Differentially Private Bipartite Consensus over Signed Networks with Time-Varying Noises

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

***Abstract*— This paper investigates the differentially private bipartite consensus problem over signed networks. To solve this problem, a new algorithm is proposed by adding noises with time-varying variances to the cooperative-competitive interactive information. In order to achieve the privacy protection, the variances of the added noises are allowed to increase, which are substantially different from the existing works. In addition, the variances of the added noises can be either decaying or constant. By using a time-varying step-size based on the stochastic approximation method, we show that the algorithm converges in mean-square and almost-surely even with increasing privacy noises. We further develop a method to design the step-size and the noise parameter, affording the algorithm to achieve the average bipartite consensus with the desired accuracy and the predefined differential privacy level. Moreover, we give the mean-square and almost-sure convergence rates of the algorithm, and the privacy level with different forms of the privacy noises. We also reveal the trade-off between the accuracy and the privacy, and extend the results to local differential privacy. Finally, a numerical example verifies the theoretical results and demonstrates the algorithm's superiority against existing methods.**

***Index Terms*— Multi-agent system; differential privacy; signed network; stochastic approximation; convergence rate.**

I. Introduction

DISTRIBUTED consensus control of multi-agent systems (MASs) is significant due to its numerous applications, such as energy internet [1], [2], cooperative guidance systems [3], and social networks [4]. Generally, it refers to designing a network protocol such that all agents asymptotically reach an agreement. To date, many works have been developed on the consensus control of MASs, including average consensus [5]–[11], max consensus [12], group consensus [13], [14], and bipartite consensus [15]–[19]. Among others, cooperative and competitive interactions exist simultaneously in many complex systems, such as social networks, duopolistic markets in economic systems, teams opposed in a sports match, competing international alliances, biology systems, and two-party political systems [15], [20], [21]. For example, in social networks, each agent controls a time-dependent state variable, which denotes its opinion on some issue. Each agent updates its opinion based on its own current opinion, the current opinions of its neighbors, and its relationships (friendship or antagonism) with its neighbors. For those neighbors with friendship, the agent trusts their opinions; for those neighbors with antagonism, the agent does not trust their opinions, and takes the opposite of their opinions in updating. For modeling such networks, signed graph theory and bipartite consensus problems were formulated in [15], where the agents achieved an agreement with identical values but opposite signs. Currently, some substantial progresses have been made for bipartite consensus control of MASs [15]–[18], [20], [21].

With the increasing need for privacy and security, preserving the privacy of data is required in many applications. For example, in social networks [22], exchanging opinions probably reveals individual privacy when potential attackers exist. Thus, privacy-preserving in social networks has become a hot research topic. In cooperative guidance systems [3], information interactions may expose the missiles' and the launch stations' location. Hence, a naturally arising problem is how to achieve a bipartite consensus while protecting each agent's sensitive information from being inferred by potential attackers.

To address the requirement for privacy protection in distributed control, some methods have been proposed recently to counteract such potential privacy breaches, such as homomorphic encryption [23], [24], adding noises [25]–[27], time-varying transformation [28], [29], and state decomposition [30]. Homomorphic encryption allows direct calculation of encrypted data without revealing any information about the original text. But, such approaches incur a heavy communication and computation overhead. Accurate consensus is achieved by adding correlated noises to interaction information while protecting the initial states from semi-honest agents [25]–[27]. However, if the potential passive attackers obtain the information received and delivered by an agent, then this agent's initial state can be estimated through an iterative observer under such correlated noises mechanism. Generally speaking, current methods considering privacy preservation in average consensus assume that the honest-but-curious adversary cannot access the entire neighborhood set of an agent

The work was supported by National Key R&D Program of China under Grant 2018YFA0703800, National Natural Science Foundation of China under Grant 62203045 and Grant T2293770. The material in this paper was not presented at any conference. Corresponding author: Ji-Feng Zhang.

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

Jieming Ke and Ji-Feng Zhang are with the Key Laboratory of Systems and Control, Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, and also with the School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. (e-mail: jif@iss.ac.cn)

[23], [25], [28], [30].

Differential privacy techniques have been widely considered when publishing data from many technology companies, such as Google and Apple. Based on the original definition given by [31], ϵ-differential privacy has been extended to the multi-agent scenario, including protecting the agent's initial states in a consensus problem [32], protecting the objection function in distributed optimization [33], [34] and games [35], and protecting the global state trajectories in Kalman filtering [36], [37]. From a system control perspective, a tutorial and comprehensive framework of privacy security on control systems is provided in [38]. By adding uncorrelated noises on information, a differentially private consensus algorithm is designed for discrete-time MASs, where agents achieve unbiased convergence to the average almost-surely [39], [40]. In particular, it is interesting to know that adding one-shot noise at the beginning can achieve the optimal privacy and accuracy trade-off [39]. An ϵ-differentially private consensus algorithm is designed in [41] for continuous-time heterogeneous MASs, while an event-triggered scheme is proposed in [42] to reduce the control updates and ensure the ϵ-differentially private of the algorithm. Overall, the above-mentioned literature has two common grounds: 1) all algorithms are designed for average consensus, and 2) in order to guarantee the convergence and satisfy the differential privacy level, the privacy noises are required to be decaying exponentially to zero (or constant) with time. In fact, the interaction between practical systems involves cooperation and competitiveness simultaneously. Although the differential privacy bipartite consensus over signed graphs is considered, the privacy noises with exponential decay to zero are required in [43], [44]. Note that decaying noises to zero potentially exposes the trajectory of the state. Then, the following questions arise. Is it possible to give a more general noise form for privacy-preserving distributed consensus algorithm with guaranteed convergence? If possible, how do the added privacy noises affect the algorithm's convergence rate and privacy level? These questions motivate us to investigate the privacy-preserving bipartite consensus algorithm and relax the limitation of the existing privacy noise forms.

