Privacy-Preserving Distributed Estimation With Limited Data Rate

Jieming Ke, Jimin Wang, *Member, IEEE*, and Ji-Feng Zhang, *Fellow, IEEE*

***Abstract*—This article focuses on the privacy-preserving distributed estimation problem with a limited data rate, where the observations are the sensitive information. Specifically, a binary-valued quantizer-based privacy-preserving distributed estimation algorithm is developed, which improves the algorithm's privacy-preserving capability and simultaneously reduces the communication costs. The algorithm's privacy-preserving capability, measured by the Fisher information matrix, is dynamically enhanced over time. Notably, the Fisher information matrix of the output signals with respect to the sensitive information converges to zero at a polynomial rate, and the improvement in privacy brought by the quantizers is quantitatively characterized as a multiplicative effect. Regarding the communication costs, each sensor transmits only 1 bit of information to its neighbors at each time step. In addition, the assumption on the negligible quantization error for real-valued messages is not required. While achieving the requirements of privacy preservation and reducing communication costs, the algorithm ensures that its estimates converge almost surely to the true value of the unknown parameter by establishing a co-design guideline for the time-varying privacy noises and step sizes. A polynomial almost sure convergence rate is obtained, and then, the tradeoff between privacy and convergence rate is established. Numerical examples demonstrate the main results.**

***Index Terms*—Distributed estimation, Fisher information, limited data rate, privacy preservation.**

I. Introduction

DISTRIBUTED estimation has received close attention in the past decade due to its extensive applications in various fields, such as biological networks, online machine learning, and smart grids [1], [2]. Different from traditional centralized estimation, the observations of distributed estimation are collected by different sensors in the communication network. Therefore, network communication is required to fuse the observations from each sensor. However, in actual distributed systems, observations may contain sensitive information, and the network communication may lead to sensitive information leakage. For example, medical research usually requires clinical observation data of patients from different hospitals, which involves the patients' personal data [3], [4]. Motivated by this practical background, this article investigates how to achieve distributed estimation while preventing leakage of the observations.

The current literature offers several privacy-preserving methods for distributed systems. One of the methods is the homomorphic encryption method [5], [6], [7], [8], which provides high-dimensional security while ensuring control accuracy. Another commonly used method is the stochastic obfuscation method [9], [10], [11], [12], [13], [14], which has the advantages of low computational complexity and high timeliness. Other methods include the state decomposition method [15] and the privacy mask method [16]. Specifically, for the distributed estimation problem, Liu et al. [17] propose an observation perturbation differential privacy method, while the authors in [18], [19], and [20] give output perturbation differential privacy methods. The methods in [17], [18], [19], and [20] provide strong privacy, but their communication relies on the transmission of real-valued messages, which causes quantization errors and high communication costs when applied to digital networks based on quantized communications.

For distributed estimation problem under quantized communications, Kar et al. [1] propose a distributed estimation algorithm under infinite-level quantized communications. Under limited data rate, the authors in [21], [22], [23], [24], [25], [26], and [27] investigate the quantization methods following the biased compression rule [21]. The realization of limited data rate relies on an assumption on the negligible quantization error for real-valued messages. Without such an assumption, Sayin and Kozat [28] design a single-bit diffusion strategy under binary-valued communications, and Nassif et al. [29] propose a distributed estimation algorithm based on variable-rate quantizers. However, the algorithms' estimates in [28] and [29] do not converge to the true value.

To achieve privacy preservation and quantized communications simultaneously, quantizer-based privacy-preserving methods have recently received significant attention [30], [31],

Received 11 August 2025; revised 9 January 2026; accepted 26 January 2026. Date of publication 3 February 2026; date of current version 1 July 2026. This work was supported by the National Natural Science Foundation of China under Grant 62433020 and Grant T2293772. Recommended by Associate Editor T. Charalambous. *(Corresponding author: Ji-Feng Zhang.)*

Jieming Ke is with the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: kejieming@amss.ac.cn).

Jimin Wang is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China, and also with the Key Laboratory of Knowledge Automation for Industrial Processes, Ministry of Education, Beijing 100083, China (e-mail: jimwang@ustb.edu.cn).

Ji-Feng Zhang is with the School of Automation and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou 450007, China, and also with the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China (e-mail: jif@iss.ac.cn).

Digital Object Identifier 10.1109/TAC.2026.3660110

1558-2523 © 2026 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

[32], [33], [34]. For example, Chen et al. [30] propose a dynamic quantization-based homomorphic encryption method. For higher computational efficiency, Lang et al. [31] design special privacy noises and dither signals in dithered lattice quantizers, and Gao et al. [32] propose a dynamic coding scheme with Laplacian privacy noises. Both of them realize ϵ-differential privacy. The authors in [33] and [34] treat the dither signals in the quantizers as privacy noises and prove that using the dithered lattice quantizer (i.e., ternary quantizer in [33] and stochastic quantizer in [34]) can achieve $(0, \delta)$-differential privacy. Intuitively, the incorporation of quantizers increases the difficulty for adversaries to infer sensitive information, but existing works lack quantitative characterization of improvement in privacy brought by quantizers.

Building on the above excellent works, this article answers several key questions. How to simultaneously achieve privacy preservation, ensure a limited data rate, and guarantee the convergence of estimates for distributed estimation problems? How to quantitatively characterize the improvement in privacy brought by quantizers? And what is the tradeoff between privacy and convergence rate under our quantizer-based method for the distributed estimation problem?

To answer these questions, a novel binary-valued quantizer-based privacy-preserving distributed estimation algorithm is proposed. For quantized communications, our algorithm achieves message transmission at a limited data rate by using the comparison of adjacent binary-valued signals. Based on this technique, the biased compression rule [21] for quantizers can be avoided, and hence, our analysis does not rely on the assumption on the negligible quantization error for real-valued messages as in [21], [22], [23], [24], [25], [26], and [27], and the information receiver is not required to know the upper bounds of the estimate's norms to decode the quantized data as in [31] and [33]. For the privacy, dither signals in quantized communications [35], [36] are also treated as privacy noises. In addition, binary-valued quantizers also make sensitive information more difficult to infer.

To quantitatively characterize the improvement in privacy of our quantizer-based method, Fisher information is adopted as the privacy metric because of its following advantages. First, Fisher information is related to the Cramér–Rao lower bound and, thereby, can intuitively quantify the capability of potential adversaries to infer sensitive information. Hence, Fisher information has been adopted as the privacy metric for the privacy-preserving smart meters [37], the privacy-preserving database query [10], and privacy-preserving average consensus [38]. More importantly, compared to differential privacy, Fisher information provides a more precise measure of privacy. Differential privacy has difficulty in accurately characterizing the improvement of deterministic quantizers on privacy [32], whereas Fisher information can explicitly capture this improvement by reducing adversaries' estimation accuracy for the sensitive information.

By using Fisher information, the binary-valued quantizer-based privacy-preserving distributed estimation algorithm is shown to achieve a dynamically enhanced privacy. The Fisher information matrix of the output signals with respect to the sensitive information converges to zero at a polynomial rate. The dynamically enhanced privacy can be achieved because under our algorithm, the privacy noises can be constant or even increasing, in contrast to the decaying ones in existing privacy-preserving distributed estimation algorithms [18], [19]. Furthermore, dynamically enhanced privacy can be used to reveal the tradeoff between privacy and convergence rate. When privacy is enhanced at a higher rate, the convergence rate will decrease.

This article proposes a novel binary-valued quantizer-based privacy-preserving distributed estimation algorithm. The main contributions of this article are summarized as follows.

1) The improvement in privacy brought by the quantizers has been quantitatively characterized. Specifically, under Gaussian privacy noises, the introduction of binary-valued quantizers can improve the privacy-preserving level by at least $\frac{\pi}{2}$ times, which reveals the impact of quantizers on the privacy-preserving level as a multiplicative effect.
2) The privacy-preserving capability of the proposed algorithm is dynamically enhanced. The Fisher information matrix of the output signals with respect to the sensitive information converges to zero at a polynomial rate. Notably, the privacy analytical framework is unified for general privacy noise types, including Gaussian, Laplacian, and even heavy-tailed ones.
3) Under the proposed algorithm, each sensor transmits only 1 bit of information to its neighbors at each time step. This is the lowest data rate among existing quantizer-based privacy-preserving distributed algorithms [30], [31], [32], [33]. In addition, the assumption on the negligible quantization error for real-valued messages [21], [22], [23], [24], [25], [26], [27] is not required.
4) A co-design guideline for the time-varying privacy noises and step sizes under quantized communications is proposed to ensure the almost sure convergence of the algorithm. A polynomial almost sure convergence rate is also obtained.
5) The tradeoff between privacy and convergence rate is established. Better privacy implies a slower convergence rate, and vice versa. Furthermore, the sensor operators can determine their own preference for the privacy and convergence rate by properly selecting privacy noises and step sizes.

The rest of this article is organized as follows. Section II formulates the problem and introduces the Fisher information-based privacy metric. Section III proposes our privacy-preserving distributed estimation algorithm. Section IV analyzes the privacy-preserving capability of the algorithm. Section V proves the almost sure convergence of the algorithm and calculates the almost sure convergence rate. Section VI establishes the tradeoff between privacy and convergence rate. Section VII uses numerical examples to demonstrate the main results. Finally, Section VIII concludes this article.

I. Notations

In the rest of this article, $\mathbb{N}$, $\mathbb{R}$, $\mathbb{R}^n$, and $\mathbb{R}^{n\times m}$ are the sets of natural numbers, real numbers, n-dimensional real vectors, and $n \times m$-dimensional real matrices, respectively. $\|x\|$ is the

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

Euclidean norm for vector x, and $\|A\|$ is the induced matrix norm for matrix A. A^+ is the pseudoinverse of matrix A. I_n is an $n \times n$ identity matrix. $\mathbb{I}_{\{\cdot\}}$ denotes the indicator function, whose value is 1 if its argument (a formula) is true, and 0, otherwise. $\mathbf{1}_n$ is the n-dimensional vector whose elements are all ones. $\mathrm{diag}\{\cdot\}$ denotes the block matrix formed in a diagonal manner of the corresponding numbers or matrices. $\mathrm{col}\{\cdot\}$ denotes the column vector stacked by the corresponding vectors. $\otimes$ denotes the Kronecker product. $\mathcal{N}(0,\sigma^2)$, $Lap(0,b)$, and $Cauchy(0,r)$ represent Gaussian distribution with density function $\frac{1}{\sqrt{2\pi}\sigma}\exp(-x^2/2\sigma^2)$, Laplacian distribution with density function $\frac{1}{2b}\exp(-|x|/b)$, and Cauchy distribution with the density function $1/\left(\pi r\left[1+(x/r)^2\right]\right)$, respectively.

II. Preliminaries and Problem Formulation

A. Preliminaries on Graph Theory

In this article, the communication graph is switching among topology graphs $\mathcal{G}^{(1)},\ldots,\mathcal{G}^{(M)}$, where $\mathcal{G}^{(u)}=(\mathcal{V},\mathcal{E}^{(u)},\mathcal{A}^{(u)})$ for all $u=1,\ldots,M$. $\mathcal{V}=\{1,\ldots,N\}$ is the set of the sensors. $\mathcal{E}^{(u)}\in\{(i,j):i,j\in\mathcal{V}\}$ is the edge set. $\mathcal{A}^{(u)}=(a_{ij}^{(u)})_{N\times N}$ represents the symmetric weighted adjacency matrix of the graph whose elements are all nonnegative. $a_{ij}^{(u)}>0$ if and only if $(i,j)\in\mathcal{E}^{(u)}$. Besides, $\mathcal{N}_i^{(u)}=\{j:(i,j)\in\mathcal{E}^{(u)}\}$ denotes the sensor i's neighbor set corresponding to the graph $\mathcal{G}^{(u)}$. Define Laplacian matrix as $\mathcal{L}^{(u)}=\mathcal{D}^{(u)}-\mathcal{A}^{(u)}$, where $\mathcal{D}^{(u)}=\mathrm{diag}\left(\sum_{i\in\mathcal{N}_1}a_{i1}^{(u)},\ldots,\sum_{i\in\mathcal{N}_N}a_{iN}^{(u)}\right)$.

The union of $\mathcal{G}^{(1)},\ldots,\mathcal{G}^{(M)}$ is denoted by $\mathcal{G}=(\mathcal{V},\mathcal{E},\mathcal{A})$, where $\mathcal{E}=\bigcup_{r=1}^{M}\mathcal{E}^{(u)}$, and $\mathcal{A}=\sum_{u=1}^{M}\mathcal{A}^{(u)}$. Besides, set $\mathcal{N}_i=\{j:(i,j)\in\mathcal{E}\}$.

Assumption 1: The union graph $\mathcal{G}$ is connected.

Remark 1: Instead of requiring instantaneous connectivity at each time step in [21], [33], and [39], Assumption 1 only requires the joint connectivity of the switching topologies $\mathcal{G}^{(1)},\ldots,\mathcal{G}^{(M)}$ over time.

The communication graph at time k, denoted by $\mathtt{G}_k$, is associated with a homogeneous Markovian chain $\{\mathtt{m}_k:k\in\mathbb{N}\}$ with a state space $\{1,\ldots,M\}$, transition probability $p_{uv}=\mathbb{P}\{\mathtt{m}_k=v|\mathtt{m}_{k-1}=u\}$, and stationary distribution $\pi_u=\lim_{k\to\infty}\mathbb{P}\{\mathtt{m}_k=u\}$. If $\mathtt{m}_k=u$, then $\mathtt{G}_k=\mathcal{G}^{(u)}$. Denote $q_{ij,k}=\mathbb{P}\{(i,j)\in\mathcal{E}^{(\mathtt{m}_k)}\}$. For convenience, $\mathcal{E}^{(\mathtt{m}_k)}$, $a_{ij}^{(\mathtt{m}_k)}$, $\mathcal{N}_i^{(\mathtt{m}_k)}$, and $\mathcal{L}^{(\mathtt{m}_k)}$ are abbreviated as $\mathtt{E}_k$, $\mathtt{a}_{ij,k}$, $\mathtt{N}_{i,k}$, and $\mathtt{L}_k$, respectively, in the rest of this article.

Remark 2: Markovian switching graphs can be used to model the link failures [40], [41]. $\mathtt{a}_{ij,k}>0$ implies that the communication link between the sensors i and j is normal. $\mathtt{a}_{ij,k}=0$ implies that the communication link fails.

Remark 3: Given $p_{u,1}=\mathbb{P}\{\mathtt{G}_1=\mathcal{G}^{(u)}\}$, $q_{ij,k}$ can be recursively obtained by $q_{ij,k}=\sum_{u\in\mathbb{G}_{ij}}p_{u,k}$, $p_{u,k+1}=\mathbb{P}\{\mathtt{G}_{k+1}=\mathcal{G}^{(u)}\}=\sum_{v=1}^{M}p_{v,k}p_{vu}$, where $\mathbb{G}_{ij}=\{u:(i,j)\in\mathcal{E}^{(u)}\}$. By [56, Thm. 1.2], we have $q_{ij,k}=\sum_{u\in\mathbb{G}_{ij}}\pi_u+O(\lambda_p^k)$ for some $\lambda_p\in(0,1)$. Specifically, when the initial distribution $\{p_{u,1}:u=1,\ldots,M\}$ is the stationary distribution $\{\pi_u:u=1,\ldots,M\}$, we have $q_{ij,k}=\sum_{u\in\mathbb{G}_{ij}}\pi_u$.

B. Observation Model

In the multisensor system coupled by the Markovian switching graphs, the sensor i observes the unknown parameter $\theta\in\mathbb{R}^n$ from the observation model

$$\mathtt{y}_{i,k}=\mathtt{H}_{i,k}\theta+\mathtt{w}_{i,k},\ i=1,\ldots,N,\ k\in\mathbb{N} \tag{1}$$

where θ is the unknown parameter, k is the time index, $\mathtt{w}_{i,k}\in\mathbb{R}^{m_i}$ is the observation noise, and $\mathtt{y}_{i,k}\in\mathbb{R}^{m_i}$ is the observation. $\mathtt{H}_{i,k}\in\mathbb{R}^{m_i\times n}$ is the random measurement matrix.

Assumptions for the observation model (1) are given as follows.

Assumption 2: The random measurement matrix $\mathtt{H}_{i,k}$ is not necessarily available, but its mean value $\bar{H}_i$ is known. $\sum_{i=1}^{N}\bar{H}_i^\top\bar{H}_i$ is invertible.

Remark 4: The invertibility on $\sum_{i=1}^{N}\bar{H}_i^\top\bar{H}_i$ is the cooperative observability assumption [1], [42], [43]. In addition, Kar et al. [1] used the unknown $\mathtt{H}_{i,k}$ to model sensor failure. Under Assumption 2, the subsystem of each sensor is not necessarily observable. $\bar{H}_i$ can be even 0 for some sensor i. Hence, communications between sensors are necessary to fuse data collected by different sensors.

Assumption 3: $\{\mathtt{w}_{i,k},\mathtt{H}_{i,k}:i\in\mathcal{V},k\in\mathbb{N}\}$ is an independent sequence[1] such that

$$\mathbb{E}\mathtt{w}_{i,k}=0,\ \sup_{i\in\mathcal{V},\ k\in\mathbb{N}}\mathbb{E}\left\|\mathtt{w}_{i,k}\right\|^{\rho}<\infty \tag{2}$$

$$\sup_{i\in\mathcal{V},\ k\in\mathbb{N}}\mathbb{E}\left\|\mathtt{H}_{i,k}-\bar{H}_i\right\|^{\rho}<\infty \tag{3}$$

for some $\rho>2$, and independent of $\{\mathtt{G}_k:k\in\mathbb{N}\}$.

