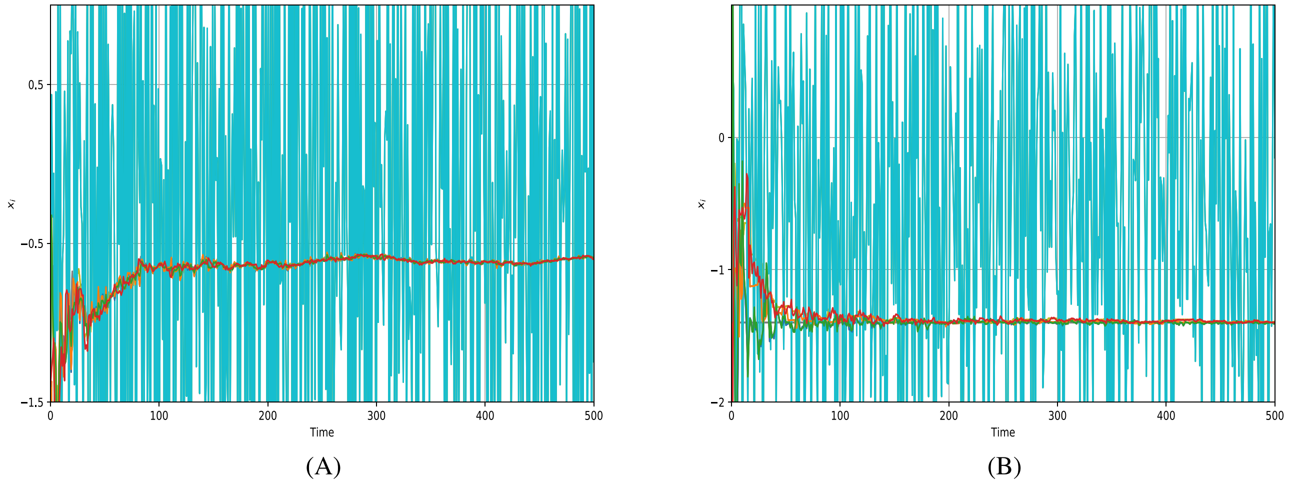
$$H_1 = H_3 = H_4 = H_5 = H_7 = \mathbf{0},\ H_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T,\ H_6 = \begin{bmatrix} 1 & 0 \end{bmatrix}^T,\ x_i(0) = -1.4,\ i = 1, \dots, 7.$$



(A)

(B)

FIGURE 2 The estimate of seven agents under the step-size given in Algorithm 1. The evolution of the faulty agent is depicted in solid green line. (A) $(2F+1)$-robust and (B) not $(2F+1)$-robust



(A)

(B)

FIGURE 3 The estimate of seven agents under the step-size given in the work of Fiore and Russo.[38] (A) $(2F+1)$-robust and (B) not $(2F+1)$-robust

$\omega_i(k)$ are independent white noises with uniform distribution on $[-0.2, 0.2]$, and the step size is $\alpha(k) = 5/(2k+2)$. The privacy level in Algorithm 1 is selected as $\epsilon = 0.1$, $\delta = 0.1$. Then, by the theoretical requirement, the scale parameter σ_k of the private noise $n_i(k)$ is $5/(2k)$, $k \geq 1$.

Under the above setting, we use Algorithm 1 to estimate the unknown parameter θ^*. If the communication topology is directed and 3-robust, then the estimate of non-faulty agents is mean square and almost sure convergence to the unknown parameter, which is confirmed in Figure 2A. However, when the directed graph is not 3-robust (see Figure 1B), as shown in Figure 2B, we cannot recover an unbiased estimate of the unknown parameter in mean square by using the algorithm. Thus, the simulation results are consistent with the theoretical analysis.

In addition, we compare the algorithm with different step sizes. If we use the step-size in the work of Fiore and Russo,[38] that is, the step size of the consensus term is a small constant, the simulation result is given in Figure 3. From Figure 3, it follows that the estimate of non-faulty agents does not converge to the unknown parameter. Thus, it shows that the step size in the work of Fiore and Russo[38] cannot be used to solve the problem considered in this article. Moreover, if we use the step-size in the work of Chen et al.,[15] that is, $\alpha(k) = 3.1 \times 10^{-2}$, the simulation result is given in Figure 4. From Figure 4, it also follows that the estimate of non-faulty agents does not converge to the unknown parameter. Thus, the algorithm used in this article is more general.