This paper designs a new differentially private bipartite consensus algorithm over signed networks. Specifically, each agent adds Laplace noises on the local state, and then transmits it to its neighbors. The added noises are with time-varying variances (which may increase with time). If the algorithm's step-size $\alpha(k)$ satisfies the stochastic approximation condition, then the algorithm can achieve the mean-square and almost-sure bipartite consensus. In summary, the contributions of this paper are fourfold:

- A new differentially private bipartite consensus algorithm is developed, compared with the existing literature [43], [44]. Specifically, in order to achieve privacy protection and avoid directly exposing the information about the state, the variances of the added noises are more general and allowed to increase. In addition, the variances of the added noises can be either decaying or constant, and cover the ones in [32], [39]–[44]. By employing a time-varying step-size based on the stochastic approximation method, both the mean-square and almost-sure average bipartite consensus of the algorithm are given even with increasing privacy noises.
- Both the mean-square and almost-sure convergence rates of the algorithm with different forms of privacy noises are given. To the best of our knowledge, it is the first to rigorously characterize both the mean-square and almost-sure convergence rates of distributed consensus with increasing noises. Even without considering privacy protection, our proof techniques fundamentally differ from existing counterparts and are of independent interest.
- A guideline for designing the time-varying step-size and the time-varying variances of the added noises is presented such that the algorithm can achieve the average bipartite consensus with the desired accuracy and predefined differential privacy level.
- The trade-off between the accuracy and the privacy is shown. When the variances of the added noises increase, both the mean-square average bipartite consensus and differential privacy with a finite privacy level over the infinite time horizon are established. Hence, our algorithm is effective for protecting the infinite time sequences of the state with guaranteed convergence, which is superior to the algorithms in [32], [39], [40], [42]–[44].

It is worth noting that this paper's results are significantly different from the literature. A comparison with the state-of-the-art is given as follows. Regarding the noise-perturbation approaches, we remove the conditions requiring the added noises are exponentially decaying [32], [39], [40], [42]–[44], or with constant variance [41]. Furthermore, compared with [45] only considering eavesdroppers, we consider eavesdroppers and honest-but-curious agents simultaneously. Compared with [23], [25], [28], [30], we remove the condition requiring that the adversary has no access to a target agent's communications with all of its neighbors, and hence, protect a more robust privacy of agents regardless of any auxiliary information an adversary may have. Compared with [7], [9]–[11], we consider the increasing noises case, and obtain both mean-square and almost-sure convergence rates of the algorithm. Moreover, we generalize communication topologies from unsigned graphs [25], [26], [39]–[42] to a class of signed graphs.

This paper is organized as follows. Section II provides the preliminaries and the problem statement. Section III introduces the algorithm's convergence and privacy analysis, while Section IV presents a numerical example. Finally, Section V concludes this work.

Notation. Denote $\mathbb{R}$, $\mathbb{N}$ as the sets of the real numbers and nonnegative integers, respectively. Let $\mathbb{R}^n$ be the n-dimensional real space, and $\mathbb{R}^{n\times m}$ be a set of $n\times m$ real matrices. I_n represents $n\times n$ identity matrix and $\mathbf{1}_n$ is an n-dimension column vector with all elements being 1. The notation $\mathrm{diag}(b_1,...,b_N)$ denotes the diagonal matrix with diagonal elements $b_1,\ldots,b_N$. For a random variable $X\in\mathbb{R}$, $\mathbb{E}X$ and $\mathrm{Var}(X)$ denote the expectation and variance of X, respectively. $\mathrm{Lap}(\mu,b)$ denotes the Laplace distribution with mean μ and scale parameter b. $\Gamma(x)=\int_0^{+\infty}t^{x-1}e^{-t}\mathrm{d}t$ is the gamma function and $\Gamma(x,z)=\int_z^{+\infty}t^{x-1}e^{-t}\mathrm{d}t$ is the upper incomplete gamma function. For sequences $\{f(k),k=$

$0, 1, ...\}$ and $\{g(k), k = 0, 1, ...\}$, $f(k) = O(g(k))$ means that there exist positive A and k_0 such that $|\frac{f(k)}{g(k)}| \leq A$ for all $k > k_0$. For any $x \in \mathbb{R}$, $\text{sgn}(x)$ is the sign function defined as $\text{sgn}(x) = 1$ if $x > 0$; -1 if $x < 0$; and 0 if $x = 0$. For square matrices $A_l, \ldots, A_k$, denote $\prod_{i=l}^{k} A_i = A_k \cdots A_l$ for $k \geq l$ and $\prod_{i=k+1}^{k} A_i = I_n$. For $x \in \mathbb{R}^n$, $\|x\|_1 = \sum_{i=1}^{n} |x_i|$, $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.