Remark 5: If ρ in (2) and (3) takes different values, for example, ρ_1 and ρ_2, respectively, then by Lyapunov inequality [44], (2) and (3) still hold for $\rho=\min\{\rho_1,\rho_2\}$.

C. Dynamically Enhanced Privacy

This section will formulate the privacy-preserving distributed estimation problem. Notably, in some medical research [4], the observation $\mathtt{y}_{i,k}$ is the private clinical observation data held by different hospitals. Such privacy scenarios motivate us to protect the observation $\mathtt{y}_{i,k}$.

The set containing all the information transmitted in the network is denoted as $\mathtt{S}=\{\mathtt{s}_{ij,k}:(i,j)\in\mathtt{E}_k,k\in\mathbb{N}\}$, where $\mathtt{s}_{ij,k}$ is the signal that the sensor i transmits to the sensor j at time k. Then, we introduce Fisher information as a privacy metric to quantify the privacy-preserving capability.

Definition 1 (Fisher information, [45]): Fisher information of $\mathtt{S}$ with respect to sensitive information $\mathtt{y}$ is defined as

$$\mathcal{I}_{\mathtt{S}}(\mathtt{y})=\mathbb{E}\left[\left[\frac{\partial\ln(\mathbb{P}(\mathtt{S}|\mathtt{y}))}{\partial\mathtt{y}}\right]\left[\frac{\partial\ln(\mathbb{P}(\mathtt{S}|\mathtt{y}))}{\partial\mathtt{y}}\right]^\top\middle|\mathtt{y}\right].$$

Given a random variable $\mathtt{x}$, the conditional Fisher information is defined as

$$\mathcal{I}_{\mathtt{S}}(\mathtt{y}|\mathtt{x})=\mathbb{E}\left[\left[\frac{\partial\ln(\mathbb{P}(\mathtt{S}|\mathtt{x},\mathtt{y}))}{\partial\mathtt{y}}\right]\left[\frac{\partial\ln(\mathbb{P}(\mathtt{S}|\mathtt{x},\mathtt{y}))}{\partial\mathtt{y}}\right]^\top\middle|\mathtt{y}\right].$$

[1] A random variable sequence is said to be independent if any pair of random variables in the sequence are independent of each other.

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

Fisher information can be used to quantify the privacy-preserving capability because of the following proposition.

Proposition 1 (Cramér–Rao lower bound, [45]): If $\mathcal{I}_{\mathtt{S}}(\mathtt{y})$ is invertible, then for any unbiased estimator $\hat{\mathtt{y}} = \hat{\mathtt{y}}(\mathtt{S})$ of $\mathtt{y}$, $\mathbb{E}(\hat{\mathtt{y}} - \mathtt{y})(\hat{\mathtt{y}} - \mathtt{y})^\top \geq \mathcal{I}_{\mathtt{S}}^{-1}(\mathtt{y})$.

Remark 6: Fisher information is a natural privacy metric [10], [37], [38], because by Proposition 1, smaller $\mathcal{I}_{\mathtt{S}}(\mathtt{y})$ implies less information leaks, and vice versa. Besides, Fisher information is closely related to other common privacy metrics. For example, Barnes et al. [46] revealed the positive correlation between ϵ-differential privacy and upper bounds of Fisher information. There are other advantages for Fisher information as the privacy metric. First, compared to mutual information [47], [48], Fisher information is unrelated to the a priori knowledge on the sensitive information, and hence, it is suitable for privacy-preserving distributed estimation where the distribution of the sensitive information $\mathtt{y}_{i,k}$ contains the unknown parameter θ. Second, compared to maximal leakage [49], Fisher information allows $\mathtt{y}_{i,k}$ to be both continuous and discrete. Third, compared to differential privacy [32], Fisher information has advantage in characterizing the improvement in privacy brought by deterministic quantizers.

Under the above privacy metric, we aim to achieve the dynamically enhanced privacy as defined below.

Definition 2: If the privacy-preserving capability of an algorithm is said to be dynamically enhanced, then given any $i \in \mathcal{V}$ and k with $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k}) > 0$, there exists $T > k$ such that $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,t}) < \mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$ for all $t \geq T$.

Remark 7: By Lemma A.1 in Appendix A, $\lim_{k\to\infty} \mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k}) = 0$ is sufficient for the dynamically enhanced privacy.

D. Problem of Interest

This article mainly seeks to develop a new privacy-preserving distributed estimation algorithm, which can simultaneously achieve the following.

1) The privacy is dynamically enhanced over time.
2) The sensor i transmits only 1 bit of information to its neighbor j at each time step.
3) The estimates for all sensors converge to the true value of the unknown parameter almost surely.

III. Privacy-Preserving Algorithm Design

This section will first give the binary-valued quantizer-based method and then propose a binary-valued quantizer-based privacy-preserving distributed estimation algorithm.

The traditional consensus+innovations type distributed estimation algorithms [1], [50] fuse the observations through the transmission of estimates $\hat{\theta}_{i,k-1}$, which would lead to sensitive information leakage. For the privacy issue, the following binary-valued quantizer-based method is designed to transform them into binary-valued signals before transmission. First, if $k = nq + l$ for some $q \in \mathbb{N}$ and $l \in \{1, \ldots, n\}$, then the sensor i generates φ_k as the n-dimensional vector whose lth element is 1 and the others are 0. The sensor i uses φ_k to compress the previous local estimate $\hat{\theta}_{i,k-1}$ into the scalar $\mathtt{x}_{i,k} = \varphi_k^\top \hat{\theta}_{i,k-1}$.

Algorithm 1: Binary-Valued Quantizer-Based Privacy-Preserving Distributed Estimation Algorithm.

Input: initial estimate sequence $\{\hat{\theta}_{i,0}\}$, threshold sequence $\{C_{ij}\}$ with $C_{ij} = C_{ji}$, noise distribution sequence $\{F_{ij,k}(\cdot)\}$ with $F_{ij,k}(\cdot) = F_{ji,k}(\cdot)$, step-size sequences $\{\alpha_{ij,k}\}$ with $\alpha_{ij,k} = \alpha_{ji,k} > 0$ and $\{\beta_{i,k}\}$ with $\beta_{i,k} > 0$.
Output: estimate sequence $\{\hat{\theta}_{i,k}\}$.
for $k = 1, 2, \ldots,$ **do**
Privacy preservation: Use the binary-valued quantizer-based method (4) to transform the previous local estimate $\hat{\theta}_{i,k-1}$ into the binary-valued signal $\mathtt{s}_{ij,k}$, and send the binary-valued signal $\mathtt{s}_{ij,k}$ to the neighbor j.
Information fusion: Fuse neighborhood information.

$$\check{\theta}_{i,k} = \hat{\theta}_{i,k-1} + \varphi_k \sum_{j\in\mathtt{N}_{i,k}} \alpha_{ij,k}\mathtt{a}_{ij,k}\left(\mathtt{s}_{ij,k} - \mathtt{s}_{ji,k}\right).$$

Estimate update: Use $\mathtt{y}_{i,k}$ to update the local estimate.

$$\hat{\theta}_{i,k} = \check{\theta}_{i,k} + \beta_{i,k}\bar{H}_i^\top\left(\mathtt{y}_{i,k} - \bar{H}_i\hat{\theta}_{i,k-1}\right)$$

where $\bar{H}_i = \mathbb{E}\mathtt{H}_{i,k}$ as in Assumption 2.
end for

Second, the sensor i generates the privacy noise $\mathtt{d}_{ij,k}$ with distribution $F_{ij,k}(\cdot)$ for all $j \in \mathtt{N}_{i,k}$. Then, given the threshold C_{ij}, the sensor i generates the binary-valued signal

$$\mathtt{s}_{ij,k} = \begin{cases} 1, & \text{if } \mathtt{x}_{i,k} + \mathtt{d}_{ij,k} \leq C_{ij} \\ -1, & \text{otherwise.} \end{cases} \tag{4}$$

Remark 8: The threshold C_{ij} can be any real number. From the communication perspective, when $\mathtt{d}_{ij,k}$ is unbiased, the communication efficiency of the quantizer (4) will be better if C_{ij} gets closer to $\mathtt{x}_{i,k}$ [36]. Since $\mathtt{x}_{i,k}$ should converge to $\varphi_k^\top\theta$, we can set $C_{ij} = \arg\min_C \mathbb{E}_{\theta\sim p(\theta)}\|\theta - C\mathbf{1}_n\|^2 = \sum_{i=1}^n \mu_{\theta,i}$, where $p(\theta)$ is the a priori distribution for θ, $\mu_{\theta,i}$ is the ith element of $\mathbb{E}_{\theta\sim p(\theta)}\theta$, and $\mathbf{1}_n$ is the n dimensional vector with all elements being 1.

By using the binary-valued quantizer-based method (4), a novel privacy-preserving distributed estimation algorithm is proposed in Algorithm 1.

Remark 9: In Algorithm 1, the expectation of $\mathtt{s}_{ij,k} - \mathtt{s}_{ji,k}$ is $F(C_{ij} - \varphi_k^\top\hat{\theta}_{i,k-1}) - F(C_{ij} - \varphi_k^\top\hat{\theta}_{j,k-1})$, which captures the difference in the statistical properties of the quantized data under different estimates. Therefore, when $F(\cdot)$ is strictly monotonically increasing, this communication mechanism is able to achieve information fusion of quantized data over a distributed network without relying on controlling an upper bound on the quantization error. This communication mechanism has the following two advantages.

1) It does not rely on the assumption that certain real-valued messages can be transmitted with negligible quantization error as required by many existing stochastic compression methods [21], [22], [23], [24], [25], [26], [27]. Compression methods in these works satisfy the biased compression rule

$$\sqrt{\mathbb{E}\left[\|\mathtt{Q}(\mathtt{x}) - \mathtt{x}\|^2|\mathtt{x}\right]} \leq \kappa\|\mathtt{x}\| + \iota \tag{5}$$

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

where $\mathtt{Q}(\cdot)$ is the stochastic compression operator, and $\kappa \in [0,1)$, $\iota \geq 0$. However, when the decoder does not know the a priori upper bound of $\|\mathtt{x}\|$, under (5), $\mathtt{Q}(\cdot)$ cannot compress a real-valued vector into finite bits. This is because under finite-level quantizer $\mathtt{Q}(\cdot)$, $\mathtt{Q}(\mathtt{x})$ is uniformly bounded, leading to

$$\lim_{\|\mathtt{x}\|\to\infty} \frac{\sqrt{\mathbb{E}\left[\|\mathtt{Q}(\mathtt{x})-\mathtt{x}\|^2|\mathtt{x}\right]}}{\|\mathtt{x}\|} = 1 > \kappa$$

which is contradictory to (5). Due to this limitation, the authors in [21], [22], [23], [24], [25], [26], and [27] assume that certain real-valued messages can be transmitted with negligible quantization error. The quantizer (4) does not follow (5), and hence, it does not require this assumption.

2) It does not require a decoder, thereby reducing the computational complexity. For comparison, dynamic quantization method [30], [32], empirical measurement method [51], and recursive projection method [52] require decoding the quantized data into approximate values of the original estimates before they can be used for information fusion. In contrast, Algorithm 1 avoids the decoder using $\mathtt{s}_{ij,k} - \mathtt{s}_{ji,k}$, and hence, it achieves a similar computational complexity with the traditional ones without quantization [18], [41].

Remark 10: Algorithm 1 does not rely on the value of $\mathtt{H}_{i,k}$ due to Assumption 2. Therefore, under Algorithm 1, preserving the privacy of $\mathtt{y}_{i,k}$ inherently ensures the privacy of $\mathtt{H}_{i,k}$.

Remark 11: Algorithm 1 achieves a data rate of 1 bit per time. Such an extremely low data rate can be achieved because it does not rely on transmission of real-valued messages, as shown in Remark 9, and the introduction of φ_k makes the data rate irrelevant of the dimension of the estimates. For comparison, the authors in [31] and [33] require no less than n and $\lceil n\log_2 3\rceil$ bits per time, respectively. When not pursuing such an extremely low data rate, φ_k can be replaced by a periodic selection matrix $\Psi_k \in \mathbb{R}^{\psi\times n}$. Under this replacement, the sensor generates a ψ-bit $s_{ij,k}$ by applying independent privacy noise and binary-valued quantization to each element of the vector $\mathtt{x}_{i,k} = \Psi_k^\top \hat{\theta}_{i,k-1}$. This approach improves the estimation accuracy with data rate of ψ bits per time.

Assumptions for privacy noises and step sizes in Algorithm 1 are given as follows.

Assumption 4: The privacy noise sequence $\{\mathtt{d}_{ij,k} : (i,j) \in \mathcal{E}, k \in \mathbb{N}\}$ satisfies

i) the density function $f_{ij,k}(\cdot)$ of $\mathtt{d}_{ij,k}$ exists;

ii) $\eta_{ij,k} = \sup_{x\in\mathbb{R}} \frac{f_{ij,k}^2(x)}{F_{ij,k}(x)(1-F_{ij,k}(x))} < \infty$;

iii) there exists a sequence $\{\zeta_{ij,k}\}$ such that for all compact set $\mathcal{X}$, $\inf_{(i,j)\in\mathcal{E},k\in\mathbb{N},x\in\mathcal{X}} \frac{f_{ij,k}(x)}{\zeta_{ij,k}} > 0$;

iv) $\{\mathtt{d}_{ij,k} : (i,j) \in \mathcal{E}, k \in \mathbb{N}\}$ is an independent sequence and independent of $\{\mathtt{w}_{i,k}, \mathtt{G}_k, \mathtt{H}_{i,k} : i \in \mathcal{V},\ k \in \mathbb{N}\}$.

Remark 12: The Assumption 4 ii) is for the privacy analysis, and iii) is for the convergence analysis.

Assumption 5: The step-size sequences $\{\alpha_{ij,k} : (i,j) \in \mathcal{E}, k \in \mathbb{N}\}$ and $\{\beta_{i,k} : i \in \mathcal{V}, k \in \mathbb{N}\}$ satisfy

i) $\sum_{k=1}^{\infty} \alpha_{ij,k}^2 < \infty$ and $\alpha_{ij,k} = O(\alpha_{ij,k+1})$ for all $(i,j) \in \mathcal{E}$;

ii) $\sum_{k=1}^{\infty} \beta_{i,k}^2 < \infty$ and $\beta_{i,k} = O(\beta_{i,k+1})$ for all $i \in \mathcal{V}$;

iii) $\sum_{k=1}^{\infty} z_k = \infty$ for $z_k = \min\{\alpha_{ij,k}\zeta_{ij,k} : (i,j) \in \mathcal{E}\} \cup \{\beta_{i,k} : i \in \mathcal{V}\}$ and $\zeta_{ij,k}$ follows Assumption 4 iii).

Remark 13: Assumption 5 is the stochastic approximation condition for distributed estimation. Such step sizes are typically set as polynomial functions of k [39]. When the step sizes are all polynomial, Assumption 5 iii) is equivalent to $\sum_{k=1}^{\infty} \alpha_{ij,k}\zeta_{ij,k} = \infty$ for all $(i,j) \in \mathcal{E}$ and $\sum_{k=1}^{\infty} \beta_{i,k} = \infty$ for all $i \in \mathcal{V}$. In Assumption 5, the step sizes are not necessarily the same for all sensors, in contrast to the centralized step sizes adopted in many distributed algorithms [1], [5], [9]. Therefore, the sensor operators can properly select their step sizes based on their own requirements.

Remark 14: Assumptions 4 and 5 indicate that Algorithm 1 establishes a time-varying design method for privacy noises and step sizes under quantized communications upon the algorithm framework of [39].

i) Diversified design of privacy noises is allowed to meet different privacy-preserving requirements. By [36, Lemma 5.3] and Lemmas B.1–B.3 in Appendix B, Assumption 4 accommodates not only standard differential privacy noises like Laplacian and Gaussian ones but also heavy-tailed Cauchy noise for outlier protection [53].

ii) Privacy noises are allowed to be time-varying for a better privacy-preserving level. By Proposition B.2 in Appendix B, Assumption 5 permits polynomially increasing privacy noises under a maximum allowable growth rate.

iii) A co-design guideline for the privacy noises and step sizes is presented to ensure the convergence of Algorithm 1 under quantized communications. Crucially, this guideline differs fundamentally from the nonquantized case [20], [54], [55]. In nonquantized settings, larger noise increases communication signal variance, requiring smaller step sizes to mitigate its impact on estimation accuracy. Under quantized communications, however, the communication signal variance remains uniformly bounded regardless of noise magnitude. Instead, larger noise reduces the previous estimate information carried by the communication signal, necessitating larger step sizes to ensure efficient information utilization.

IV. Privacy Analysis

The section will analyze the privacy-preserving capability of Algorithm 1. Theorem 1 proves that privacy-preserving capability of Algorithm 1 is dynamically enhanced over time. Theorem 2 quantifies the improvement of the privacy-preserving capability brought by the binary-valued quantizers.