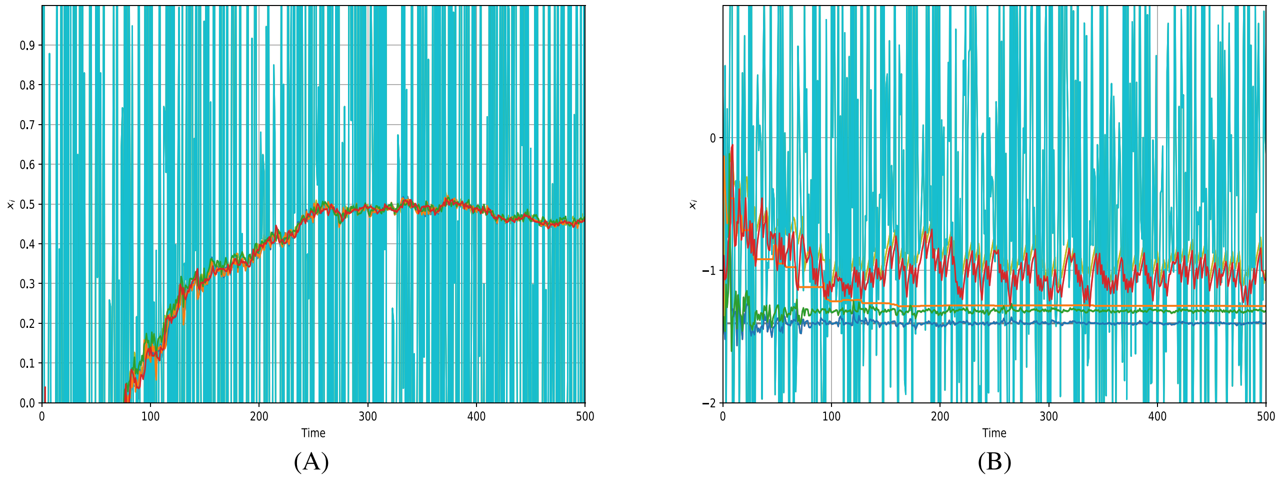


FIGURE 4 The estimate of seven agents under the step-size given in the work of Chen et al.[15] (A) $(2F + 1)$-robust and (B) not $(2F + 1)$-robust

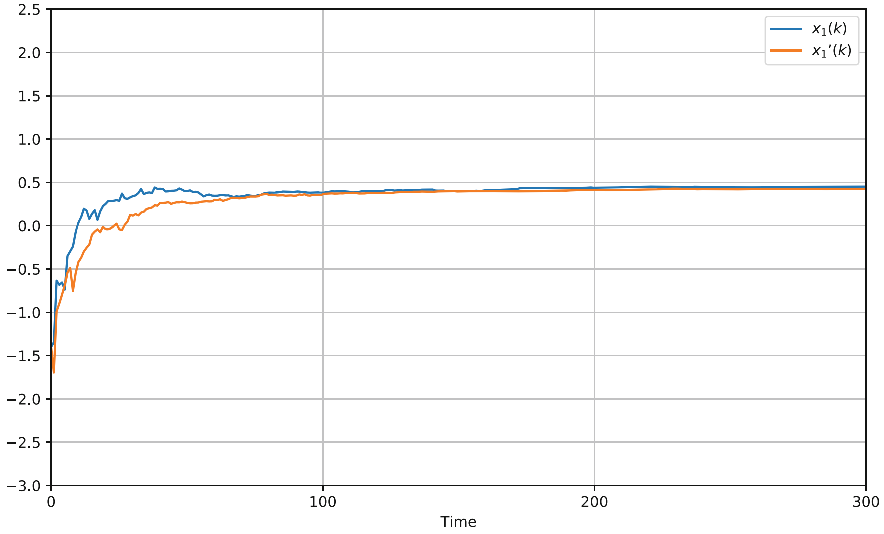


FIGURE 5 Outputs $x_1(k)$ and $x_1'(k)$ related to the adjacent relations

In the following, we verify the differential privacy properties of the algorithm. When $\{y_i(k), i = 1, 2, 4, 5, 6, 7\}$ and $\{y_i'(k), i = 1, 2, 4, 5, 6, 7\}$ are δ-adjacent, two outputs $x_1(k)$ and $x_1'(k)$ related to the adjacent relations are given in Figure 5. Figure 5 describes that the two outputs are fully fitted, which are almost indistinguishable from the eavesdropper. In this case, the sensitive information is protected, which is consistent with the theoretical analysis.

5 | CONCLUSION

In this article, the differentially private resilient DOE over digraphs has been studied. To protect the private data of each agent, the differentially private method is used. We design a differentially private resilient distributed online algorithm to estimate the unknown parameter, where additional noise is introduced by privacy protection. Moreover, we characterize the ϵ-differential privacy and convergence of the algorithm. Specifically, the differential privacy of the non-faulty agents is independent of the incorrect information that faulty agents deliver to the network. By using the stochastic approximation-type conditions, both mean square and almost sure convergence are proved regardless of the actions taken by the faulty agents and the introduced private noise. Finally, a simulation result is provided to verify the effectiveness and advantage of the algorithm. Future work will focus on how to design the optimal private noise and privacy budgets while guaranteeing the convergence of the algorithm.