II. Preliminaries and Problem formulation

A. Graph theory

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ be an undirected signed graph with a set of agents $\mathcal{V} = \{1, 2, \ldots, N\}$, a set of edges $\mathcal{E} \in \mathcal{V} \times \mathcal{V}$, and a weighted adjacency matrix $\mathcal{A} = (a_{ij})_{N\times N}$. Agent i represents the i-th system, and an edge e_{ji} in the graph is denoted by the ordered pair agents $\{j, i\}$. $\{j, i\} \in \mathcal{E}$ if and only if Agent i can obtain the information from Agent j. For the adjacency matrix $\mathcal{A}$, $a_{ij} \neq 0$ if $\{j, i\} \in \mathcal{E}$, and $a_{ij} = 0$, otherwise. Specifically, the interaction between Agents i and j is cooperative if $a_{ij} > 0$, and competitive if $a_{ij} < 0$. We assume there is no self-loop in the graph $\mathcal{G}$, i.e., $a_{ii} = 0$. Let $\mathcal{N}_i = \{j | \{j, i\} \in \mathcal{E}\}$ be the set of Agent i's neighbors. The Laplacian matrix $\mathcal{L} = (l_{ij})_{N\times N}$ of graph $\mathcal{G}$ is defined as $l_{ii} = \sum_{k=1,k\neq i}^{N} |a_{ik}|$ and $l_{ij} = -a_{ij}$ if $i \neq j$. We denote $c_i = \sum_{j\in\mathcal{N}_i} |a_{ij}|$ as the degree of Agent i. For a signed graph, we define the greatest degree and the smallest degree as $c_{\max} = \max\{c_i, i \in \mathcal{V}\}$ and $c_{\min} = \min\{c_i, i \in \mathcal{V}\}$. Furthermore, structural balance is defined as follows.

Definition 2.1 (Structural balance, [15]): A signed graph $\mathcal{G}$ is structurally balanced if $\mathcal{V}$ can be divided into two disjoint subsets $\mathcal{V}_1$ and $\mathcal{V}_2$ (i.e., $\mathcal{V}_1 \bigcup \mathcal{V}_2 = \mathcal{V}$ and $\mathcal{V}_1 \bigcap \mathcal{V}_2 = \emptyset$) such that $a_{ij} \geq 0$ for $\forall i, j \in \mathcal{V}_h (h \in \{1, 2\})$, and $a_{ij} \leq 0$ for $\forall i \in \mathcal{V}_h, j \in \mathcal{V}_q, h \neq q, (h, q \in \{1, 2\})$.

Assumption 2.1: The signed graph $\mathcal{G}$ is connected and structurally balanced.

Remark 2.1: Assumption 2.1 is an important assumption on the signed graph $\mathcal{G}$, which is commonly used to ensure the bipartite consensus [15]–[18]. From Definition 2.1, a graph is still said to be structurally balanced if $\mathcal{V}_1$ or $\mathcal{V}_2$ is empty. Obviously, the graph with nonnegative weights in traditional consensus problem [32], [39] satisfies Assumption 2.1.

Lemma 2.1 ([15]): If Assumption 2.1 holds, then

1) A diagonal matrix $S = \text{diag}(s_1, s_2, \ldots, s_N)$ exists, such that $S\mathcal{A}S$ has all nonnegative elements, where $s_i \in \{1, -1\}$, for all $i \in \mathcal{V}$.
2) The Laplacian matrix associated with the corresponding unsigned graph $\mathcal{L}_S = S\mathcal{L}S$ is positive semi-definite.
3) The eigenvalue $\lambda_k(\mathcal{L}), k = 1, 2, \ldots, N$ of the Laplacian matrix $\mathcal{L}$ satisfies $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \leq \ldots \leq \lambda_N(\mathcal{L})$.

Remark 2.2: Under Assumption 2.1, we have $\mathbf{1}_N^T S\mathcal{L} = 0$. In particular, if $S = I$, then the problem studied in this paper is reduced to the traditional consensus problem [32], [39]. Besides, although we consider the undirected graph case here, it is not difficult to extend it to the digraph case [5].

B. Problem formulation

Consider a set of N agents coupled by an undirected signed graph $\mathcal{G}$. The dynamics of the i-th agent are as follows

$$x_i(k + 1) = x_i(k) + u_i(k), \tag{1}$$

where $x_i(k) \in \mathbb{R}$ is the state of Agent i with initial value $x_i(0)$, and $u_i(k) \in \mathbb{R}$ is the control input. To achieve the bipartite consensus of the system (1), [16] designs the following distributed controller: $u_i(k) = -\sum_{j\in\mathcal{N}_i} |a_{ij}|(x_i(k) - \text{sgn}(a_{ij})x_j(k))$, where $x_j(k)$ is the information that Agent i receives from its neighbors j.