Theorem 1: Suppose Assumptions 2, 3, 4 i), ii), iv), and 5 ii), iii) hold, and

i) $\beta_{i,k}\lambda_{\max}(Q_i) < 1$, where $Q_i = \bar{H}_i^\top \bar{H}_i$ and $\lambda_{\max}(Q_i)$ is the maximum eigenvalue of Q_i;

ii) $\sum_{t=1}^{\infty} \prod_{l=1}^{t} \eta_{ij,t}\,(1-\lambda_{\min}^{+}(Q_i)\,\beta_{i,l})^2 < \infty$, where $\lambda_{\min}^{+}(Q_i)$ is the minimum positive eigenvalue of Q_i. Then

$\mathbb{E}\mathcal{I}_{\mathtt{s}}(\mathtt{y}_{i,k})$

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

$$\leq \sum_{j\in\mathcal{N}_i}\sum_{t=k+1}^{\infty}\beta_{i,k}^2 q_{ij,t}\eta_{ij,t}\left(\prod_{l=k+1}^{t-1}\left(1-\lambda_{\min}^{+}(Q_i)\beta_{i,l}\right)\right)^2 \bar{H}_i\bar{H}_i^{\top} < \infty \tag{6}$$

where $q_{ij,k}$ is given in Section II-A. Furthermore, if

iii) $p_{u,1}=\mathbb{P}\{\mathtt{G}_1=\mathcal{G}^{(u)}\}=\pi_u$;

iv) $\eta_{ij,k}\leq\frac{\eta_{ij,1}}{k^{2\epsilon_{ij}}}$ with $\eta_{ij,1}>0$ and $\epsilon_{ij}\geq 0$;

v) $\beta_{i,k}=\frac{\beta_{i,1}}{k^{\delta_i}}$ if $k\geq k_{i,0}$, and 0, otherwise, where $\delta_i\in(1/2,1]$, $\beta_{i,1}\in(0,k_{i,0}^{\delta_i})$ and $2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}>1$;

then

$$\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k}) \leq \sum_{j\in\mathcal{N}_i}\sum_{u\in\mathbb{G}_{ij}}\pi_u R_{ij,k}\beta_{i,k}\eta_{ij,k}\bar{H}_i\bar{H}_i^{\top} = O\left(\sum_{j\in\mathcal{N}_i}\frac{1}{k^{\delta_i+2\epsilon_{ij}}}\right) \tag{7}$$

where

$$R_{ij,k}=\begin{cases}\frac{\beta_{i,1}}{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}-1}\frac{(k+1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}k^{2\epsilon_{ij}}}{(k-1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}}}, & \text{if } \delta_i=1\\ \frac{\beta_{i,1}}{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}-(\delta_i-2\epsilon_{ij})k^{\delta_i-1}}, & \text{if } \delta_i\in(1/2,1).\end{cases}$$

Therefore, Algorithm 1 achieves the dynamically enhanced privacy.

Proof: First, we expand the sequence $\mathtt{S}=\{\mathtt{s}_{ij,k}:(i,j)\in\mathtt{E}_k,k\in\mathbb{N}\}$ to $\breve{\mathtt{S}}=\{\mathtt{s}_{ij,k}:(i,j)\in\mathcal{E},k\in\mathbb{N}\}$. Note that we have expanded the noise sequence $\{\mathtt{d}_{ij,k}:(i,j)\in\mathtt{E}_k,k\in\mathbb{N}\}$ to $\{\mathtt{d}_{ij,k}:(i,j)\in\mathcal{E},k\in\mathbb{N}\}$ in Assumption 4. Then, for all $(i,j)\in\mathcal{E}$, define

$$\mathtt{a}'_{ij,k}=\begin{cases}1, & \text{if } (i,j)\in\mathtt{E}_k\\ 0, & \text{otherwise}\end{cases}$$

$$\mathtt{s}'_{ij,k}=\begin{cases}1, & \text{if } \mathtt{x}_{i,k}+\mathtt{d}_{ij,k}\leq C_{ij}\\ -1, & \text{otherwise.}\end{cases} \tag{8}$$

For $(i,j)\in\mathcal{E}\setminus\mathtt{E}_k$, define $\mathtt{s}_{ij,k}=0$. Then, $\mathtt{s}_{ij,k}=\mathtt{a}'_{ij,k}\mathtt{s}'_{ij,k}$ and $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})=\mathbb{E}\mathcal{I}_{\breve{\mathtt{S}}}(\mathtt{y}_{i,k})$.

Note that $\mathcal{I}_{\{\mathtt{y}_{i,l}:l\neq k\}}(\mathtt{y}_{i,k})=0$. Then, by Corollary A.1 in Appendix A, we have

$$\mathcal{I}_{\breve{\mathtt{S}}}(\mathtt{y}_{i,k})\leq\mathcal{I}_{\breve{\mathtt{S}}}(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,l}:l\neq k\}). \tag{9}$$

Note that for any $(u,v)\in\mathcal{E}$, $\mathtt{d}_{uv,t}$ is independent of $\mathtt{M}^{-}_{i,t-1,k}$ and $\mathtt{y}_{i,k}$, and $\mathtt{x}_{u,t}$ is $\sigma(\mathtt{M}^{-}_{i,t-1,k}\cup\{\mathtt{y}_{i,k}\})$-measurable, where $\sigma(\cdot)$ is the minimum σ-algebra containing the corresponding set, and $\mathtt{M}^{-}_{i,t,k}=\{\mathtt{y}_{i,l}:l\neq k\}\cup\{\mathtt{s}_{uv,l}:(u,v)\in\mathcal{E},l\leq t\}$. Then, given $\mathtt{M}^{-}_{i,t-1,k}$ and $\mathtt{y}_{i,k}$, one can see that $\{\mathtt{s}'_{uv,t}:(u,v)\in\mathcal{E}\}$ is independent. Besides, given $\mathtt{M}^{-}_{i,t-1,k}$ and $\mathtt{y}_{i,k}$, we have that $\mathtt{s}'_{uv,t}$ is uniquely determined by $\mathtt{d}_{uv,t}$, and $\mathtt{a}'_{uv,t}$ is uniquely determined by $\mathtt{G}_k$. Then, by Assumption 4, given $\mathtt{M}^{-}_{i,t-1,k}$ and $\mathtt{y}_{i,k}$, one can see that $\{\mathtt{s}'_{uv,t}:(u,v)\in\mathcal{E}\}$ is independent of $\{a'_{uv,t}:(u,v)\in\mathcal{E}\}$. Therefore, by Lemma A.3 in Appendix A, one can get

$$\mathcal{I}_{\breve{\mathtt{S}}}(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,l}:l\neq k\})=\sum_{t=1}^{\infty}\sum_{(u,v)\in\mathcal{E}}\mathcal{I}_{\mathtt{s}_{uv,t+1}}(\mathtt{y}_{i,k}|\mathtt{M}^{-}_{i,t,k}) = \sum_{t=1}^{\infty}\sum_{j\in\mathcal{N}_i}\mathcal{I}_{\mathtt{s}_{ij,t+1}}(\mathtt{y}_{i,k}|\mathtt{M}^{-}_{i,t,k}). \tag{10}$$

Denote $\bar{\mathtt{q}}_{ij,t}=\mathbb{P}\{(i,j)\in\mathcal{E}|\mathtt{G}_{t-1}\}$, and note that $\{\mathtt{d}_{ij,k}:k\in\mathbb{N}\}$ is independent. Then, we have

$$\begin{aligned}&\ln\mathbb{P}\left\{\mathtt{s}_{ij,t}\middle|\mathtt{y}_{i,k},\mathtt{M}^{-}_{i,t,k-1}\right\}\\ &=\ln\left(\bar{\mathtt{q}}_{ij,t}F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})\right)\mathbb{I}_{\{\mathtt{s}_{ij,t}=1\}}+\ln(1-\bar{\mathtt{q}}_{ij,t})\mathbb{I}_{\{\mathtt{s}_{ij,t}=0\}}\\ &\quad+\ln\left(\bar{\mathtt{q}}_{ij,t}\left(1-F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})\right)\right)\mathbb{I}_{\{\mathtt{s}_{ij,t}=-1\}}\end{aligned}$$

which implies

$$\begin{aligned}&\frac{\partial}{\partial\mathtt{y}_{i,k}}\ln\left(\mathbb{P}\left\{\mathtt{s}_{ij,t}\middle|\mathtt{y}_{i,k},\mathtt{M}^{-}_{i,t-1,k}\right\}\right)\\ &=\frac{\partial}{\partial\mathtt{y}_{i,k}}\ln\left(\bar{\mathtt{q}}_{ij,t}F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})\right)\mathbb{I}_{\{\mathtt{s}_{ij,t}=1\}}\\ &\quad+\frac{\partial}{\partial\mathtt{y}_{i,k}}\ln\left(\bar{\mathtt{q}}_{ij,t}\left(1-F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})\right)\right)\mathbb{I}_{\{\mathtt{s}_{ij,t}=-1\}}\\ &=-\frac{f_{ij,t}\left(C_{ij}-\mathtt{x}_{i,t}\right)}{F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})}\frac{\partial\mathtt{x}_{i,t}}{\partial\mathtt{y}_{i,k}}\mathbb{I}_{\{\mathtt{s}_{ij,t}=1\}}\\ &\quad+\frac{f_{ij,t}\left(C_{ij}-\mathtt{x}_{i,t}\right)}{1-F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})}\frac{\partial\mathtt{x}_{i,t}}{\partial\mathtt{y}_{i,k}}\mathbb{I}_{\{\mathtt{s}_{ij,t}=-1\}}.\end{aligned} \tag{11}$$

Now, we calculate $\frac{\partial\mathtt{x}_{i,t}}{\partial\mathtt{y}_{i,k}}$. If $k\geq t$, then $\frac{\partial\mathtt{x}_{i,t}}{\partial\mathtt{y}_{i,k}}=0$. If $k<t$, then by Lemma A.4 in Appendix A, one can get

$$\begin{aligned}\frac{\partial\mathtt{x}_{i,t}}{\partial\mathtt{y}_{i,k}}&=\beta_{i,k}\bar{H}_iJ_i\left(\prod_{l=k+1}^{t-1}\left(I_n-\beta_{i,l}Q_i\right)\right)^{\top}\varphi_t\\ &=\beta_{i,k}\bar{H}_i\left(\prod_{l=k+1}^{t-1}\left(J_i-\beta_{i,l}Q_i\right)\right)^{\top}\varphi_t\end{aligned} \tag{12}$$

where $J_i=Q_i^{+}Q_i$. Hence, by (9)–(12) and Lemmas A.5 and A.6 in Appendix A, one can get

$$\begin{aligned}&\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})=\mathbb{E}\mathcal{I}_{\breve{\mathtt{S}}}(\mathtt{y}_{i,k})\\ &=\sum_{t=1}^{\infty}\sum_{j\in\mathcal{N}_i}\mathbb{E}\left[\left(\frac{\partial}{\partial\mathtt{y}_{i,k}}\ln\left(\mathbb{P}\left\{\mathtt{s}_{ij,t}\middle|\mathtt{y}_{i,k},\mathtt{M}^{-}_{i,t-1,k}\right\}\right)\right)\right.\\ &\quad\left.\cdot\left(\frac{\partial}{\partial\mathtt{y}_{i,k}}\ln\left(\mathbb{P}\left\{\mathtt{s}_{ij,t}\middle|\mathtt{y}_{i,k},\mathtt{M}^{-}_{i,t-1,k}\right\}\right)\right)^{\top}\right]\\ &=\sum_{j\in\mathcal{N}_i}\sum_{t=k+1}^{\infty}\beta_{i,k}^2\mathbb{E}\left[\frac{\bar{\mathtt{q}}_{ij,t}f_{ij,t}^2\left(C_{ij}-\mathtt{x}_{i,t}\right)}{F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})\left(1-F_{ij,t}(C_{ij}-\mathtt{x}_{i,t})\right)}\right]\\ &\quad\cdot\bar{H}_i\left(\prod_{l=k+1}^{t-1}\left(J_i-\beta_{i,l}Q_i\right)\right)^{\top}\varphi_t\varphi_t^{\top}\left(\prod_{l=k+1}^{t-1}\left(J_i-\beta_{i,l}Q_i\right)\right)\bar{H}_i^{\top}\\ &\leq\sum_{j\in\mathcal{N}_i}\sum_{t=k+1}^{\infty}\beta_{i,k}^2q_{ij,t}\eta_{ij,t}\left(\prod_{l=k+1}^{t-1}\left(1-\lambda_{\min}^{+}(Q_i)\beta_{i,l}\right)\right)^2\bar{H}_i\bar{H}_i^{\top}\\ &<\infty.\end{aligned} \tag{13}$$

Now, we prove (7). If $k<k_{i,0}$, then $\beta_{i,k}=0$, which together with (6) implies $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})=0$.

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

If $k \geq k_{i,0}$, then by [55, Lemma A.2], one can get

$$\begin{aligned}
&\bar{H}_i\left(\prod_{l=k+1}^{t-1}(J_i-\beta_{i,l}Q_i)\right)^{\top}\varphi_t\varphi_t^{\top}\left(\prod_{l=k+1}^{t-1}(J_i-\beta_{i,l}Q_i)\right)\bar{H}_i^{\top}\\
&\leq\left(\prod_{l=k+1}^{t-1}\left(1-\frac{\lambda_{\min}^{+}(Q_i)\beta_{i,1}}{l^{\delta_i}}\right)\right)^2\bar{H}_i\bar{H}_i^{\top}\\
&\leq\begin{cases}\left(\frac{k+1}{t-1}\right)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}\bar{H}_i\bar{H}_i^{\top}\\ \qquad\qquad\text{if } \delta_i=1\\ \exp\left(\frac{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}{1-\delta_i}\left((k+1)^{1-\delta_i}-t^{1-\delta_i}\right)\right)\bar{H}_i\bar{H}_i^{\top}\\ \qquad\qquad\text{if } \delta_i<1.\end{cases}
\end{aligned}\tag{14}$$

Therefore, if $\delta_i = 1$, then

$$\begin{aligned}
\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})&\leq\sum_{j\in\mathcal{N}_i}\sum_{t=k+1}^{\infty}\beta_{i,k}^2q_{ij,t}\eta_{ij,t}\left(\frac{k+1}{t-1}\right)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}\bar{H}_i\bar{H}_i^{\top}\\
&\leq\sum_{j\in\mathcal{N}_i}\beta_{i,1}^2\eta_{ij,1}\left(\sum_{u\in\mathbb{G}_{ij}}\pi_u\right)\frac{(k+1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}}{k^2}\\
&\quad\cdot\sum_{t=k+1}^{\infty}\frac{\bar{H}_i\bar{H}_i^{\top}}{(t-1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}}}\\
&\leq\sum_{j\in\mathcal{N}_i}\beta_{i,1}^2\eta_{ij,1}\left(\sum_{u\in\mathbb{G}_{ij}}\pi_u\right)\frac{(k+1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}}{k^2}\\
&\quad\cdot\frac{(k-1)^{1-2\lambda_{\min}^{+}(Q_i)\beta_{i,1}-2\epsilon_{ij}}}{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}-1}\bar{H}_i\bar{H}_i^{\top}\\
&\leq\sum_{j\in\mathcal{N}_i}\left(\sum_{u\in\mathbb{G}_{ij}}\pi_u\right)\frac{\beta_{i,1}}{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}-1}\\
&\quad\cdot\frac{(k+1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}k^{2\epsilon_{ij}}}{(k-1)^{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}+2\epsilon_{ij}}}\beta_{i,k}\eta_{ij,k}\bar{H}_i\bar{H}_i^{\top}.
\end{aligned}\tag{15}$$

If $\delta_i < 1$, then $2\lambda_{\min}^{+}(Q_i)\beta_{i,1}k^{1-\delta_i} > 1-2\epsilon_{ij} > \delta_i - 2\epsilon_{ij}$, which together with Lemma A.7 in Appendix A implies

$$\begin{aligned}
&\mathbb{E}\mathcal{I}_{\{\mathtt{s}_{ij,t}:j\in\mathcal{N}_i,t\in\mathbb{N}\}}(\mathtt{y}_{i,k})\\
&\leq\sum_{j\in\mathcal{N}_i}\sum_{t=k+1}^{\infty}\beta_{i,k}^2\eta_{ij,t}q_{ij,t}\frac{\exp\left(\frac{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}{1-\delta_i}(k+1)^{1-\delta_i}\right)}{\exp\left(\frac{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}{1-\delta_i}t^{1-\delta_i}\right)}\bar{H}_i\bar{H}_i^{\top}\\
&\leq\sum_{j\in\mathcal{N}_i}\left(\sum_{u\in\mathbb{G}_{ij}}\pi_u\right)\frac{\beta_{i,1}^2\eta_{ij,1}}{k^{2\delta_i}}\exp\left(\frac{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}{1-\delta_i}(k+1)^{1-\delta_i}\right)\\
&\quad\cdot\sum_{t=k+1}^{\infty}\frac{\exp\left(-\frac{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}}{1-\delta_i}t^{1-\delta_i}\right)}{t^{2\epsilon_{ij}}}\bar{H}_i\bar{H}_i^{\top}\\
&\leq\sum_{j\in\mathcal{N}_i}\left(\sum_{u\in\mathbb{G}_{ij}}\pi_u\right)\frac{\beta_{i,1}}{2\lambda_{\min}^{+}(Q_i)\beta_{i,1}-(\delta_i-2\epsilon_{ij})k^{\delta_i-1}}\\
&\quad\cdot\beta_{i,k}\eta_{ij,k}\bar{H}_i\bar{H}_i^{\top}.
\end{aligned}\tag{16}$$

Hence, by $\beta_{i,k}\eta_{ij,k} = O\left(\frac{1}{k^{\delta_i+2\epsilon_i}}\right)$, (7) is obtained. Then, the theorem can be proof by Lemma A.1. □

Remark 15: In Theorem 1, i) is easy to be achieved since $\beta_{i,k}$ converges to 0. ii) provides a general condition on minimum privacy noises required to ensure finite $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$, with iv) and v) specifying polynomial step-size and noise design schemes satisfying this condition. iii) implies that the initial distribution of the Markovian chain is its stationary distribution. This condition is naturally satisfied if the chain has been running from the infinite past.