ACKNOWLEDGMENTS

The work is supported by National Key R&D Program of China under Grant 2018YFA0703800 and National Natural Science Foundation of China (61877057, 62103156).

CONFLICT OF INTEREST

We declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

ORCID

Ji-Feng Zhang https://orcid.org/0000-0002-0656-2886

REFERENCES

1. Zhang Q, Zhang JF. Distributed parameter estimation over unreliable networks with Markovian switching topologies. *IEEE Trans Automat Contr*. 2012;57(10):2545-2560.
2. You K, Xie L, Song S. Asymptotically optimal parameter estimation with scheduled measurements. *IEEE Trans Signal Process*. 2013;61(14):3521-3531.
3. Kar S, Moura JM. Convergence rate analysis of distributed gossip (linear parameter) estimation: fundamental limits and tradeoffs. *IEEE J Sel Top Signal Process*. 2011;5(4):674-690.
4. Kar S, Moura JM, Ramanan K. Distributed parameter estimation in sensor networks: nonlinear observation models and imperfect communication. *IEEE Trans Inf Theory*. 2012;58(6):3575-3605.
5. Kar S, Moura JM, Poor HV. Distributed linear parameter estimation: asymptotically efficient adaptive strategies. *SIAM J Control Optim*. 2013;51(3):2200-2229.
6. Mei J, Ren W. Distributed parameter estimation under unreliable directed networks. Proceedings of the 2015 54th IEEE Conference on Decision and Control; 2015:4284-4289; IEEE.
7. Zong X, Li T, Zhang JF. Consensus control of second-order delayed multiagent systems with intrinsic dynamics and measurement noises. *Int J Robust Nonlinear Control*. 2018;28:5050-5070.
8. Huang M, Dey S, Nair G, Manton J. Stochastic consensus over noisy networks with Markovian and arbitrary switches. *Automatica*. 2010;46(10):1571-1583.
9. Li T, Zhang JF. Mean square average-consensus under measurement noises and fixed topologies: necessary and sufficient conditions. *Automatica*. 2009;45(8):1929-1936.
10. Li T, Zhang JF. Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises. *IEEE Trans Automat Contr*. 2010;55(9):2043-2057.
11. Yan L, Chen WS, Li C, Dai H, Zhu X, Gan ZX. Consensus-based distributed power control in power grids. *Sci China Inf Sci*. 2020;63(4):149202.
12. Chen Y, Kar S, Moura JMF. Resilient distributed parameter estimation with heterogeneous data. *IEEE Trans Signal Process*. 2019;67(19):4918-4933.
13. Chen Y, Kar S, Moura JMF. Resilient distributed estimation: sensor attacks. *IEEE Trans Automat Contr*. 2019;64(9):3772-3779.
14. Chen Y, Kar S, Moura JMF. Resilient distributed field estimation. *SIAM J Control Optim*. 2020;8(3):1429-1456.
15. Chen Y, Kar S, Moura JMF. Resilient distributed estimation through adversary detection. *IEEE Trans Signal Process*. 2018;66(9):2455-2469.
16. LeBlanc HJ, Hassan F. Resilient distributed parameter estimation in heterogeneous time-varying networks. Proceedings of the 2014 3rd International Conference on High Confidence Networked Systems; 2014:19-28.
17. LeBlanc HJ, Zhang H, Koutsoukos X, Sundaram S. Resilient asymptotic consensus in robust networks. *IEEE J Sel Areas Commun*. 2013;31(4):766-781.
18. Dibaji SM, Ishii H. Resilient consensus of second-order agent networks: asynchronous update rules with delays. *Automatica*. 2017;81:123-132.
19. Dibaji SM, Ishii H, Tempo R. Resilient randomized quantized consensus. *IEEE Trans Automat Contr*. 2018;63(8):2508-2522.
20. Sundaram S, Hadjicostis CN. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Trans Automat Contr*. 2011;56(7):1495-1508.
21. Sundaram S, Gharesifard B. Distributed optimization under adversarial nodes. *IEEE Trans Automat Contr*. 2019;64(3):1063-1076.
22. Su L, Shahrampour S. Finite-time guarantees for Byzantine-resilient distributed state estimation with noisy measurements. *IEEE Trans Automat Contr*. 2020;65(9):3758-3771.
23. Lu Y, Zhu MH. Privacy preserving distributed optimization using homomorphic encryption. *Automatica*. 2018;96:314-325.