In distributed bipartite consensus, *eavesdroppers* or *honest-but-curious (semi-honest) agents* may exist in the network [29], [35]. Note that the *honest-but-curious agents* might collude and attempt to deduce information about the initial state values of the other honest agents from the information they receive. *Eavesdroppers* are external adversaries who steal information through wiretapping all communication channels and intercepting exchanged information between agents. An *honest-but-curious agent* i has access to the internal state x_i, which is unavailable to external *eavesdroppers*. However, an *eavesdropper* has access to all shared information in the network, whereas an *honest-but-curious agent* can only access the shared information destined to it. These two attacker types are collectively called *passive attackers*. If the network has *passive attackers*, then delivering $\{x_i(k)|k \geq 0\}$ directly for each agent may leak its privacy, including the state $x_i(k)$ and the initial opinion or belief $x_i(0)$. Therefore, direct communication of intermediate results in the above controller can lead to severe privacy leakage of each agent's sensitive information. It is imperative to provide a theoretical privacy guarantee on each agent's sensitive information. To do so, each agent i sends to its neighbors the masking information $y_i(k)$ instead of the original information $x_i(k)$.

C. Differential privacy

A mechanism $\mathcal{M}(\cdot)$ is a stochastic map from a private dataset D to an observation O. In this paper, we focus on protecting the initial states of each agent against passive attackers. Thus, the private dataset is $D = \{x_i(0), i \in \mathcal{V}\}$, and the observation is $O = \{y_i(k), i \in \mathcal{V}\}_{k=0}^{T}$ with the time horizon $T \geq 1$. Then, we introduce the ϵ-differential privacy for the private dataset.

Definition 2.2 ([39]): Given $\delta > 0$, the initial states $D = \{x_i(0), i \in \mathcal{V}\}$ and $D' = \{x_i'(0), i \in \mathcal{V}\}$ are δ-adjacent if there exists $i_0 \in \mathcal{V}$, such that

$$|x_i(0) - x_i'(0)| \leq \begin{cases} \delta & \text{if } i = i_0; \\ 0 & \text{if } i \neq i_0. \end{cases}$$

Based on the above definition, inspired by [36] and [39], a definition of differential privacy is given for the differentially private bipartite consensus as follows.

Definition 2.3 (Differential privacy): Given $\delta > 0$, a mechanism $\mathcal{M}(\cdot)$ is ϵ-differentially private if $\mathbb{P}\{\mathcal{M}(D) \in \mathcal{O}\} \leq e^{\epsilon}\mathbb{P}\{\mathcal{M}(D') \in \mathcal{O}\}$ holds for any two δ-adjacent initial state sets $D = \{x_i(0), i \in \mathcal{V}\}$, $D' = \{x_i'(0), i \in \mathcal{V}\}$ and an observation set $\mathcal{O} \subseteq (\mathbb{R}^N)^{\mathbb{N}}$.

By Definition 2.3, the privacy level ϵ is non-negative, and a smaller ϵ corresponds to a stronger privacy protection.

Next, the mean-square average bipartite consensus, almost-sure average bipartite consensus, and (m, r)-accuracy are defined as follows, respectively.

Definition 2.4 (Mean-square average bipartite consensus): The system (1) is said to achieve the mean-square average bipartite consensus if there exists a random variable $x^\star$ with $\mathbb{E}x^\star = \frac{1}{N}\sum_{i=1}^{N} s_i x_i(0)$, $\mathbb{E}[x^\star]^2 < \infty$, such that $\lim_{k\to\infty}\mathbb{E}[x_i(k) - s_i x^\star]^2 = 0, s_i \in \{1, -1\}, \forall\ i \in \mathcal{V}$.

Definition 2.5 (Almost-sure average bipartite consensus): The system (1) is said to achieve the almost-sure average bipartite consensus if there exists a random variable $x^\star$ with $\mathbb{E}x^\star = \frac{1}{N}\sum_{i=1}^{N} s_i x_i(0)$, $\mathbb{E}[x^\star]^2 < \infty$, such that $\lim_{k\to\infty} x_i(k) = s_i x^\star, s_i \in \{1, -1\}, \forall\ i \in \mathcal{V}$.

Definition 2.6 (Accuracy): For $m \in [0, 1]$ and $r > 0$, the system (1) is said to achieve an (m, r)-accuracy, if the mean-square and almost-sure average bipartite consensus is achieved, and the $x^\star$ in Definition 2.4 (or 2.5) satisfies $\mathbb{P}\{|x^\star - \frac{1}{N}\sum_{i=1}^{N} s_i x_i(0)| \le r\} \ge 1 - m$.

Problems of interest: In this paper, the following questions are answered:

- How to design a more general privacy noise form for average bipartite consensus to achieve a better privacy protection with guaranteed convergence?
- What are the mean-square and almost-sure convergence rates under the influence of privacy noises?
- How to design the distributed protocol $u_i(k)$ and privacy noises for the desired accuracy $(m^\star, r^\star)$ and predefined differential privacy level $\epsilon^\star$?

III. Main Result

This section first presents the convergence analysis, ensuring that the mean-square and almost-sure average bipartite consensus is achieved under certain conditions. In addition to the mean-square and almost-sure convergence rates, the privacy analysis is also given by introducing the definition of the sensitivity on private datasets. Finally, we discuss the trade-off between accuracy and privacy, and extend the results to local differential privacy.

A. Algorithm

This subsection introduces a differentially private bipartite consensus algorithm, which is given in Algorithm 1.