Remark 16: By (7), $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k}) = O\left(\sum_{j\in\mathcal{N}_i}\beta_{i,k}\eta_{ij,k}\right)$, which indicates that the privacy can be enhanced over time. The sensor i's operator can control the convergence rate of $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$ by properly selecting the step size $\beta_{i,k}$ and the privacy noise distributions. First, smaller $\beta_{i,k}$ implies a better privacy. Besides, privacy noises can be designed as increasing ones to enhance privacy. In contrast, the constant dithered lattice quantizer methods [31], [33], [34] and the decaying noise method [32] are difficult to utilize for improving the convergence rate of $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$. In addition, (6) reveals that the privacy-preserving level measured by Fisher information is proportional to communication frequency. Therefore, the Markovian switching topology can improve privacy by reducing communication frequency.

Remark 17: By Propositions B.1 and B.2, the privacy noise distributions satisfying the condition of Theorem 1 include $\mathcal{N}(0,\sigma_{ij,k}^2)$ with $\sigma_{ij,k} = \sigma_{ij,1}k^{\epsilon_{ij}}$ and $Lap(0,b_{ij,k})$ with $b_{ij,k} = b_{ij,1}k^{\epsilon_{ij}}$. Under such a choice of noise distribution, by Theorem 1, we have $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k}) = O\left(\sum_{j\in\mathcal{N}_i}\frac{\beta_{i,k}}{\mathbb{E}\mathtt{d}_{ij,k}^2}\right)$. Besides, the variances of the privacy noises are not necessarily finite. For example, by Propositions B.1 and B.2, the privacy noises can obey Cauchy distribution, which is heavy tailed with infinite variance.

The following theorem takes Gaussian privacy noise as an example to quantify the improvement of privacy brought by the binary-valued quantizers. In the theorem, the conditional Fisher information given $\{\mathtt{y}_{i,t}: t\neq k\}$ is considered as the privacy metric to eliminate privacy-preserving effects between different observations $\mathtt{y}_{i,k}$.

Theorem 2: Under the condition of Theorem 1, when the noise $d_{ij,k}$ is Gaussian distributed, we have

$$\mathcal{I}_{\mathtt{S}}\left(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,t}: t\neq k\}\right)\leq\frac{2}{\pi}\mathcal{I}_{\bar{\mathtt{x}}}\left(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,t}\ : t\neq k\}\right)$$

where $\bar{\mathtt{x}}_{ij,k} = \mathtt{x}_{i,k} + \mathtt{d}_{ij,k}$ and $\bar{\mathtt{X}} = \{\bar{\mathtt{x}}_{ij,k}: (i,j)\in\mathtt{E}_k, k\in\mathbb{N}\}$.

Proof: Set $\mathtt{d}_{ij,k}\sim\mathcal{N}(\mu_{ij,k},\sigma_{ij,k}^2)$. Similar to (13) and by [36, Lemma 5.3], we have

$$\begin{aligned}
&\mathcal{I}_{\mathtt{S}}\left(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,t}: t\neq k\}\right)=\mathcal{I}_{\bar{\mathtt{S}}}\left(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,t}: t\neq k\}\right)\\
&\leq\sum_{t=k+1}^{\infty}\sum_{j\in\mathcal{N}_i}\frac{2}{\pi\sigma_{ij,k}^2}q_{ij,t}\bar{\varphi}_{i,k,t}\bar{\varphi}_{i,k,t}^{\top}
\end{aligned}$$

where $\bar{\varphi}_{i,k,t} = \beta_{i,k}\bar{H}_i\left(\prod_{l=k+1}^{t-1}(I_n-\beta_{i,l}Q_i)\right)^{\top}\varphi_t$.

Similarly, one can get

$$\mathcal{I}_{\bar{\mathtt{X}}}\left(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,t}: t\neq k\}\right)=\mathcal{I}_{\bar{\mathtt{X}}}\left(\mathtt{y}_{i,k}|\{\mathtt{y}_{i,t}: t\neq k\}\right)$$

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

$$= \sum_{t=k+1}^{\infty} \sum_{j\in\mathcal{N}_i} \frac{1}{\sigma_{ij,k}^2} q_{ij,t} \bar{\varphi}_{i,k,t} \bar{\varphi}_{i,k,t}^\top$$

where $\breve{\mathtt{X}} = \{\bar{\mathtt{x}}_{ij,k}\mathtt{a}'_{ij,k} : (i,j) \in \mathtt{E}_k, k \in \mathbb{N}\}$ and $\mathtt{a}'_{ij,k}$ is defined in (8). Thus, the theorem is proved. □

Remark 18: Theorem 2 proves that in the Gaussian privacy noise case, the introduction of quantizers improves the privacy-preserving capability of the algorithm by at least $\frac{\pi}{2}$ times. Similarly, according to Lemma B.3, in the Cauchy noise case, the improvement is at least $\frac{\pi^2}{8}$ times. Therefore, the impact of quantizers in the privacy-preserving level is revealed as a multiplicative effect. And, by Lemma B.2, in the Laplacian noise case, the introduction of the quantizers also improves the privacy-preserving capability of the algorithm, except for the case that $\mathtt{x}_{i,k} = C$ for all k, which will not happen almost surely due to the randomness of $\mathtt{x}_{i,k}$.

V. Convergence Analysis

This section will focus on the convergence properties of Algorithm 1. First, the almost sure convergence will be proved. Then, the almost sure convergence rate will be obtained.

For convenience, denote

$$\tilde{\theta}_{i,k} = \hat{\theta}_{i,k} - \theta,\ \tilde{\Theta}_k = \mathrm{col}\{\tilde{\theta}_{1,k}, \ldots, \tilde{\theta}_{N,k}\},\ \bar{a}_{ij} = \sum_{r=1}^{M} \pi_r a_{ij}^{(r)}$$

$$\bar{\mathbb{H}} = \mathrm{diag}\{\bar{H}_1^\top \bar{H}_1, \ldots, \bar{H}_N^\top \bar{H}_N\},\ \hat{\mathtt{F}}_{ij,k} = F_{ij,k}(C_{ij} - \mathtt{x}_{i,k})$$

$$\bar{\mathbb{H}}_{\beta,k} = \mathrm{diag}\{\beta_{1,k}\bar{H}_1^\top \bar{H}_1, \ldots, \beta_{N,k}\bar{H}_N^\top \bar{H}_N\}$$

$$\Phi_{i,k} = \varphi_k \sum_{j\in\mathcal{N}_i} \alpha_{ij,k}(\mathtt{a}_{ij,k} - \bar{a}_{ij})(\mathtt{s}_{ij,k} - \mathtt{s}_{ji,k})$$

$$\Phi'_{i,k} = \varphi_k \sum_{j\in\mathcal{N}_i} \alpha_{ij,k}\bar{a}_{ij}\left((\mathtt{s}_{ij,k} - \mathtt{s}_{ji,k}) - 2(\hat{\mathtt{F}}_{ij,k} - \hat{\mathtt{F}}_{ji,k})\right)$$

$$\mathtt{W}_k = \mathrm{col}\{\beta_{1,k}\left(\mathtt{y}_{1,k} - \bar{H}_1\theta\right), \ldots, \beta_{N,k}\left(\mathtt{y}_{N,k} - \bar{H}_N\theta\right)\} + \mathrm{col}\{\Phi_{1,k}, \ldots, \Phi_{N,k}\} + \mathrm{col}\{\Phi'_{1,k}, \ldots, \Phi'_{N,k}\}$$

$$\mathcal{F}_k = \sigma(\{\mathtt{w}_{i,t}, \mathtt{G}_t, \mathtt{H}_{i,t}, \mathtt{d}_{ij,t} : i \in \mathcal{V}, (i,j) \in \mathtt{E}_t, 1 \le t \le k\}).$$

Then, $\tilde{\Theta}_k$ is $\mathcal{F}_k$-measurable.

The following theorem proves the almost sure convergence of Algorithm 1.

Theorem 3: Suppose that Assumptions 1, 2, 3, 4 i), iii), iv) and 5 hold. Then, the estimate $\hat{\theta}_{i,k}$ in Algorithm 1 converges to the true value θ almost surely.

Proof: By [56, Thm. 1.2], there exists $\lambda_a \in (0,1)$ such that $\mathbb{E}\mathtt{a}_{ij,k} = \bar{a}_{ij} + O(\lambda_a^k)$. Then, by Assumptions 3 and 4 iv), we have $\mathbb{E}[\mathtt{a}_{ij,k}\mathtt{s}_{ij,k} \,|\, \mathcal{F}_{k-1}] = \bar{a}_{ij}F(C_{ij} - \mathtt{x}_{i,k}) + O(\lambda_a^k)$. Therefore, one can get

$$\mathbb{E}\left[\|\tilde{\theta}_{i,k}\|^2 \middle| \mathcal{F}_{k-1}\right]$$

$$= \|\tilde{\theta}_{i,k-1}\|^2 - 2\beta_{i,k}\left(\bar{H}_i\tilde{\theta}_{i,k}\right)^2 + 2\varphi_k^\top\tilde{\theta}_{i,k-1}\sum_{j\in\mathcal{N}_i}\alpha_{ij,k}\bar{a}_{ij}\left(\hat{\mathtt{F}}_{ij,k} - \hat{\mathtt{F}}_{ji,k}\right)$$

$$+ O\left(\beta_{i,k}^2\left(\|\tilde{\theta}_{i,k-1}\|^2 + 1\right) + \sum_{j\in\mathcal{N}_i}\alpha_{ij,k}^2 + \lambda_a^k\right).$$

Define $\tilde{\mathtt{x}}_{i,k} = \varphi_k^\top\tilde{\theta}_{i,k-1}$. By $\mathtt{x}_{i,k} = \varphi_k^\top\hat{\theta}_{i,k-1} = \tilde{\mathtt{x}}_{i,k} + \varphi_k^\top\theta$, we have $\mathtt{x}_{i,k} - \mathtt{x}_{j,k} = \tilde{\mathtt{x}}_{i,k} - \tilde{\mathtt{x}}_{j,k}$. Then, we have

$$\sum_{i\in\mathcal{V}}\varphi_k^\top\tilde{\theta}_{i,k-1}\sum_{j\in\mathcal{N}_i}\alpha_{ij,k}\bar{a}_{ij}\left(\hat{\mathtt{F}}_{ij,k} - \hat{\mathtt{F}}_{ji,k}\right) = 2\sum_{(i,j)\in\mathcal{E}}\alpha_{ij,k}\bar{a}_{ij}\left(\mathtt{x}_{i,k} - \mathtt{x}_{j,k}\right)\left(\hat{\mathtt{F}}_{ij,k} - \hat{\mathtt{F}}_{ji,k}\right) \le 0$$

which implies

$$\mathbb{E}\left[\sum_{i\in\mathcal{V}}\|\tilde{\theta}_{i,k}\|^2 \middle| \mathcal{F}_{k-1}\right] \le \sum_{i\in\mathcal{V}}\|\tilde{\theta}_{i,k-1}\|^2 + O\left(\sum_{i\in\mathcal{V}}\beta_{i,k}^2\left(\|\tilde{\theta}_{i,k-1}\|^2 + 1\right) + \sum_{(i,j)\in\mathcal{E}}\alpha_{ij,k}^2 + \lambda_a^k\right).$$

Hence, by [57, Thm. 1], $\sum_{i\in\mathcal{V}}\|\tilde{\theta}_{i,k}\|^2$ converges to a finite value almost surely. Therefore, $\tilde{\theta}_{i,k}$, $\hat{\theta}_{i,k}$, and $\mathtt{x}_{i,k}$ are all bounded almost surely.

By the Lagrange mean value theorem [58], there exists $\xi_{ij,k}$ between $C_{ij} - \mathtt{x}_{i,k}$ and $C_{ij} - \mathtt{x}_{j,k}$ such that

$$\hat{\mathtt{F}}_{ij,k} - \hat{\mathtt{F}}_{ji,k} = f_{ij,k}(\xi_{ij,k})\left(\mathtt{x}_{j,k} - \mathtt{x}_{i,k}\right) = f_{ij,k}(\xi_{ij,k})\left(\tilde{\mathtt{x}}_{j,k} - \tilde{\mathtt{x}}_{i,k}\right).$$

For convenience, set $\check{\mathtt{f}}_{ij,k} = f_{ij,k}(\xi_{ij,k})$. By the almost sure boundedness of $\mathtt{x}_{i,k}$ and Assumption 4, there exists $\underline{\mathtt{f}} > 0$ such that $\check{\mathtt{f}}_{ij,k} \ge \underline{\mathtt{f}}\zeta_{ij,k}$ almost surely.

Define $\mathtt{L}_{F,k}$ as a Laplacian matrix whose element in the ith row and jth column is $-\alpha_{ij,k}\bar{a}_{ij}\check{\mathtt{f}}_{ij,k}$ if $i \ne j$, and $\sum_{l\in\mathcal{N}_i}\alpha_{1j,k}\bar{a}_{il}\check{\mathtt{f}}_{il,k}$ if $i = j$. Then, we have

$$\tilde{\Theta}_k = \left(I_{N\times n} - \mathbb{H}_{\beta,k} - \mathtt{L}_{F,k}\otimes\varphi_k\varphi_k^\top\right)\tilde{\Theta}_{k-1} + \mathtt{W}_k \tag{17}$$

and $\mathtt{L}_{F,k} \ge z_k\underline{\mathtt{f}}\bar{\mathcal{L}}$, where z_k is given in Assumption 5 and $\bar{\mathcal{L}} = \sum_{r=1}^{M}\pi_r\mathcal{L}^{(r)}$. In addition, by [59, Lemma 5.4], one can get

$$\sum_{t=k-n+1}^{k}\frac{1}{z_t}\left(\bar{\mathbb{H}}_{\beta,t} + \mathtt{L}_{F,t}\otimes\varphi_t\varphi_t^\top\right) \ge \sum_{t=k-n+1}^{k}\left(\bar{\mathbb{H}} + \underline{\mathtt{f}}\bar{\mathcal{L}}\otimes\varphi_t\varphi_t^\top\right) \ge n\bar{\mathbb{H}} + \underline{\mathtt{f}}\bar{\mathcal{L}}\otimes I_n > 0. \tag{18}$$

Hence, by Corollary A.2 in Appendix A, $\tilde{\Theta}_k$ and then $\tilde{\theta}_{i,k}$ converge to 0 almost surely. □

Remark 19: Note that in Algorithm 1, each sensor transmits 1 bit of information to its neighbors at each time step, and as analyzed in Proposition B.2, the privacy noises are allowed to be increasing. Then, by Theorem 3, the estimates of Algorithm 1 can converge to the true value θ even under 1 communication data rate and increasing privacy noises, which is the first to be achieved among existing privacy-preserving distributed algorithms [9], [11], [54].

Remark 20: In Assumption 4, the privacy noise can be heavy tailed. Therefore, the results in Theorem 3 can also be applied to the heavy-tailed communication noise case [42], [43]. For Algorithm 1, the key to achieving convergence with heavy-tailed

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

noises lies in the binary-valued quantizer, which transmits noisy signals with probably infinite variances to binary-valued signals with uniformly bounded variances.

Then, the following theorem calculates the almost sure convergence rate of Algorithm 1.

Theorem 4: Suppose that Assumptions 1–4 hold, $\rho > 4$, and the distribution of privacy noise $\mathtt{d}_{ij,k}$ is $\mathcal{N}(0,\sigma_{ij,k}^2)$ (resp., $Lap(0,b_{ij,k})$, $Cauchy(0,r_{ij,k})$) with $\sigma_{ij,k}=\sigma_{ij,1}k^{\epsilon_{ij}}$ (resp., $b_{ij,k}=b_{ij,1}k^{\epsilon_{ij}}$, $r_{ij,k}=r_{ij,1}k^{\epsilon_{ij}}$) and $\sigma_{ij,1}=\sigma_{ji,1}>0$ (resp., $b_{ij,1}=b_{ji,1}>0$, $r_{ij,1}=r_{ji,1}>0$). Given $k_{i,0}$, set $\alpha_{ij,k}=\frac{\alpha_{ij,1}}{k^{\gamma_{ij}}}$, $\beta_{i,k}=\frac{\beta_{i,1}}{k^{\delta_i}}$ if $k\geq k_{i,0}$, and 0, otherwise, where

i) $\alpha_{ij,1}=\alpha_{ji,1}>0$, $\gamma_{ij}=\gamma_{ji}>\frac{1}{2}$ and $\epsilon_{ij}=\epsilon_{ji}\geq 0$ for all $(i,j)\in\mathcal{E}$, and $\beta_{i,1}>0$ for all $i\in\mathcal{V}$;

ii) $\max_{(i,j)\in\mathcal{E}}\gamma_{ij}+\epsilon_{ij}<\min_{i\in\mathcal{V}}\delta_i\leq\max_{i\in\mathcal{V}}\delta_i\leq 1$.