24. Zhang JF, Tan JW, Wang JM. Privacy security in control systems. *Sci China Inf Sci*. 2021;64:176201:1-176201:3.
25. Mo YL, Murray RM. Privacy preserving average consensus. *IEEE Trans Automat Contr*. 2017;62(2):753-765.
26. Ruan M, Gao H, Wang YQ. Secure and privacy-preserving consensus. *IEEE Trans Automat Contr*. 2019;64(10):4035-4049.
27. Wang YQ. Privacy-preserving average consensus via state decomposition. *IEEE Trans Automat Contr*. 2019;64(11):4711-4716.
28. Dwork C. Differential privacy. Paper presented at: 2006 33rd International Colloquium on Automata, Languages and Programming 2006;1-12.
29. Yan CL, Ni ZY, Cao B, Lu RX, Wu SH, Zhang QY. UMBRELLA: user demand privacy preserving framework based on association rules and differential privacy in social networks. *Sci China Inf Sci*. 2019;62(3):039106.
30. Abadi M, Chu A, Goodfellow I, McMahan H, Mironov I, Talwar K, Zhang L. Deep learning with differential privacy. Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security; 2016:308-318.
31. Liang WJ, Chen H, Zhang J, Zhao D, Li CP. An effective scheme for top-k frequent itemsets mining under differential privacy conditions. *Sci China. Inf Sci*. 2020;63:159101:1-159101:3.
32. Li C, Zhou P, Xiong L, Wang Q, Wang T. Differentially private distributed online learning. *IEEE Trans Knowl Data En*. 2018;30(8):1440-1453.
33. Zhu JL, Xu CQ, Guan JF, Wu DO. Differentially private distributed online algorithms over time-varying directed networks. *IEEE Trans Signal Inf Process Netw*. 2018;4(1):4-17.
34. Han S, Topcu U, Pappas GJ. Differentially private distributed constrained optimization. *IEEE Trans Automat Contr*. 2017;62(1):50-64.
35. Lü QG, Liao XF, Xiang T, Li HQ, Huang TW. Privacy masking stochastic subgradient-push algorithm for distributed online optimization. *IEEE Trans Cybern*. 2020;51(6):3224-3237. doi:10.1109/TCYB.2020.2973221
36. Huang Z, Mitra S, Dullerud GE. Differentially private iterative synchronous consensus. Proceedings of the 2012 CCS Workshop on Privacy in the Electronic Society; 2012:81-90.
37. Nozari E, Tallapragada P, Cortes J. Differentially private average consensus: obstructions, trade-offs, and optimal algorithm design. *Automatica*. 2017;81:221-231.
38. Fiore D, Russo G. Resilient consensus for multi-agent systems subject to differential privacy requirements. *Automatica*. 2019;106:18-26.
39. Liu XK, Zhang JF, Wang JM. Differentially private consensus algorithm for continuous-time heterogeneous multi-agent systems. *Automatica*. 2020;122:109283.
40. Ny JL, Pappas GJ. Differentially private filtering. *IEEE Trans Automat Contr*. 2014;59(2):341-354.
41. Katewa V, Chakrabortty A, Gupta V. Differential privacy for network identification. *IEEE Trans Control Netw Syst*. 2020;7(1):266-277.
42. Liu Y, Liu J, Başar T. Differentially private gossip gradient descent. Proceedings of the 2018 IEEE Conference on Decision and Control; 2018:2777-2782.
43. Liu, Y, Zhang, X, Qin, S, Lei, X. Differentially private linear regression over fully decentralized datasets. Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS); 2019.
44. Ren W, Cao Y. *Distributed Coordination of Multi-agent Networks: Emergent Problems, Models, and Issues*. Springer-Verlag; 2011.
45. Berman A, Plemmons RJ. *Nonnegative Matrices in the Mathematical Sciences*. Academic Press, Inc.; 1979.
46. Goodwin G, Sin K. *Adaptive Filtering, Prediction and Control*. Prentice-Hall; 1984.
47. Chow Y, Teicher H. *Probability Theory: Independence, Interchangeability, Martingales*. Springer-Verlag; 1978.
48. Lü QG, Liao XF, Li HQ, Huang TW. A Nesterov-like gradient tracking algorithm for distributed optimization over directed networks. *IEEE Trans Syst Man Cybern Syst*. 2021;51(10):6258-6270.
49. Lü QG, Liao XF, Li HQ, Huang TW. Achieving acceleration for distributed economic dispatch in smart grids over directed networks. *IEEE Trans Netw Sci Eng*. 2020;7(3):1988-1999.

How to cite this article: Wang J, Zhang J-F, Liu X-K. Differentially private resilient distributed cooperative online estimation over digraphs. *Int J Robust Nonlinear Control*. 2022;1-19. doi: 10.1002/rnc.6303