Set

$$x(k) = \begin{bmatrix} x_1(k) & x_2(k) & \dots & x_N(k) \end{bmatrix}^T,$$
$$y(k) = \begin{bmatrix} y_1(k) & y_2(k) & \dots & y_N(k) \end{bmatrix}^T,$$
$$\omega(k) = \begin{bmatrix} \omega_1(k) & \omega_2(k) & \dots & \omega_N(k) \end{bmatrix}^T.$$

Then, the equation (4) can be rewritten in a compact form as follows:

$$x(k+1) = (I_N - \alpha(k)\mathcal{L})\, x(k) + \alpha(k)\mathcal{A}\omega(k). \tag{5}$$

Remark 3.1: In order to achieve the privacy protection, we add noises to Agent i's state before transmitting it to its neighbors. Different from the existing literature, the privacy noises added in (2) are more general. Specifically, the privacy noises used in this paper are random with variances of $2b^2(k)$, which are not required to decay to zero. Therefore, the state's information is not directly inferred with time. However, this brings convergence difficulties with the corresponding privacy analysis described next. Moreover, it is worth noting that we employ a time-varying step-size $\alpha(k)$, making the controller more flexible than that utilizing a constant step-size. Note that if $\alpha(k)$ is set to constant, as in current literature, the above closed-loop system cannot achieve the convergence because of the influence of privacy noises. To this end, we apply the stochastic approximation method to design a time-varying step-size.

Algorithm 1 A differentially private bipartite consensus algorithm

Input: Initial state sequence $\{x_i(0)\}$, step-size sequence $\{\alpha(k)\}$, and noise parameter sequence $\{b(k)\}$.
Output: State sequence $\{x_i(k)\}$.
for $k = 0, 1, ...,$ **do**

• **Information transmission**: Each agent i generates a privacy noise $\omega_i(k)$ with distribution $\mathrm{Lap}(0, b(k))$, and sends to its neighbors the following information instead of the original information $x_i(k)$.

$$y_i(k) = x_i(k) + \omega_i(k), i \in \mathcal{V}, k \in \mathbb{N}, \tag{2}$$

• **State update**: Each agent i receives $y_j(k)$ from its neighbor j and updates its own state by using the following privacy-preserving distributed controller:

$$u_i(k) = -\alpha(k)\sum_{j\in\mathcal{N}_i} |a_{ij}|(x_i(k) - \mathrm{sgn}(a_{ij})y_j(k)), \tag{3}$$

where $\alpha(k)$ is a positive time-varying step-size. Then, each agent i updates its own state as follows.

$$x_i(k+1) = x_i(k) - \alpha(k)\sum_{j\in\mathcal{N}_i} |a_{ij}|(x_i(k) - \mathrm{sgn}(a_{ij})y_j(k)). \tag{4}$$

end for

B. Convergence analysis

This subsection first proves that Algorithm 1 can achieve the mean-square and almost-sure average bipartite consensus. Then, we provide a method to design the step-size $\alpha(k)$ and the noise parameter $b(k)$ to ensure the $(m^\star, r^\star)$-accuracy.

For the step-size $\alpha(k)$ and the noise parameter $b(k)$, we give the following assumption.

Assumption 3.1: The step-size $\alpha(k)$ and the noise parameter $b(k)$ are positive and satisfy one of the following conditions:

a) $\sup_k \alpha(k) \le \frac{1}{\lambda_N(\mathcal{L})}$, $\sum_{k=0}^{\infty}\alpha(k) = \infty$, $\lim_{k\to 0}\alpha(k)b^2(k) = 0$;

b) $\sup_k \alpha(k) \le \frac{1}{\lambda_N(\mathcal{L})}$, $\sum_{k=0}^{\infty}\alpha(k) = \infty$, $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k) < \infty$.

Remark 3.2: Assumption 3.1 a) is weaker than Assumption 3.1 b). For example, if we take $\alpha(k) = \frac{a_1}{(k+a_2)^\beta}$ with a sufficiently small a_1 and $\beta \in [0, 1]$, and $b(k) = (k + a_2)^\gamma$,

then Assumption 3.1 a) holds when $2\gamma < \beta$. Assumption 3.1 b) holds when $2\gamma < 2\beta - 1$. Especially, when $b(k)$ is a constant, Assumption 3.1 becomes the commonly used stochastic approximation step-size [7]. Furthermore, when the step-size $\alpha(k)$ is a constant, and the privacy noises decay exponentially to zero [32], [39], [40], [42]–[44], Assumption 3.1 still holds.

Remark 3.3: The step-size $\alpha(k)$ and the noise parameter $b(k)$ are assumed to be same for all agents in Algorithm 1. Its practical implementation is an issue worthy of attention. This issue can be solved by implementing the following protocol before running Algorithm 1. Firstly, different $a_{1,i}$, $a_{2,i}$, β_i, and γ_i are chosen by the agents. Secondly, a consensus protocol (e.g., finite-time average consensus protocol [8] or max-consensus protocol [12]) is applied to obtain the same a_1, a_2, β, and γ. Thirdly, set $\alpha(k) = \frac{a_1}{(k+a_2)^\beta}$ and $b(k) = (k+a_2)^\gamma$.

To get the convergence results, the following independence assumption on the privacy noises is required.

Assumption 3.2: $\omega_i(k)$ and $\omega_j(l)$ are independent when $i \neq j$ or $k \neq l$.