Then, the almost sure convergence rate of the estimation error for the sensor i is

$$\tilde{\theta}_{i,k}=\begin{cases}O\left(1\Big/k^{\frac{\lambda_H\min_{i\in\mathcal{V}}\beta_{i,1}}{N}}\right)\\ \qquad\text{if } \bar{b}=1,\ 2\underline{b}-\frac{2\lambda_H\min_{i\in\mathcal{V}}\beta_{i,1}}{N}>1\\ O\left(\ln k/k^{\underline{b}-1/2}\right)\\ \qquad\text{if } \bar{b}=1,\ 2\underline{b}-\frac{2\lambda_H\min_{i\in\mathcal{V}}\beta_{i,1}}{N}\leq 1\\ O\left(1\Big/k^{\underline{b}-\bar{b}/2}\right)\\ \qquad\text{if } \bar{b}<1\end{cases}\quad\text{a.s.}$$

where $\lambda_H=\lambda_{\min}\left(\sum_{i=1}^N\bar{H}_i^\top\bar{H}_i\right)$, $\underline{b}=\min_{(i,j)\in\mathcal{E}}\gamma_{ij}$, and $\bar{b}=\max_{i\in\mathcal{V}}\delta_i$.

Proof: By Lemma B.1, $\zeta_{ij,k}$ in Assumption 4 can be $\frac{1}{k^{\epsilon_{ij}}}$. In this case, z_k in Assumption 5 is $\min\{\frac{\alpha_{ij,1}}{k^{\gamma_{ij}+\epsilon_{ij}}}:(i,j)\in\mathcal{E}\}\cup\{\frac{\beta_{i,1}}{k^{\delta_i}}:i\in\mathcal{V}\}$.

If $\bar{b}<1$, then the theorem can be proved by (17), (18), and Corollary A.3 in Appendix A.

If $\bar{b}=1$, then $k\sum_{i=1}^N\beta_{i,k}\bar{H}_i^\top\bar{H}_i\geq\lambda_H\min_{i\in\mathcal{V}}\beta_{i,1}$. Hence, by [59, Lemma 5.4], one can get

$$\frac{1}{n}\sum_{t=k-n+1}^{k}t\left(\bar{\mathbb{H}}_{\beta,t}+\mathtt{L}_{F,t}\otimes\varphi_t\varphi_t^\top\right)\geq\frac{\lambda_H\min_{i\in\mathcal{V}}\beta_{i,1}}{N}I_{nN}+O\left(\frac{1}{k^\tau}\right)$$

for some $\tau>0$, which together with (17) and Corollary A.3 implies the theorem. □

Remark 21: For all $\upsilon\in\left(0,\frac{1}{2}\right)$, when $\delta_i=1$, $\gamma_{ij}>\upsilon+\frac{1}{2}$, and $\beta_{i,1}$ is sufficiently large, by Theorem 4, Algorithm 1 can achieve an almost sure convergence rate of $o(1/k^\upsilon)$. The convergence rate is consistent with the classical one [50] of distributed estimation without considering the quantized communications and privacy issues.

Remark 22: Theorem 4 i) and ii) provide a design guideline for polynomial step sizes and noise levels that satisfy Assumption 5. In particular, a larger noise restricts the admissible choice of $\alpha_{ij,k}$, which in turn slows down the convergence rate of the algorithm, and vice versa. In addition, by Theorems and 1 and 4, when $\delta_i=1$, the best privacy level and convergence rate will be achieved simultaneously.

VI. Tradeoff Between Privacy and Convergence Rate

Based on the privacy and convergence analysis in Theorems 1–4, this section will establish the tradeoff between the privacy level and the convergence rate of Algorithm 1.

Theorem 5: Suppose that Assumptions 1–5 hold. Then, given $\nu\in\left(\frac{1}{2},1\right)$, there exist step-size sequences $\{\alpha_{ij,k}:(i,j)\in\mathtt{E}_k,k\in\mathbb{N}\}$, $\{\beta_{i,k}:i\in\mathcal{V},k\in\mathbb{N}\}$, and the privacy noise distribution sequence $\{F_{ij,k}(\cdot):(i,j)\in\mathtt{E}_k,k\in\mathbb{N}\}$ such that $\mathbb{E}\mathcal{I}_{\mathtt{s}}(\mathtt{y}_{i,k})=O\left(\frac{1}{k^\chi}\right)$ and $\tilde{\theta}_{i,k}=O\left(\frac{1}{k^{\nu-\chi/2}}\right)$ almost surely for all $i\in\mathcal{V}$ and $\chi\in[1,2\nu)$.

Proof: Consider the privacy noises obeying the Gaussian distribution $\mathcal{N}(0,\sigma_{ij,k}^2)$ with $\sigma_{ij,k}=\sigma_{ij,1}k^{\epsilon_{ij}}$, $\sigma_{ij,1}=\sigma_{ji,1}>0$, and $\epsilon_{ij}=\epsilon_{ji}\geq 0$ as Theorem 4 and Proposition B.1.

Set $k_{i,0}=\exp\left(\left\lfloor\frac{1}{\delta_i}\ln\beta_{i,1}\right\rfloor+1\right)$, $\delta_i=1$, $\epsilon_{ij}=\frac{\chi-1}{2}$, and $\gamma_{ij}=\frac{2+\nu-\chi}{2}$, and let $\beta_{i,1}$ be any number bigger than $\frac{2-\chi}{2\lambda_{\min}^+(Q_i)}$, where $\lfloor\cdot\rfloor$ is the floor function. The step size $\alpha_{ij,k}=\frac{\alpha_{ij,1}}{k^{\gamma_{ij}}}$ and $\beta_{i,k}=\frac{\beta_{i,1}}{k^{\delta_i}}$ if $k\geq k_{i,0}$, and 0, otherwise. Then, the step-size conditions in Theorems 1 and 4 are achieved simultaneously. By Theorem 1, $\mathbb{E}\mathcal{I}_{\mathtt{s}}(\mathtt{y}_{i,k})=O\left(\frac{1}{k^{\delta_i+2\epsilon_{ij}}}\right)=O\left(\frac{1}{k^\chi}\right)$. By Theorem 4, $\tilde{\theta}_{i,k}=O(\ln k/k^{(1+\nu-\chi)/2})=O\left(\frac{1}{k^{\nu-\chi/2}}\right)$ almost surely. The theorem is proved. □

Remark 23: In Theorem 5, χ decides the convergence rate of $\mathbb{E}\mathcal{I}_{\mathtt{s}}(\mathtt{y}_{i,k})$, and hence, it can be regarded as the coefficient representing the desired privacy level. ν can be any scalar smaller than but very close to 1. By Theorem 5, better privacy implies a slower convergence rate, and vice versa. The proof of Theorem 5 provides a practical selection for privacy noises and step sizes to achieve the tradeoff under given χ, where the privacy noises are set to be increasing at a rate of $k^{\frac{\chi-1}{2}}$, and the step sizes $\alpha_{ij,k}$ and $\beta_{i,k}$ are set to be convergent at rates of $\frac{1}{k^{(2+\nu-\chi)/2}}$ and $\frac{1}{k}$, respectively.

VII. Simulations

This section will demonstrate the main results of the article by simulation examples.

A. Numerical Examples

Consider an eight-sensor system. The communication graph sequence $\{\mathtt{G}_k:k\in\mathbb{N}\}$ is switching among $\mathcal{G}^{(1)}$, $\mathcal{G}^{(2)}$, $\mathcal{G}^{(3)}$, and $\mathcal{G}^{(4)}$, as shown in Fig. 1. For all $u=1,2,3,4$, $a_{ij}^{(u)}=1$ if $(i,j)\in\mathcal{E}^{(u)}$, and 0, otherwise. The communication graph sequence $\{\mathtt{G}_k:k\in\mathbb{N}\}$ is associated with a Markovian chain $\{\mathtt{m}_k:k\in\mathbb{N}\}$. The initial probability $p_{u,1}=\mathbb{P}\{\mathtt{G}_1=\mathcal{G}^{(u)}\}=\frac{1}{4}$. The transition probability matrix

$$P=(p_{uv})_{4\times4}=\begin{bmatrix}\frac{1}{2}&\frac{1}{2}&0&0\\0&\frac{1}{2}&\frac{1}{2}&0\\0&0&\frac{1}{2}&\frac{1}{2}\\\frac{1}{2}&0&0&\frac{1}{2}\end{bmatrix}$$

where $p_{uv}=\mathbb{P}\{m_k=v|m_{k-1}=u\}$. Therefore, the stationary distribution $\pi_u=\frac{1}{4}$ for all $u=1,2,3,4$.

In the observation model, the unknown parameter $\theta=[1\ \ -1]^\top$. Sensors fail with probability $\frac{1}{2}$. When the sensor i

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

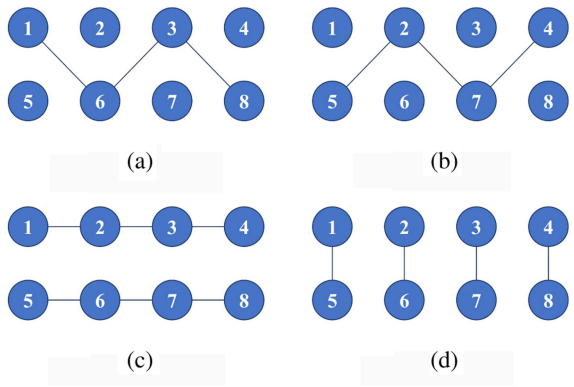


Fig. 1. Communication graphs. (a) Graph $\mathcal{G}^{(1)}$. (b) Graph $\mathcal{G}^{(2)}$. (c) Graph $\mathcal{G}^{(3)}$. (d) Graph $\mathcal{G}^{(4)}$.

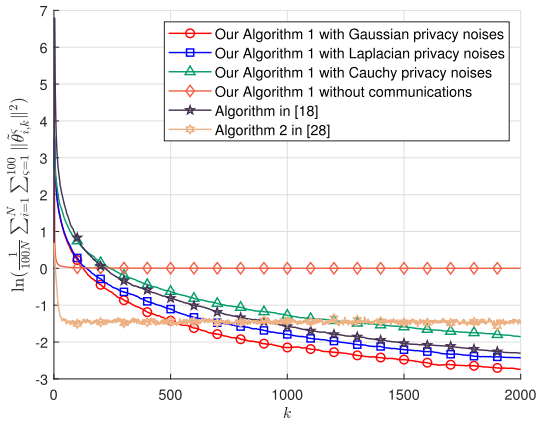


Fig. 2. Trajectories of $\ln\left(\frac{1}{100N}\sum_{i=1}^{N}\sum_{\varsigma=1}^{100}\|\tilde{\theta}_{i,k}^{\varsigma}\|^2\right)$.

does not fail at time k, the measurement matrix $\mathtt{H}_{i,k}=[2\ \ 0]$ if i is odd, and $[0\ \ 2]$ if i is even. When the sensor i fails, $\mathtt{H}_{i,k}=0$. Therefore, $\bar{H}_i=[1\ \ 0]$ if i is odd, and $[0\ \ 1]$ if i is even. The observation noise $\mathtt{w}_{i,k}$ is i.i.d. Gaussian with zero mean and standard deviation 0.1.

In Algorithm 1, the threshold $C_{ij}=0$. The step sizes $\alpha_{ij,k}=\frac{3}{k^{0.8}}$, and $\beta_{i,k}=\frac{3}{k}$ if $k\geq 8$, and 0, otherwise. Three types of privacy noise distributions are considered, including Gaussian distribution $\mathcal{N}(0,\sigma_{ij,k}^2)$ with $\sigma_{ij,k}=k^{0.15}$, Laplacian distribution $Lap(0,b_{ij,k})$ with $b_{ij,k}=k^{0.15}$, and Cauchy distribution $Cauchy(0,r_{ij,k})$ with $r_{ij,k}=k^{0.15}$.

We repeat the simulation 100 times, and Fig. 2 illustrates the trajectories of $\frac{1}{100N}\sum_{i=1}^{N}\sum_{\varsigma=1}^{100}\|\tilde{\theta}_{i,k}^{\varsigma}\|^2$, where $\tilde{\theta}_{i,k}^{\varsigma}$ is the estimate of θ by sensor i at time k in the ςth run. The figure demonstrates that the estimates can converge to the true value θ even under increasing noises and one communication data rate. In addition, Fig. 2 shows that when sensors do not communicate with each other, the estimates do not converge to the true value. Therefore, the communication is necessary for the distributed estimation.

Fig. 3 draws the upper bounds of the nonzero elements in $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$ given by Theorem 1. To avoid duplicate presentation of similar figures, Fig. 3 only takes the sensors 1 and 2 as

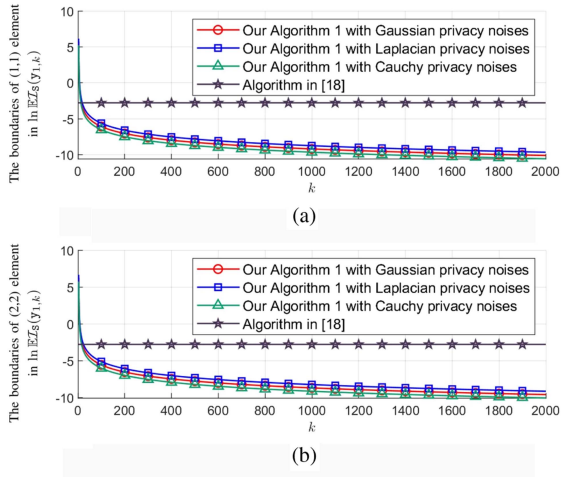


Fig. 3. Upper boundaries of the nonzero elements in $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$ for the sensors 1 and 2. (a) Boundaries of (1,1) element in $\ln\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{1,k})$. (b) Boundaries of (2,2) element in $\ln\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{2,k})$.

representative examples. The figure indicates that the privacy-preserving capability of Algorithm 1 is dynamically enhanced under the three types of privacy noise distributions.

Remark 24: Under our setting, $\bar{H}_i\bar{H}_i^{\top}=\begin{bmatrix}1&0\\0&0\end{bmatrix}$ if i is odd, and $\bar{H}_i\bar{H}_i^{\top}=\begin{bmatrix}0&0\\0&1\end{bmatrix}$ if i is even. Then, by Theorem 1, only one element in the matrix $\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$ is nonzero. Therefore, it is sufficient to depict the trajectory of nonzero element in the matrix $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$ in Fig. 3.

Figs. 2 and 3 also compare Algorithm 1 with existing ones in [18] and [28]. From Figs. 2 and 3, one can get that Algorithm 1 can achieve similar estimation error and much better privacy simultaneously compared with the algorithm in [18]. Besides, the algorithm in [18] requires sensors to transmit real-valued information to each other, in contrast to the binary-valued communications of our Algorithm 1. The work in [28, Algorithm 2] also requires binary-valued communications. The mean-square errors of its estimates quickly decrease to a certain value, but do not converge to 0. Therefore, after about 1000 iterations, the estimation error of our Algorithm 1 is smaller than that of [28, Algorithm 2]. Besides, Sayin and Kozat [28] do not consider the privacy-preserving issue.

Fig. 4 demonstrates the tradeoff between privacy and convergence rate for Algorithm 1. In Algorithm 1, the step size $\alpha_{ij,k}=\frac{3}{k^{(2.9-\chi)/2}}$, and the privacy noises are Cauchy distributed with $r_{ij,k}=k^{\frac{\chi-1}{2}}$, where $\chi=1.3,\ 1.6$, and 1.9. Fig. 4(a) depicts the log–log plot for the boundaries of $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{1,k})$. It is observed that a better privacy level is achieved with a larger χ. Fig. 4(b) shows the log–log plot for the trajectories of $\frac{1}{100N}\sum_{i=1}^{N}\sum_{\varsigma=1}^{100}\|\tilde{\theta}_{i,k}^{\varsigma}\|^2$. It is observed that a better convergence rate is achieved with a smaller χ. Therefore, the tradeoff can be shown under different χ.

By Remark 11, when not pursuing 1-bit communication data rate, φ_k in Algorithm 1 can be replaced by a higher dimensional

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

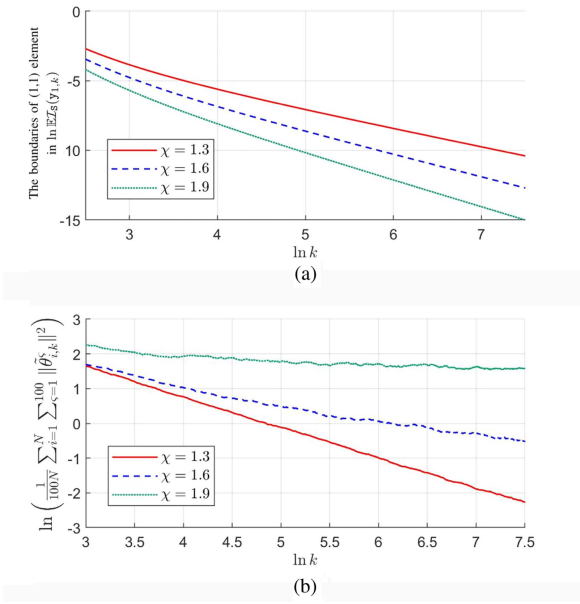


Fig. 4. Tradeoff between privacy and convergence rate. (a) Boundaries of (1,1) element in $\ln \mathbb{E}\mathcal{I}_S(y_{1,k})$ with different χ. (b) Log–log plot for $\frac{1}{100N}\sum_{i=1}^{N}\sum_{\varsigma=1}^{100}\|\tilde{\theta}_{i,k}^{\varsigma}\|^2$ with different χ.

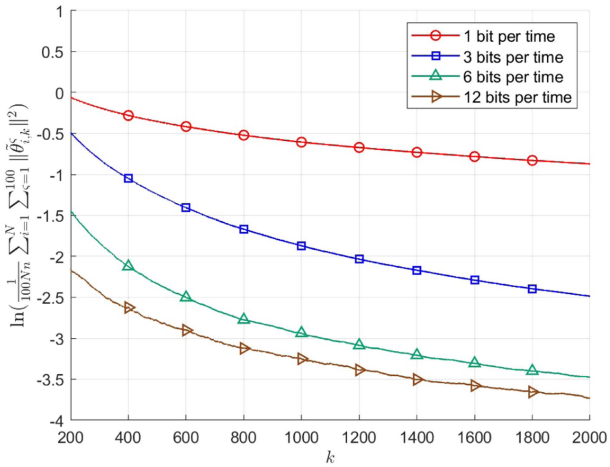


Fig. 5. Trajectories of $\ln\left(\frac{1}{100Nn}\sum_{i=1}^{N}\sum_{\varsigma=1}^{100}\|\tilde{\theta}_{i,k}^{\varsigma}\|^2\right)$ with different data rate.

selection matrix to make the modified algorithm perform better in high-dimensional settings. To show this improvement, consider the case of $n = 12$. The unknown parameter θ is uniformly generated within $[-1, 1]^{12}$. $\bar{H}_i$ is expanded to $[I_6 \;\; O_6]$ if i is odd, and $[O_6 \;\; I_6]$ if i is even, which ensures Assumption 2 in the high dimensional θ case. Under this settings, from Fig. 5, one can see that the modified algorithm converges faster than the original Algorithm 1. In addition, since the influence of φ_k was neglected in the analysis of Theorem 1, the upper bound of privacy-preserving level obtained from Theorem 1 still holds after removing φ_k, which implies that the modified algorithm still has strong privacy-preserving capability.

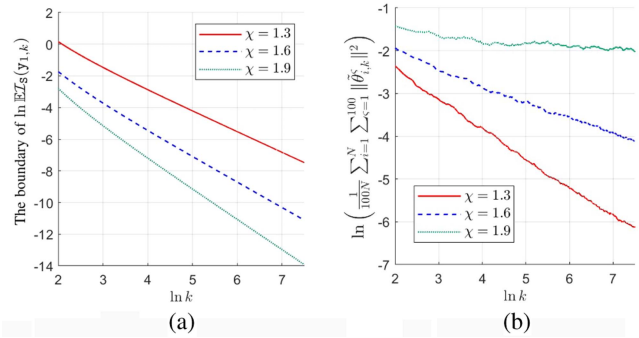


Fig. 6. Privacy and convergence of Algorithm 1 for the event rate analysis of essential hypertension. (a) Boundary of $\ln \mathbb{E}\mathcal{I}_S(y_{1,k})$. (b) Estimation error.

B. Experiment on the Event Rate Analysis of Essential Hypertension

In this section, Algorithm 1 is applied in the event rate analysis of essential hypertension in 281 299 white British participants.[2] In the experiment, $H_{i,k} = 1$ if there is a participant for the sensor i at time k, and $H_{i,k} = 0$, otherwise. $H_{i,k} = 1$ with probability 0.7. The observation $y_{i,k} = 1$ if $H_{i,k} = 1$, and the participant suffers from the essential hypertension, and $y_{i,k} = 0$, otherwise. Such clinical information $y_{i,k}$ is private and needs to be protected in practical scenarios.

About four-fifths of the database is used as the training set, while the rest is the test set. From the test set, we have the event rate $\theta \approx 0.2699$. Data in the training set are distributed in a 20-sensor network. In the network, $a_{ij,k} = 1$ if $(i, j) \in E_k$, and 0, otherwise. The initial probability $\mathbb{P}\{a_{ij,1} = 1\} = 0.5$, and the transition probability $\mathbb{P}\{a_{ij,k} = 1 | a_{ij,k-1} = 1\} = \mathbb{P}\{a_{ij,k} = 0 | a_{ij,k-1} = 0\} = 0.7$.

In Algorithm 1, the threshold $C_{ij} = 0$. The step sizes $\alpha_{ij,k} = \frac{0.2}{k^{(2.9-\chi)/2}}$, $\beta_{i,k} = \frac{0.4}{k}$, and the privacy noise is Gaussian $\mathcal{N}(0, \sigma_{ij,k}^2)$ with $\sigma_{ij,k} = k^{\frac{\chi-1}{2}}$, where $\chi = 1.3,\ 1.6$, and 1.9. Under the settings, Fig. 6(a) shows the dynamically enhanced privacy of our algorithm, and Fig. 6(b) demonstrates the convergence.

VIII. Conclusion

This article proposes a binary-valued quantizer-based privacy-preserving distributed estimation algorithm with multiple advantages. In terms of privacy, the proposed algorithm achieves the dynamically enhanced privacy, and the Fisher information-based privacy metric $\mathbb{E}\mathcal{I}_S(y_{i,k})$ is proved to converge to 0 at a polynomial rate. In terms of communication costs, each sensor transmits only 1 bit of information to its neighbors at each time step. Besides, the assumption on the negligible quantization error for real-valued messages is not required. In terms of effectiveness, the proposed algorithm can achieve almost sure convergence even with increasing privacy noises. A polynomial convergence rate is also obtained. Besides, the tradeoff between privacy and convergence rate is established.

[2]The data come from U.K. Biobank (Application: 78793).

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

When the step sizes and privacy noise distributions are properly selected, a better privacy-preserving capability implies a slower convergence rate, and vice versa.

There are still many interesting topics worth further investigation, for example, how to apply the proposed method to distributed optimization problems to achieve the dynamically enhanced privacy and a limited data rate, and how to protect the observation matrices.

APPENDIX A
LEMMAS AND COROLLARIES

Lemma A.1: If $\lim_{k\to\infty}\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})=0$, then the privacy-preserving capability is dynamically enhanced.

Proof: Since $\lim_{k\to\infty}\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})=0$, for any $A>0$, there exists $T\in\mathbb{N}$ such that $\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,t})\le A$ for all $t\ge T$. Then, the lemma can be proved by setting $A=\mathbb{E}\mathcal{I}_{\mathtt{S}}(\mathtt{y}_{i,k})$. □

Lemma A.2 (Chain rule for Fisher information, [45]): For random variables $\mathtt{X},\mathtt{Y}$, and θ, $\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta)=\mathcal{I}_{\mathtt{X}}(\theta)+\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X})\ge\mathcal{I}_{\mathtt{X}}(\theta)$.

Corollary A.1: For random variables $\mathtt{X},\mathtt{Y},\mathtt{Z}$, and θ, we have

a) $\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta|\mathtt{Z})=\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Z})+\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X},\mathtt{Z})$;

b) If $\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X})=0$, then $\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Y})\le\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta)=\mathcal{I}_{\mathtt{X}}(\theta)$;

c) If $\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Z})=0$, then $\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{Z})\le\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X},\mathtt{Z})$.

Proof:

a) By Lemma A.2, we have

$$\begin{aligned}\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta|\mathtt{Z})&=\mathcal{I}_{\mathtt{X},\mathtt{Y},\mathtt{Z}}(\theta)-\mathcal{I}_{\mathtt{Z}}(\theta)\\&=\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Y},\mathtt{Z})+\mathcal{I}_{\mathtt{X},\mathtt{Z}}(\theta|\mathtt{Y})-\mathcal{I}_{\mathtt{Z}}(\theta)=\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Z})+\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X},\mathtt{Z}).\end{aligned}$$

b) By Lemma A.2, we have

$$\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Y})\le\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta)=\mathcal{I}_{\mathtt{X}}(\theta)+\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X})=\mathcal{I}_{\mathtt{X}}(\theta).$$

c) By a), we have

$$\begin{aligned}\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{Z})&=\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta|\mathtt{Z})-\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Y},\mathtt{Z})\le\mathcal{I}_{\mathtt{X},\mathtt{Y}}(\theta|\mathtt{Z})\\&=\mathcal{I}_{\mathtt{X}}(\theta|\mathtt{Z})+\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X},\mathtt{Z})=\mathcal{I}_{\mathtt{Y}}(\theta|\mathtt{X},\mathtt{Z}).\end{aligned}$$

□

Lemma A.3: For random variables $\mathtt{X},\theta$, and random variable sequences $\mathtt{Y}_k=\{\mathtt{Y}_{i,k}:i=1,\dots,N\},\mathtt{Z}_k=\{\mathtt{Z}_{i,k}:i=1,\dots,N\}$ for all $k\in\mathbb{N}$, if

i) $\mathtt{Y}_{1,k},\dots,\mathtt{Y}_{N,k}\ne0,\mathtt{Z}_{1,k},\dots,\mathtt{Z}_{N,k}\in\{0,1\}$;

ii) Given θ, $\mathtt{X}$, and $\breve{\mathtt{Z}}_{k-1}$, the sequence $\mathtt{Y}_k$ is independent and independent of $\mathtt{Z}_k$, where $\breve{\mathtt{Z}}_k=\bigcup_{t=1}^k\hat{\mathtt{Z}}_t$ and $\hat{\mathtt{Z}}_k=\{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k},i=1,\dots,N\}$;

iii) $\mathcal{I}_{\mathtt{Z}_k}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})=0$,

then $\mathcal{I}_{\breve{\mathtt{Z}}_\infty}(\theta|\mathtt{X})=\sum_{k=1}^\infty\sum_{i=1}^N\mathcal{I}_{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})$.

Proof: Note that by i), we have that $\mathtt{Z}_{i,k}$ can be uniquely determined by $\mathtt{Z}_{i,k}\mathtt{Y}_{i,k}$. Then, by ii), given θ, $\mathtt{X}$, $\breve{\mathtt{Z}}_{k-1}$, and $\mathtt{Z}_k$, we have that $\hat{\mathtt{Z}}_k$ is independent. Hence, by Corollary A.1, we have

$$\begin{aligned}\mathcal{I}_{\breve{\mathtt{Z}}_\infty}(\theta|\mathtt{X})&=\sum_{k=1}^\infty\mathcal{I}_{\hat{\mathtt{Z}}_k}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})=\sum_{k=1}^\infty\mathcal{I}_{\hat{\mathtt{Z}}_k,\mathtt{Z}_k}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})\\&=\sum_{k=1}^\infty\mathcal{I}_{\hat{\mathtt{Z}}_k}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1},\mathtt{Z}_k)+\mathcal{I}_{\mathtt{Z}_k}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})\\&=\sum_{k=1}^\infty\sum_{i=1}^N\mathcal{I}_{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1},\mathtt{Z}_k).\end{aligned}\tag{A.1}$$

By iii), given θ, $\mathtt{X}$, $\breve{\mathtt{Z}}_{k-1}$, and $\mathtt{Z}_{i,k}$, we have that $\mathtt{Y}_{i,k}$ is independent of $\mathtt{Z}_{j,k}$ for all $j\ne i$. Therefore, by Corollary A.1, we have

$$\begin{aligned}&\mathcal{I}_{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1},\mathtt{Z}_k)=\mathcal{I}_{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1},\mathtt{Z}_{i,k})\\&=\mathcal{I}_{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k},\mathtt{Z}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})-\mathcal{I}_{\mathtt{Z}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})\\&=\mathcal{I}_{\mathtt{Z}_{i,k}\mathtt{Y}_{i,k}}(\theta|\mathtt{X},\breve{\mathtt{Z}}_{k-1})\end{aligned}$$

which together with (A.1) implies the lemma. □

Lemma A.4: For a matrix H, set $Q=H^\top H$ and $J=Q^+Q$. Then, $HJ=H$.

Proof: By [60, Thm. 1], one can get

$$\begin{aligned}HJ&=(H^\top)^+H^\top HQ^+Q=(H^\top)^+QQ^+Q\\&=(H^\top)^+Q=(H^\top)^+H^\top H=H.\end{aligned}$$

□

Lemma A.5: For a positive semidefinite matrix Q, set $J=Q^+Q$. Then, $\lambda_{\max}(J-\beta Q)=1-\beta\lambda^+_{\min}(Q)$, where $\beta\in[0,\frac{1}{\lambda^+_{\min}(Q)}]$, and $\lambda_{\max}(\cdot)$ and $\lambda^+_{\min}(\cdot)$ are defined in Theorem 1.

Proof: By [61, Thm. 5], all the eigenvectors v for Q are eigenvectors for $J-\beta Q$. If $Qv=0$, then $(J-\beta Q)v=0$. If $Qv=\lambda v$ for some $\lambda>0$, then $(J-\beta Q)v=(1-\beta\lambda)v$. The lemma is thereby proved. □

Lemma A.6: If sequences $\{a_k:k\in\mathbb{N}\}$, $\{b_k:k\in\mathbb{N}\}$, and $\{\eta_k:k\in\mathbb{N}\}$ satisfy

i) $a_k\in[0,\bar{a}]$ for some $\bar{a}<1$, and $\eta_k>0$;

ii) $\sum_{t=1}^\infty\prod_{l=1}^t\eta_t(1-a_l)^p<\infty$ for some positive integer p;

iii) $b_k>0$ and $\sum_{k=1}^\infty b_k<\infty$,

then $\sum_{t=k}^\infty\prod_{l=k}^t\eta_t(1-a_l+b_l)^p<\infty$.

Proof: First, we have

$$\begin{aligned}\sum_{t=k}^\infty\prod_{l=k}^t\eta_t(1-a_l)^p&=\frac{\sum_{t=k}^\infty\prod_{l=1}^t\eta_t(1-a_l)^p}{\prod_{l=1}^{k-1}(1-a_l)^p}\\&\le\frac{\sum_{t=1}^\infty\prod_{l=1}^t\eta_t(1-a_l)^p}{\prod_{l=1}^{k-1}(1-a_l)^p}<\infty.\end{aligned}$$

Then, one can get

$$\begin{aligned}&\sum_{t=k}^\infty\prod_{l=k}^t\eta_t(1-a_l+b_l)^p\\&\le\sum_{t=k}^\infty\prod_{l=k}^t\eta_t(1-a_l)^p\left(1+\frac{b_l}{1-\bar{a}}\right)^p\\&\le\left(\sum_{t=k}^\infty\prod_{l=k}^t\eta_t(1-a_l)^p\right)\left(\prod_{t=1}^\infty\left(1+\frac{b_t}{1-\bar{a}}\right)\right)^p<\infty.\end{aligned}$$

□

Lemma A.7: If $c,k_0>0$, $g\ge0$, and $p\in(0,1]$ satisfy $cpk_0^p\ge1-p-g$, then

$$\begin{aligned}&\sum_{t=1}^k\frac{\exp\left(-c(t+k_0)^p\right)}{(t+k_0)^g}\\&\le\frac{k_0^{1-p-g}\exp(-ck_0^p)-(k+k_0)^{1-p-g}\exp(-c(k+k_0)^p)}{cp-(1-p-g)k_0^{-p}}.\end{aligned}$$

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

Proof: From the condition of the lemma, we have

$$\frac{\sum_{t=1}^{k}\exp\left(-c(t+k_0)^p\right)}{(t+k_0)^g} \le \int_{k_0}^{k+k_0}\frac{\exp\left(-ct^p\right)}{t^g}dt$$

$$\le \int_{k_0}^{k+k_0}\frac{cp-(1-p-g)t^{-p}}{cp-(1-p-g)k_0^{-p}}\frac{\exp\left(-ct^p\right)}{t^g}dt$$

$$= \frac{\int_{k_0}^{k+k_0} cpt^{-g}\exp\left(-ct^p\right)-(1-p-g)t^{-p-g}\exp\left(-ct^p\right)dt}{cp-(1-p-g)k_0^{-p}}$$

$$= \frac{k_0^{1-p-g}\exp(-ck_0^p)-(k+k_0)^{1-p-g}\exp(-c(k+k_0)^p)}{cp-(1-p-g)k_0^{-p}}.$$

The lemma is thereby proved. □

Lemma A.8: Assume that

i) $\{\alpha_k : k\in\mathbb{N}\}$, $\{\beta_k : k\in\mathbb{N}\}$ and $\{\gamma_k : k\in\mathbb{N}\}$ are positive sequences satisfying $\sum_{k=1}^{\infty}\alpha_k=\infty$, $\sum_{k=1}^{\infty}\beta_k^2<\infty$ and $\sum_{k=1}^{\infty}\gamma_k^2<\infty$;

ii) $\{\mathcal{F}_k : k\in\mathbb{N}\}$ is a σ-algebra sequence with $\mathcal{F}_{k-1}\subseteq\mathcal{F}_k$ for all k;

iii) $\{\mathtt{W}_k,\mathcal{F}_k : k\in\mathbb{N}\}$ is a sequence of adaptive random variables satisfying $\sum_{k=1}^{\infty}\|\mathbb{E}[\mathtt{W}_k\,|\,\mathcal{F}_{k-1}]\|<\infty$ and $\mathbb{E}[\|\mathtt{W}_k-\mathbb{E}[\mathtt{W}_k\,|\,\mathcal{F}_{k-1}]\|^{\rho}\,|\,\mathcal{F}_{k-1}]=O(\beta_k^{\rho})$ almost surely for some $\rho>2$;

iv) $\{\mathtt{U}_k : k\in\mathbb{N}\}$ is a sequence with $\sum_{k=1}^{\infty}\alpha_k^2\|\mathtt{U}_k\|^2<\infty$. And, $\mathtt{U}_k$ is $\mathcal{F}_{k-1}$-measurable;

v) $\mathtt{U}_k+\mathtt{U}_k^{\top}\ge 2aI_n$ for some $p\in\mathbb{N}$, $a>0$ and all $k\in\mathbb{N}$ almost surely;

vi) $\{\mathtt{X}_k,\mathcal{F}_k : k\in\mathbb{N}\}$ is a sequence of adaptive random variables with

$$\mathtt{X}_k=(I_n-\alpha_k\mathtt{U}_k+O(\gamma_k))\,\mathtt{X}_{k-1}+\mathtt{W}_k,\ \text{a.s.} \tag{A.2}$$

Then, $\mathtt{X}_k$ converges to 0 almost surely.