Theorem 3.1: If Assumptions 2.1, 3.1 a), and 3.2 hold, then $\lim_{k\to\infty} \mathbb{E}[s_i x_i(k) - s_j x_j(k)]^2 = 0$, $\forall\ i, j \in \mathcal{V}$.

Proof: Let $z(k) = Sx(k)$ and $\mathcal{L}_S = S\mathcal{L}S$. Then, from (5) and $S^{-1} = S$ it follows that

$$z(k+1) = (I_N - \alpha(k)\mathcal{L}_S)\, z(k) + \alpha(k) S\mathcal{A}\omega(k). \quad (6)$$

Let $J = (1/N)\mathbf{1}_N\mathbf{1}_N^T$, $\delta(k) = (I_N - J)z(k)$ and $V(k) = \|\delta(k)\|^2$. Note that $\mathbf{1}_N^T\mathcal{L}_S = 0$. Then, $\mathcal{L}_S J = 0$, and from (6) we further have

$$\begin{aligned}
&\delta(k+1)\\
=&(I_N - J)z(k+1)\\
=&z(k) - \alpha(k)\mathcal{L}_S z(k) + \alpha(k)S\mathcal{A}\omega(k)\\
&- Jz(k) - \alpha(k)JS\mathcal{A}\omega(k)\\
=&\delta(k) - \alpha(k)\mathcal{L}_S z(k) + \alpha(k)(I_N - J)S\mathcal{A}\omega(k)\\
=&\big[I_N - \alpha(k)\mathcal{L}_S\big]\delta(k) + \alpha(k)(I_N - J)S\mathcal{A}\omega(k).
\end{aligned}$$

Note that $J\delta(k) = 0$. Then, we have

$$\begin{aligned}
&\delta(k+1)\\
=&\big[I_N - J - \alpha(k)\mathcal{L}_S\big]\delta(k) + \alpha(k)(I_N - J)S\mathcal{A}\omega(k). \quad (7)
\end{aligned}$$

Since $\lambda_2(\mathcal{L}_S) = \lambda_2(\mathcal{L})$ and $\lambda_N(\mathcal{L}_S) = \lambda_N(\mathcal{L})$, from Theorem 2.1 in [7], we have $\lambda_2(\mathcal{L})(I_N - J) \leq \mathcal{L}_S \leq \lambda_N(\mathcal{L})(I_N - J)$. Note that $\sup_k \alpha(k) \leq \frac{1}{\lambda_N(\mathcal{L})}$. Then, $I_N - J - \alpha(k)\mathcal{L}_S \geq (1 - \alpha(k)\lambda_N(\mathcal{L}))(I_N - J) \geq 0$. From (7), we have

$$\begin{aligned}
&V(k+1)\\
\leq&\big[1 - \alpha(k)\lambda_2(\mathcal{L})\big]^2 V(k)\\
&+ 2\alpha(k)\delta^T(k)\big[I_N - J - \alpha(k)\mathcal{L}_S\big]^T(I_N - J)S\mathcal{A}\omega(k)\\
&+ \alpha^2(k)w^T(k)\mathcal{A}^T S^T(I_N - J)^T(I_N - J)S\mathcal{A}\omega(k).
\end{aligned}$$

Define σ-algebra $\mathcal{F}_k^\omega = \sigma\{\omega(0), \omega(1), \omega(2), \ldots, \omega(k-1)\}$. Note that $\omega(k)$ is the zero-mean noise. Then, taking the conditional expectation with respect to $\mathcal{F}_k^\omega$ on both sides of the above equations, one can get

$$\begin{aligned}
&\mathbb{E}\left[V(k+1)|\mathcal{F}_k^\omega\right]\\
\leq&\big[1 - \alpha(k)\lambda_2(\mathcal{L})\big]^2 V(k) + 2\alpha^2(k)b^2(k)N\|\mathcal{A}\|^2. \quad (8)
\end{aligned}$$

Note that $\mathbb{E}\left[\mathbb{E}\left[V(k+1)|\mathcal{F}_k^\omega\right]\right] = \mathbb{E}V(k+1)$. Then, taking mathematical expectation on both sides of (8), we obtain

$$\begin{aligned}
&\mathbb{E}V(k+1)\\
\leq&\big[1 - \alpha(k)\lambda_2(\mathcal{L})\big]^2\mathbb{E}V(k) + 2\alpha^2(k)b^2(k)N\|\mathcal{A}\|^2\\
\leq&\mathbb{E}V(k) - \alpha(k)\lambda_2(\mathcal{L})\mathbb{E}V(k) + 2\alpha^2(k)b^2(k)N\|\mathcal{A}\|^2. \quad (9)
\end{aligned}$$

Then, by Lemma A.1 of [7], we have $\mathbb{E}V(k) = 0$, which further implies the result. □

Theorem 3.2: If Assumptions 2.1, 3.1 b), and 3.2 hold, then Algorithm 1 achieves the mean-square average bipartite consensus with $\mathrm{Var}(x^\star) = \frac{2\sum_{i\in\mathcal{V}} c_i^2}{N^2}\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)$.