Proof: Consider $\mathtt{X}'_k=\mathtt{X}_k-\mathtt{Y}_k$, where $Y_0=0$ and $\mathtt{Y}_k=(I_n-\alpha_k\mathtt{U}_k+O(\gamma_k))\mathtt{Y}_{k-1}+\mathbb{E}[\mathtt{W}_k\,|\,\mathcal{F}_{k-1}]$. Since $\|\mathtt{Y}_k\|\le(1-a\alpha_k+O(\alpha_k^2\|\mathtt{U}_k\|^2+\gamma_k))\|\mathtt{Y}_{k-1}\|+\|\mathbb{E}[\mathtt{W}_k\,|\,\mathcal{F}_{k-1}]\|$, by [54, Lemma 2], $\mathtt{Y}_k$ converges to 0 almost surely. Therefore, it suffices to prove the convergence of $\mathtt{X}'_k$, which satisfies

$$\mathtt{X}'_k=(I_n-\alpha_k\mathtt{U}_k+O(\gamma_k))\,\mathtt{X}'_{k-1}+\mathtt{W}_k-\mathbb{E}\left[\mathtt{W}_k|\mathcal{F}_{k-1}\right]. \tag{A.3}$$

By (A.3), one can get

$$\mathbb{E}\left[\|\mathtt{X}'_k\|^2\big|\mathcal{F}_{k-1}\right] \le \left(1-2a\alpha_k+\alpha_k^2\|\mathtt{U}_k\|^2+O(\gamma_k)\right)\|\mathtt{X}'_{k-1}\|^2+O\left(\beta_k^2\right). \tag{A.4}$$

Then by [54, Lemma 2], $\mathtt{X}_k$ converges to 0 almost surely. □

Corollary A.2: If i)–iv) and vi) in Lemma A.8 hold, $\alpha_k=O(\alpha_{k-1})$ and

$$\frac{1}{p}\sum_{t=k-p+1}^{k}\left(\mathtt{U}_t+\mathtt{U}_t^{\top}\right)\ge 2aI_n \tag{A.5}$$

for some $p\in\mathbb{N}$, $a>0$, and all $k\in\mathbb{N}$ almost surely, then $\mathtt{X}_k$ converges to 0 almost surely.

Proof: By (A.2), $\mathtt{X}_k=\prod_{t=k-p+1}^{k}(I_n-\alpha_t\mathtt{U}_t+O(\gamma_t))\mathtt{X}_{k-p}+\sum_{t=k-p+1}^{k}\prod_{l=t+1}^{k}(I_n-\alpha_l\mathtt{U}_l)\mathtt{W}_t$. In this recursive function, $\prod_{t=k-p+1}^{k}(I_n-\alpha_t\mathtt{U}_t+O(\gamma_t))=I_n-\sum_{t=k-p+1}^{k}\alpha_t\mathtt{U}_t+O(\sum_{t=k-p+1}^{k}(\gamma_t+\alpha_k^2\|\mathtt{U}_k\|^2))$. Note that by (A.5), one can get

$$\frac{1}{p}\sum_{t=k-p+1}^{k}\alpha_t\left(\mathtt{U}_t+\mathtt{U}_t^{\top}\right) \ge\left(\min_{k-p<t\le k}\alpha_t\right)\frac{1}{p}\sum_{t=k-p+1}^{k}\left(\mathtt{U}_t+\mathtt{U}_t^{\top}\right)\ge 2a\left(\min_{k-p<t\le k}\alpha_t\right)I_n$$

and by [39, Lemma A.2], $\sum_{k=p+1}^{\infty}\min_{k-p<t\le k}\alpha_t=\infty$. Then, the corollary can be proved by Lemma A.8. □

Lemma A.9: If an adaptive sequence $\{\mathtt{V}_k,\mathcal{F}_k : k\in\mathbb{N}\}$ satisfies $\mathbb{E}[\mathtt{V}_k\,|\,\mathcal{F}_{k-1}]\le(1-\frac{a}{k}+\gamma_k)\mathtt{V}_{k-1}+O\left(\frac{1}{k^b}\right)$ with $a>0$, $b>1$ and $\sum_{k=1}^{\infty}\gamma_k<\infty$, then

$$\mathtt{V}_k=\begin{cases}O\left(\frac{1}{k^a}\right), & \text{if } b-a>1\\ O\left(\frac{(\ln k)^2}{k^{b-1}}\right), & \text{if } b-a\le 1.\end{cases}$$

Proof: Due to the space limitation, we leave the proof to the arXiv version [62] of this article. □

Lemma A.10: If sequences $\{V_k : k\in\mathbb{N}\}$, $\{\xi_k : k\in\mathbb{N}\}$, $\{\eta_k : k\in\mathbb{N}\}$, and $\{\gamma_k : k\in\mathbb{N}\}$ satisfy

i) $\xi_k\ge 0$, $\overline{\lim}_{k\to\infty}\xi_k<1$;

ii) $\sum_{k=1}^{\infty}\eta_k<\infty$, $\sum_{k=1}^{\infty}|\gamma_k|<\infty$;

iii) $V_k\le(1-\xi_k+\gamma_k)V_{k-1}+\eta_k+O(\xi_k)$,

then V_k is uniformly upper bounded.

Proof: Due to the space limitation, we leave the proof to the arXiv version [62] of this article. □

Lemma A.11: If i)–vi) in Lemma A.8 hold, $\rho>4$, $\alpha_k=\frac{1}{k^c}$, $\beta_k=\frac{1}{k^b}$ for $c\in\left(\frac{1}{2},1\right]$ and $b>1$, and $\|\mathbb{E}[\mathtt{W}_k\,|\,\mathcal{F}_{k-1}]\|\le\lambda^k$ for some $\lambda\in(0,1)$, then

$$\mathtt{X}_k=\begin{cases}O\left(\frac{1}{k^a}\right), & \text{if } c=1, 2b-2a>1\\ O\left(\frac{\ln k}{k^{b-1/2}}\right), & \text{if } c=1, 2b-2a\le 1\ \text{a.s.}\\ O\left(\frac{1}{k^{b-c/2}}\right), & \text{if } c\in\left(\frac{1}{2},1\right).\end{cases} \tag{A.6}$$

Proof: Due to the space limitation, we leave the proof to the arXiv version [62] of this article. □

Corollary A.3: Suppose that i)–iv), vi) in Lemma A.8, and (A.5) in Corollary A.2 hold. ρ, α_k, and β_k are set as Lemma A.11. Then, $\mathtt{X}_k$ achieves the almost sure convergence rate as (A.6).

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

The proof of Corollary A.3 is similar to that of Corollary A.2 and thereby omitted here.

APPENDIX B
LEMMAS AND PROPOSITIONS ON GAUSSIAN, LAPLACIAN, AND CAUCHY DISTRIBUTIONS

Following lemmas provide some useful properties on Gaussian, Laplacian, and Cauchy distributions.

Lemma B.1: If the noise $\mathtt{d}_{ij,k}$ obeys the distribution $\mathcal{N}(0,\sigma_{ij,k}^2)$ with $\inf_k \sigma_{ij,k}>0$ (resp., $Lap(0,b_{ij,k})$ with $\inf_k b_{ij,k}>0$, then $\zeta_{ij,k}, Cauchy(0,r_{ij,k})$ with $\inf_k r_{ij,k}>0$), then $\zeta_{ij,k}$ in iii) of Assumption 4 can be $\frac{\sigma_{ij,1}}{\sigma_{ij,k}}$ (resp., $\frac{b_{ij,1}}{b_{ij,k}}$, $\frac{r_{ij,1}}{r_{ij,k}}$).

Proof: For the Gaussian distribution case, denote $f_G^\star(\cdot)$ as the density function of the standard Gaussian distribution. Then, $f_{ij,k}(x)=\frac{1}{\sigma_{ij,k}}f_G^\star\left(\frac{x}{\sigma_{ij,k}}\right)$. Since $\inf_k\sigma_{ij,k}>0$, there exists a compact set $\mathcal{X}'$ such that $\frac{x}{\sigma_{ij,k}}\in\mathcal{X}'$ for all $(i,j)\in\mathcal{E}$, $k\in\mathbb{N}$, and $x\in\mathcal{X}'$. Therefore, when $\zeta_{ij,k}=\frac{1}{\sigma_{ij,k}}$, $\inf_{(i,j)\in\mathcal{E},k\in\mathbb{N},x\in\mathcal{X}}\frac{f_{ij,k}(x)}{\zeta_{ij,k}}\geq\inf_{z\in\mathcal{X}'}\frac{f_G^\star(z)}{\sigma_{ij,1}}>0$, which implies the lemma. The proofs for Laplacian and Cauchy distribution cases are similar to the Gaussian one and, hence, omitted here. □

Lemma B.2: Given $b>0$, if $F_L(\cdot;b)$ and $f_L(\cdot;b)$ are the distribution function and the density function of the distribution $Lap(0,b)$, respectively, then

$$\sup_{x\in\mathbb{R}}\frac{f_L^2(x;b)}{F_L(x;b)(1-F_L(x;b))}=\frac{1}{b^2}.$$

Proof: By $f_L(x;b)=\frac{1}{2b}\exp(-\frac{|x|}{b})$, $F_L(x;b)=\frac{1}{2}\exp\left(\frac{x}{b}\right)$ if $x<0$, and $1-\frac{1}{2}\exp\left(-\frac{x}{b}\right)$, otherwise.

By symmetry, one can get

$$\sup_{x\in\mathbb{R}}\frac{f_L^2(x;b)}{F_L(x;b)(1-F_L(x;b))}=\sup_{x\geq0}\frac{f_L^2(x;b)}{F_L(x;b)(1-F_L(x;b))}$$
$$=\sup_{x\geq0}\frac{1}{2b^2\left(\exp\left(\frac{x}{b}\right)-\frac{1}{2}\right)}=\frac{1}{b^2}.$$

The lemma is thereby proved. □

Lemma B.3: Given $r>0$, if $F_C(\cdot;r)$ and $f_C(\cdot;r)$ are the distribution function and the density function of the distribution $Cauchy(0,r)$, respectively, then

$$\sup_{x\in\mathbb{R}}\frac{f_C^2(x;r)}{F_C(x;r)(1-F_C(x;r))}=\frac{4}{\pi^2r^2}.$$

Proof: Since $f_C(x;r)=\frac{1}{\pi r[1+(x/r)^2]}$, one can get $F_C(x;r)=\frac{1}{2}+\frac{1}{\pi}\arctan\left(\frac{x}{r}\right)$.

Note that $F_C(x;r)=F_C\left(\frac{x}{r};1\right)$ and $f_C(x;r)=\frac{1}{r}f_C\left(\frac{x}{r};1\right)$. Then, it suffices to consider the case of $r=1$. In this case, $F_C(x;1)$ and $f_C(\cdot;r)$ are abbreviated as $F_C(x)$ and $f_C(x)$, respectively. Denote

$$h_{C,1}(x)=\frac{F_C(x)(1-F_C(x))}{f_C^2(x)}=(1+x^2)^2\left(\frac{\pi^2}{4}-\arctan^2x\right).$$

Then, $h'_{C,1}(x)=(1+x^2)(\pi^2x-2\arctan x-4x\arctan^2x)$.

Furthermore, denote $h_{C,2}(x)=\pi^2x-2\arctan x-4x\arctan^2x$. Then, $h'_{C,1}(x)=(1+x^2)h_{C,2}(x)$, and

$$h'_{C,2}(x)=\pi^2-\frac{2}{1+x^2}-4\arctan^2x-\frac{8x\arctan x}{1+x^2}$$
$$h''_{C,2}(x)=\frac{-4x-16\arctan x}{(1+x^2)^2}.$$

Note that $h''_{C,2}(x)>0$ when $x<0$; $h''_{C,2}(x)<0$ when $x>0$; and $\lim_{x\to\infty}h'_{C,2}(x)=\lim_{x\to-\infty}h'_{C,2}(x)=0$. Then, $h'_{C,2}(x)>0$, which implies that $h_{C,2}(x)$ is strictly monotonously increasing. Furthermore, by $h_{C,2}(0)=0$, we have $h_{C,2}(x)<0$ when $x<0$, and $h_{C,2}(x)>0$ when $x>0$.

Note that $h'_{C,1}(x)=(1+x^2)h_{C,2}(x)$. Then, $h'_{C,1}(x)<0$ when $x<0$, and $h'_{C,1}(x)>0$ when $x>0$. Therefore, one has

$$\sup_{x\in\mathbb{R}}\frac{f_C^2(x)}{F_C(x)(1-F_C(x))}=\frac{1}{\inf_{x\in\mathbb{R}}h_{C,1}(x)}=\frac{1}{h_{C,1}(0)}=\frac{4}{\pi^2}.$$

The lemma is thereby proved. □

The following propositions give sufficient conditions on privacy noises satisfying Assumptions 4 and 5, when the privacy noises are Gaussian, Laplacian, and Cauchy.

Proposition B.1: For the noise distribution $\mathcal{N}(0,\sigma_{ij,k}^2)$ (resp., $Lap(0,b_{ij,k})$, $Cauchy(0,r_{ij,k})$), Assumption 4 ii) holds when $\sigma_{ij,k}>0$ (resp., $b_{ij,k}>0$, $r_{ij,k}>0$), and Assumption 4 iii) holds when $\inf_{k\in\mathbb{N}}\sigma_{ij,k}>0$ (resp., $\inf_{k\in\mathbb{N}}b_{ij,k}>0$, $\inf_{k\in\mathbb{N}}r_{ij,k}>0$). In addition, if $\epsilon_{ij}\geq0$ and $\sigma_{ij,k}=\sigma_{ij,1}k^{\epsilon_{ij}}$ (resp., $b_{ij,k}=b_{ij,1}k^{\epsilon_{ij}}$, $r_{ij,k}=r_{ij,1}k^{\epsilon_{ij}}$), then v) of Theorem 1 holds.

Proof: Consider Gaussian noise case. By [36, Lemma 5.3], one can get

$$\eta_{ij,k}=\sup_{x\in\mathbb{R}}\frac{f_{ij,k}^2(x)}{F_{ij,k}(x)\left(1-F_{ij,k}(x)\right)}$$
$$=\frac{f_{ij,k}^2(0)}{F_{ij,k}(0)\left(1-F_{ij,k}(0)\right)}=\frac{2}{\pi\sigma_{ij,k}^2}.\tag{B.1}$$

Therefore, when $\sigma_{ij,k}>0$, $\eta_{ij,k}<\infty$. Besides, Lemma B.1 implies that $\inf_{k\in\mathbb{N}}\sigma_{ij,k}>0$ is sufficient to achieve Assumption 4 ii). In addition, if $\epsilon_{ij}\geq0$ and $\sigma_{ij,k}=\sigma_{ij,1}k^{\epsilon_{ij}}$, then (B.1) implies v) of Theorem 1.

The analysis for the Laplacian and Cauchy noise cases is similar, and thereby omitted here. □

Remark B.1: By Proposition B.1, for Gaussian and Laplacian privacy noises, Assumption 4 ii) and iii) can be replaced with the condition that there is a uniform positive lower bound of the noise variances. The reasons to adopt the assumption are twofold. For the privacy, sufficient privacy noises can ensure the privacy-preserving capability of the algorithm. For the effectiveness, the privacy noises are also necessary dithered signals in the quantizers [39]. The lack of sufficient dithered signals in the quantizers will result in the algorithm failing to converge.

Proposition B.2: For the noise distribution $\mathcal{N}(0,\sigma_{ij,k}^2)$ (resp., $Lap(0,b_{ij,k})$, $Cauchy(0,r_{ij,k})$) with $\sigma_{ij,k}=\sigma_{ij,1}k^{\epsilon_{ij}}$ (resp., $b_{ij,k}=b_{ij,1}k^{\epsilon_{ij}}$, $r_{ij,k}=r_{ij,1}k^{\epsilon_{ij}}$) and $\zeta_{ij,k}=k^{-\epsilon_{ij}}$, there exists step-size sequences $\{\alpha_{ij,k}:(i,j)\in\mathcal{E},k\in\mathbb{N}\}$ and $\{\beta_{i,k}:i\in\mathcal{V},k\in\mathbb{N}\}$ satisfying Assumption 5 if and only if $\epsilon_{ij}\leq\frac{1}{2}$.

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

Proof: By Hölder inequality [58], $(\sum_{k=1}^{\infty}\alpha_{ij,k}\zeta_{ij,k})^2 \leq (\sum_{k=1}^{\infty}\alpha_{ij,k}^2)(\sum_{k=1}^{\infty}\zeta_{ij,k}^2)$. Then, under Assumption 5, $\sum_{k=1}^{\infty}\zeta_{ij,k}^2=\infty$, which implies $\epsilon_{ij}\leq\frac{1}{2}$. When $\epsilon_{ij}\leq\frac{1}{2}$, set $\alpha_{ij,k}=\frac{\alpha_{ij,1}}{k^{1-\epsilon_{ij}}\ln k}$ and $\beta_{i,k}=\frac{\beta_{i,1}}{k}$. Then, Assumption 5 holds. □

Remark B.2: $\zeta_{ij,k}$ in Proposition B.2 is consistent with the one given in Lemma B.1.

References

[1] S. Kar, J. M. F. Moura, and K. Ramanan, "Distributed parameter estimation in sensor networks: Nonlinear observation models and imperfect communication," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3575–3605, Jun. 2012.

[2] A. H. Sayed, S. Y. Tu, J. Chen, X. Zhao, and Z. J. Towfic, "Diffusion strategies for adaptation and learning over networks: An examination of distributed strategies and network behavior," *IEEE Signal Proc. Mag.*, vol. 30, no. 3, pp. 155–171, May 2013.

[3] C. Li, P. Zhou, L. Xiong, Q. Wang, and T. Wang, "Differentially private distributed online learning," *IEEE Trans. Knowl. Data Eng.*, vol. 30, no. 8, pp. 1440–1453, Aug. 2018.

[4] A. Smith, "Privacy-preserving statistical estimation with optimal convergence rates," in *Proc. 43rd Annu. ACM Symp. Theory Comput.*, San Jose, CA, USA, Jun. 2011, pp. 813–821.

[5] Y. Lu and M. H. Zhu, "Privacy preserving distributed optimization using homomorphic encryption," *Automatica*, vol. 96, pp. 314–325, 2018.

[6] C. N. Hadjicostis and A. D. Domínguez-García, "Privacy-preserving distributed averaging via homomorphically encrypted ratio consensus," *IEEE Trans. Autom. Control*, vol. 65, no. 9, pp. 3887–3894, Sep. 2020.

[7] C. Regueiro, I. Seco, S. de Diego, O. Lage, and L. Etxebarria, "Privacy-enhancing distributed protocol for data aggregation based on blockchain and homomorphic encryption," *Inf. Process. Manage.*, vol. 58, no. 6, 2021, Art. no. no. 102745.

[8] X. Li, N. Wang, L. Zhu, S. Yuan, and Z. Guan, "FUSE: A federated learning and U-shape split learning-based electricity theft detection framework," *Sci. China Inf. Sci.*, vol. 67, no. 4, 2024, Art. no. no. 149302.

[9] M. J. Ye, G. Q. Hu, L. H. Xie, and S. Y. Xu, "Differentially private distributed Nash equilibrium seeking for aggregative games," *IEEE Trans. Autom. Control*, vol. 67, no. 5, pp. 2451–2458, May 2022.

[10] F. Farokhi and H. Sandberg, "Ensuring privacy with constrained additive noise by minimizing fisher information," *Automatica*, vol. 99, pp. 275–288, 2019.

[11] C. Gratton, N. K. D. Venkategowda, R. Arablouei, and S. Werner, "Privacy-preserved distributed learning with zeroth-order optimization," *IEEE Trans. Inf. Forensics Secur.*, vol. 17, pp. 265–279, 2022.

[12] K. Wei et al., "Gradient sparsification for efficient wireless federated learning with differential privacy," *Sci. China Inf. Sci.*, vol. 67, no. 4, 2024, Art. no. 142303.

[13] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design," *Automatica*, vol. 81, no. 7, pp. 221–231, 2017.

[14] B. Jayaraman, L. X. Wang, D. Evans, and Q. Q. Gu, "Distributed learning without distress: Privacy-preserving empirical risk minimization," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst.*, Montreal, QC, Canada, Dec. 2018, pp. 6343–6354.

[15] Y. Q. Wang, "Privacy-preserving average consensus via state decomposition," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4711–4716, Nov. 2019.

[16] C. Altafini, "A system-theoretic framework for privacy preservation in continuous-time multiagent dynamics," *Automatica*, vol. 122, 2020, Art. no. 109253.

[17] Y. Liu, J. Liu, and T. Başar, "Differentially private gossip gradient descent," in *Proc. 57th IEEE Conf. Decis. Control*, Miami, FL, USA, Dec. 2018, pp. 2777–2782.

[18] J. M. Wang, J. F. Zhang, and X. K. Liu, "Differentially private resilient distributed cooperative online estimation over digraphs," *Int. J. Robust Nonlinear Control*, vol. 32, no. 15, pp. 8670–8688, 2022.

[19] J. M. Wang, J. W. Tan, and J. F. Zhang, "Differentially private distributed parameter estimation," *J. Syst. Sci. Complex.*, vol. 36, no. 1, pp. 187–204, 2023.

[20] Z. Q. Chen and Y. Q. Wang, "Locally differentially private distributed online learning with guaranteed optimality," *IEEE Trans. Autom. Control*, vol. 70, no. 4, pp. 2521–2536, Apr. 2025.

[21] N. Michelusi, G. Scutari, and C. S. Lee, "Finite-bit quantization for distributed algorithms with linear convergence," *IEEE Trans. Inf. Theory*, vol. 68, no. 11, pp. 7254–7280, Nov. 2022.

[22] D. Alistarh, D. Grubic, J. Z. Li, R. Tomioka, and M. Vojnovic, "QSGD: Communication-efficient SGD via gradient quantization and encoding," in *Proc. 31st Int. Conf. Neural Inf. Process. Syst.*, Long Beach, CA, USA, Dec. 2017, pp. 1709–1720.

[23] J. Wangni, J. Wang, J. Liu, and T. Zhang, "Gradient sparsification for communication-efficient distributed optimization," in *Proc. 32nd Int. Conf. Neural Inf. Process. Syst.*, Montreal, QC, Canada, Dec. 2018, pp. 1299–1309.

[24] A. Koloskova, S. U. Stich, and M. Jaggi, "Decentralized stochastic optimization and gossip algorithms with compressed communication," in *Proc. 36th Int. Conf. Mach. Learn.*, Long Beach, CA, USA, Jun. 2019, pp. 3478–3487.

[25] D. Kovalev, A. Koloskova, M. Jaggi, P. Richtarik, and S. U. Stich, "A linearly convergent algorithm for decentralized optimization: Sending less bits for free!," in *Proc. 24th Int. Conf. Artif. Intell. Statist.*, San Diego, CA, USA, Apr. 2021, pp. 4087–4095.

[26] M. Carpentiero, V. Matta, and A. H. Sayed, "Compressed regression over adaptive networks," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 10, pp. 851–867, 2024.

[27] M. Carpentiero, V. Matta, and A. H. Sayed, "Distributed adaptive learning under communication constraints," *IEEE Open J. Signal Process.*, vol. 5, pp. 321–358, 2024.

[28] M. O. Sayin and S. S. Kozat, "Compressive diffusion strategies over distributed networks for reduced communication load," *IEEE Trans. Signal Process.*, vol. 62, no. 20, pp. 5308–5323, Oct. 2014.

[29] R. Nassif, S. Vlaski, M. Carpentiero, V. Matta, M. Antonini, and A. H. Sayed, "Quantization for decentralized learning under subspace constraints," *IEEE Trans. Signal Process.*, vol. 71, pp. 2320–2335, 2023.

[30] W. Chen, L. Liu, and G. P. Liu, "Privacy-preserving distributed economic dispatch of microgrids: A dynamic quantization-based consensus scheme with homomorphic encryption," *IEEE Trans. Smart Grid*, vol. 14, no. 1, pp. 701–713, Jan. 2023.

[31] N. Lang, E. Sofer, T. Shaked, and N. Shlezinger, "Joint privacy enhancement and quantization in federated learning," *IEEE Trans. Signal Process.*, vol. 71, pp. 295–310, 2023.

[32] L. Gao, S. J. Deng, W. Ren, and C. Q. Hu, "Differentially private consensus with quantized communication," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 4075–4088, Aug. 2021.

[33] Y. Q. Wang and T. Başar, "Quantization enabled privacy protection in decentralized stochastic optimization," *IEEE Trans. Autom. Control*, vol. 68, no. 7, pp. 4038–4052, Jul. 2023.

[34] L. Liu, Y. Kawano, and M. Cao, "Design of stochastic quantizers for privacy-preserving control," *IEEE Trans. Autom. Control*, vol. 71, no. 2, pp. 1160–1174, Feb. 2026, doi: 10.1109/TAC.2025.3608135.

[35] Y. Wang, Y. L. Zhao, and J. F. Zhang, "Asymptotically efficient quasi-Newton type identification with quantized observations under bounded persistent excitations," *Automatica*, vol. 166, 2024, Art. no. 111722.

[36] Y. Wang, X. Li, Y. L. Zhao, and J. F. Zhang, "Threshold selection and resource allocation for quantized identification," *J. Syst. Sci. Complex.*, vol. 37, no. 1, pp. 204–229, 2024.

[37] F. Farokhi and H. Sandberg, "Fisher information as a measure of privacy: Preserving privacy of households with smart meters using batteries," *IEEE Trans. Smart Grid*, vol. 9, no. 5, pp. 4726–4734, Sep. 2018.

[38] J. M. Ke, J. M. Wang, and J. F. Zhang, "Differentiated output-based privacy-preserving average consensus," *IEEE Control Syst. Lett.*, vol. 7, pp. 1369–1347, 2023.

[39] J. M. Ke, X. D. Lu, Y. L. Zhao, and J. F. Zhang, "Signal-comparison-based distributed estimation under decaying average data rate communications," *SIAM J. Control Optim.*, vol. 63, no. 2, pp. 1129–1155, 2025.

[40] S. Kar and J. M. F. Moura, "Distributed consensus algorithms in sensor networks with imperfect communication: Link failures and channel noise," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 355–369, Jan. 2009.

[41] Q. Zhang and J. F. Zhang, "Distributed parameter estimation over unreliable networks with Markovian switching topologies," *IEEE Trans. Autom. Control*, vol. 57, no. 10, pp. 2545–2560, Oct. 2012.

[42] D. Jakovetic, M. Vukovic, D. Bajovic, A. K. Sahu, and S. Kar, "Distributed recursive estimation under heavy-tail communication noise," *SIAM J. Control Optim.*, vol. 61, no. 3, 1582–1609, 2023.

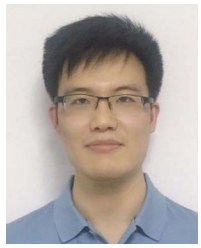
[43] M. Vukovic, D. Jakovetic, D. Bajovic, and S. Kar, "Nonlinear consensus innovations under correlated heavy-tailed noises: Mean square convergence rate and asymptotics," *SIAM J. Control Optim.*, vol. 62, no. 1, pp. 376–399, 2024.

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.

[44] A. N. Shiryaev, *Probability*, 2nd ed. New York, NY, USA: Springer, 1996.
[45] R. Zamir, "A proof of the Fisher information inequality via a data processing argument," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1246–1250, May 1998.
[46] L. P. Barnes, W. N. Chen, and A. Özgür, "Fisher information under local differential privacy," *IEEE J. Sel. Areas Inf. Theory*, vol. 1, no. 3, pp. 645–659, Nov. 2020.
[47] W. Wang, L. Ying, and J. Zhang, "On the relation between identifiability, differential privacy, and mutual-information privacy," *IEEE Trans. Inf. Theory*, vol. 62, no. 9, pp. 5018–5029, Sep. 2016.
[48] E. Nekouei, H. Sandberg, M. Skoglund, and K. H. Johansson, "A model randomization approach to statistical parameter privacy," *IEEE Trans. Autom. Control*, vol. 68, no. 2, pp. 839–850, Feb. 2023.
[49] I. Issa, A. B. Wagner, and S. Kamath, "An operational approach to information leakage," *IEEE Trans. Inf. Theory*, vol. 66, no. 3, pp. 1625–1657, Mar. 2020.
[50] S. Kar and J. M. F. Moura, "Convergence rate analysis of distributed gossip (linear parameter) estimation: Fundamental limits and tradeoffs," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 4, pp. 674–690, Aug. 2011.
[51] Y. L. Zhao, T. Wang, and W. J. Bi, "Consensus protocol for multiagent systems with undirected topologies and binary-valued communications," *IEEE Trans. Autom. Control*, vol, 64, no. 1, pp. 206–221, Jan. 2019.
[52] Y. Wang, Y. L. Zhao, and J. F. Zhang, "Distributed estimation with quantized measurements and communication over Markovian switching topologies," *Automatica*, vol. 183, 2026, Art. no. 112658.
[53] K. Ito, Y. Kawano, and K. Kashima, "Privacy protection with heavy-tailed noise for linear dynamical systems," *Automatica*, vol. 131, 2021, Art. no. 109732.
[54] Y. Q. Wang and A. Nedić, "Tailoring gradient methods for differentially private distributed optimization," *IEEE Trans. Autom. Control*, vol. 69, no. 2, pp. 872–887, Feb. 2024.
[55] J. M. Wang, J. M. Ke, and J. F. Zhang, "Differentially private bipartite consensus over signed networks with time-varying noises," *IEEE Trans. Autom. Control*, vol. 69, no. 9, pp. 5788–5803, Sep. 2024.
[56] E. Seneta, *Non-Negative Matrices and Markov Chains*, 2nd ed. New York, NY, USA, Springer: 2006.
[57] H. Robbins and D. Siegmund, "A convergence theorem for nonnegative almost supermartingales and some applications," in *Optimizing Methods in Statistics*. New York, NY, USA, Academic, 1971, pp. 233–257.
[58] V. A. Zorich, *Mathematical Analysis I*, 2nd ed. New York, NY, USA: Springer, 2015.
[59] S. Y. Xie and L. Guo, "Analysis of normalized least mean squares-based consensus adaptive filters under a general information condition," *SIAM J. Control Optim.*, vol. 56, pp. 3404–3431, 2018.
[60] T. N. E. Greville, "Note on the generalized inverse of a matrix product," *SIAM Rev.*, vol. 8, no. 4, pp. 518–521, 1966.
[61] J. E. Scroggs and P. L. Odell, "An alternate definition of a pseudoinverse of a matrix," *SIAM J. Appl. Math.*, vol. 14, no. 4, pp. 796–810, 1966.
[62] J. M. Ke, J. M. Wang, and J. F. Zhang, "Privacy-preserving distributed estimation with limited data rate," 2026, *arXiv:2510.12549*.

Jieming Ke received the B.S. degree in mathematics from the University of Chinese Academy of Science, Beijing, China, in 2020, and the Ph.D. degree in system theory from the Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, in 2025.

He is currently a Postdoctoral Researcher with the Department of Information Engineering, University of Padova, Padova, Italy. His research interests include identification and control of quantized systems, privacy, and security in stochastic systems.



Jimin Wang (IEEE Member) received the B.S. degree in mathematics from Shandong Normal University, Jinan, China, in 2012, and the Ph.D. degree in mathematics from the School of Mathematics, Shandong University, Jinan, in 2018.

From 2017 to 2018, he was a joint Ph.D. student with the School of Electrical Engineering and Computing, The University of Newcastle, Callaghan, NSW, Australia. From 2018 to 2020, he was a Postdoctoral Researcher with the Institute of Systems Science, Chinese Academy of Sciences, Beijing, China. He is currently an Associate Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing. His current research interests include privacy and security in cyber-physical systems, stochastic systems, and networked control systems.

Dr. Wang was a recipient of Shandong University's excellent doctoral dissertation. He is an Associate Editor for *Systems and Control Letters* and *Journal of Automation and Intelligence*. He is a Member of the IEEE CSS Technical Committee on Security and Privacy, the IEEE CSS Technical Committee on Networks and Communication Systems, the IFAC Technical Committee 1.5 on Networked Systems.

Ji-Feng Zhang (IEEE Fellow) received the B.S. degree in mathematics from Shandong University, Jinan, China, in 1985, and the Ph.D. degree in operational research and cybernetics from the Institute of Systems Science, Chinese Academy of Sciences (CAS), Beijing, China, in 1991.

He is currently with the School of Automation and Electrical Engineering, Zhongyuan University of Technology, Zhengzhou, China, and the State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, CAS. His current research interests include system modeling, adaptive control, stochastic systems, and multiagent systems .

Dr. Zhang received the Second Prize of the State Natural Science Award of China in 2010 and 2015, respectively. He was a Vice President of the Chinese Association of Automation, the Chinese Mathematical Society, and the Systems Engineering Society of China. He was a Vice Chair of the IFAC Technical Board; a Member of the Board of Governors of the IEEE Control Systems Society; a Convenor of Systems Science Discipline of the Academic Degree Committee of the State Council of China. He was an Editor-in-Chief, Deputy Editor-in-Chief, or Associate Editor for more than ten journals, including *Science China Information Sciences*, IEEE TRANSACTIONS ON AUTOMATIC CONTROL, and *SIAM Journal on Control and Optimization*. He is an IFAC Fellow, CAA Fellow, SIAM Fellow, Member of the European Academy of Sciences and Arts, and Academician of the International Academy for Systems and Cybernetic Sciences.

Authorized licensed use limited to: Jifeng Zhang. Downloaded on July 01,2026 at 01:29:39 UTC from IEEE Xplore. Restrictions apply.