Proof: Since the graph is structurally balanced, from Lemma 2.1, it follows that $\mathbf{1}_N^T\mathcal{L}_S = 0$, and

$$\begin{aligned}
\mathbf{1}_N^T z(k) =& (\mathbf{1}_N^T(I_N - \alpha(k-1)\mathcal{L}_S))z(k-1)\\
&+ \alpha(k-1)(\mathbf{1}_N^T S\mathcal{A})\omega(k-1)\\
=& \mathbf{1}_N^T z(k-1) + \alpha(k-1)(\mathbf{1}_N^T S\mathcal{A})\omega(k-1). \quad (10)
\end{aligned}$$

By iteration, we have

$$\mathbf{1}_N^T z(k) = \sum_{i\in\mathcal{V}} z_i(0) + \sum_{j=1}^{k}\alpha(j-1)(\mathbf{1}_N^T S\mathcal{A})\omega(j-1), \quad (11)$$

which immediately follows that

$$\lim_{k\to\infty}\mathbf{1}_N^T z(k) = \sum_{i\in\mathcal{V}} z_i(0) + \sum_{j=1}^{\infty}\sum_{i\in\mathcal{V}}\alpha(j-1)s_i c_i\omega_i(j-1).$$

By Theorem 3.1, set

$$x^\star = \frac{1}{N}\sum_{i\in\mathcal{V}} z_i(0) + \frac{1}{N}\sum_{j=1}^{\infty}\sum_{i\in\mathcal{V}}\alpha(j-1)s_i c_i\omega_i(j-1).$$

Then, we have

$$\begin{aligned}
&\lim_{k\to\infty}\sqrt{\mathbb{E}\left[s_i x_i(k) - x^\star\right]^2}\\
\leq&\lim_{k\to\infty}\sqrt{\mathbb{E}\left[s_i x_i(k) - \frac{1}{N}\mathbf{1}_N^T z(k)\right]^2}\\
&+ \lim_{k\to\infty}\sqrt{\mathbb{E}\left[\frac{1}{N}\mathbf{1}_N^T z(k) - x^\star\right]^2}\\
=&0.
\end{aligned}$$

By the fact that $\omega_i(k)$ are independent for all $i \in \mathcal{V}$, $k \in \mathbb{N}$, it is obtained that

$$\begin{aligned}
\mathbb{E}x^\star =& \mathbb{E}\left[\frac{1}{N}\sum_{i\in\mathcal{V}} z_i(0) + \frac{1}{N}\sum_{j=0}^{\infty}\sum_{i\in\mathcal{V}}\alpha(j)s_i c_i\omega_i(j)\right]\\
=& \frac{1}{N}\sum_{i\in\mathcal{V}} z_i(0) = \frac{1}{N}\sum_{i\in\mathcal{V}} s_i x_i(0),
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}\left(x^\star\right) =& \frac{1}{N^2}\sum_{j=0}^{\infty}\sum_{i\in\mathcal{V}}\alpha^2(j)\mathbb{E}\left[s_i c_i\omega_i(j)\right]^2\\
=& \frac{2\sum_{i\in\mathcal{V}} c_i^2}{N^2}\sum_{k=0}^{\infty}\alpha^2(k)b^2(k). \quad (12)
\end{aligned}$$

Since $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)<\infty$, $\mathrm{Var}\,(x^\star)$ is bounded. This completes the proof. □

Remark 3.4: $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)<\infty$ is necessary for a finite $\mathbb{E}[x^\star]^2$. Otherwise, when $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)=\infty$, one can get $\mathrm{Var}(x^\star)=\infty$ by (12).

The almost-sure convergence properties are important, because they represent what happen to individual trajectories of the stochastic iterations, which are instantiations of the algorithm actually used in practice. From the following theorem, the almost-sure average bipartite consensus of Algorithm 1 is achieved as well under Assumptions 2.1 and 3.1.

Theorem 3.3: If Assumptions 2.1, 3.1 b), and 3.2 hold, then Algorithm 1 achieves the almost-sure average bipartite consensus.

Proof: From (8) it follows that

$$\begin{aligned}&\mathbb{E}\left[V(k+1)|\mathcal{F}_k^{\omega}\right]\\ \le&V(k)-\alpha(k)\lambda_2(\mathcal{L})V(k)+2\alpha^2(k)b^2(k)N\|\mathcal{A}\|^2.\end{aligned}$$

Notice that $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)<\infty$. Then, by Lemma A.1, $V(k)$ converges almost-surely, and $\sum_{k=0}^{\infty}\alpha(k)V(k)<\infty$ almost-surely. Since $\sum_{k=0}^{\infty}\alpha(k)=\infty$, $V(k)$ converges to 0 almost-surely.

By (11), $\frac{1}{N}\sum_{i=1}^{N}z_i(k)=\frac{1}{N}\mathbf{1}_N^{\top}z_i(k)$ is a martingale. By Assumption 3.1, we have

$$\begin{aligned}&\mathbb{E}\left\|\sum_{j=1}^{k}\alpha(j-1)(\mathbf{1}_N^TS\mathcal{A})\omega(j-1)\right\|^2\\ \le&2\left\|\mathbf{1}_N^TS\mathcal{A}\right\|^2\sum_{j=1}^{k}\alpha^2(j-1)b^2(j-1)<\infty,\end{aligned}$$

which implies that $\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N}z_i(k)\right]^2<\infty$. This together with Theorem 7.6.10 of [46] and Theorem 3.2 implies that $\frac{1}{N}\sum_{i=1}^{N}z_i(k)$ converges to $x^\star$ almost-surely. Since $V(k)$ converges to 0 almost-surely, we have $z_i(k)$ converges to $\frac{1}{N}\sum_{i=1}^{N}z_i(k)$ almost-surely. This proves the theorem. □

Remark 3.5: By assuming that the graph is structurally balanced, the mean-square and almost-sure average bipartite consensus of Algorithm 1 is achieved. The results can be extended to the case where the graph is structurally unbalanced and the weighted adjacency matrix $\mathcal{A}$ satisfies the signed Perron-Frobenius property [19]. In this case, there exists $t_0>0$ such that $\mathcal{A}^{t_0}$ is the weighted adjacency matrix of a structurally balanced graph. The existence of t_0 is ensured by Theorem 2 of [19]. In Algorithm 1, instead of (3), each agent updates its own state by using $u_i(k)=-\alpha(k)\sum_{j\in\mathcal{N}_i}|a_{ij}|(x_i(k)-\mathrm{sgn}(a_{ij})\tilde{y}_j(k))$, where $\tilde{y}_j(k)$ is the jth component of $\mathcal{A}^{t_0-1}y(k)$. Similar to the proof of Theorems 3.1-3.3, the mean-square and almost-sure average bipartite consensus of the modified algorithm for the structurally unbalanced graph can be achieved.

Remark 3.6: Theorems 3.1-3.3 give a unified framework of the consensus analysis under different types of step-sizes $\alpha(k)$ and noise parameters $b(k)$, including the decaying $\alpha(k)$ and constant $b(k)$ considered in [7], [41], the constant $\alpha(k)$ and exponentially decaying $b(k)$ considered in [32], [39], [40], [42]–[44], and the decaying $\alpha(k)$ and increasing $b(k)$.

The following theorem provides a way to design the step-size $\alpha(k)$ and the noise parameter $b(k)$ to ensure the $(m^\star,r^\star)$-accuracy.

Theorem 3.4: Under Assumptions 2.1, 3.1 b), and 3.2, for any given a pair of parameters $(m^\star,r^\star)$, if

$$\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)\le\frac{m^\star(r^\star)^2N^2}{2\sum_{i\in\mathcal{V}}c_i^2},$$

then Algorithm 1 achieves the $(m^\star,r^\star)$-accuracy.

Proof: From the Chebyshev's inequality [46] it follows that

$$\mathbb{P}\left\{\frac{(x^\star-\mathbb{E}x^\star)^2}{\mathrm{Var}(x^\star)}<\epsilon\right\}\ge1-\frac{1}{\epsilon}.$$

Taking (12) into the above inequality yields $\mathbb{P}\left\{|x^\star-\mathbb{E}x^\star|<\sqrt{\epsilon\kappa}\right\}\ge1-\frac{1}{\epsilon}$, where $\kappa=\frac{2\sum_{i\in\mathcal{V}}c_i^2}{N^2}\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)$.

Set $r=\sqrt{\epsilon\kappa}$. Then, $\epsilon=\frac{r^2}{\kappa}$ and $\mathbb{P}\left\{|x^\star-\mathbb{E}x^\star|<r\right\}\ge1-\frac{\kappa}{r^2}$. Therefore, the (m,r)-accuracy is achieved with $m=\frac{2\sum_{i\in\mathcal{V}}c_i^2}{N^2r^2}\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)$.

Clearly, as long as $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)\le\frac{m^\star(r^\star)^2N^2}{2\sum_{i\in\mathcal{V}}c_i^2}$, the $(m^\star,r^\star)$-accuracy is ensured. □

Next, we further analyze the $(m^\star,r^\star)$-accuracy of Algorithm 1 with $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and $b(k)=\underline{b}(k+a_2)^\gamma$.

Corollary 3.1: Under Assumption 2.1, for any given a pair of parameters $(m^\star,r^\star)$, set $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and $b(k)=\underline{b}(k+a_2)^\gamma$, $\beta\in(0,1]$, $\gamma<\beta-1/2$, $a_1,a_2,\underline{b}>0$, such that

$$\frac{a_1^2\underline{b}^2a_2^{2\gamma-2\beta+1}}{2\beta-2\gamma-1}+a_1^2\underline{b}^2a_2^{2\gamma-2\beta}\le\frac{m^\star(r^\star)^2N^2}{2\sum_{i\in\mathcal{V}}c_i^2}.\tag{13}$$

Then, Algorithm 1 achieves the $(m^\star,r^\star)$-accuracy.

Proof: By the fact that $f(x)=\frac{a_1\underline{b}(x+a_2)^\gamma}{(x+a_2)^\beta}$ with $\beta\in(0,1]$, $\gamma<\beta-1/2$, a_1, a_2, $\underline{b}>0$, is a strictly decreasing function of $x>0$. Then, for $k\ge1$, we have

$$\left(\frac{a_1\underline{b}(k+a_2)^\gamma}{(k+a_2)^\beta}\right)^2\le\int_{k-1}^{k}\left(\frac{a_1\underline{b}(x+a_2)^\gamma}{(x+a_2)^\beta}\right)^2\mathrm{d}x,$$

and thus,

$$\begin{aligned}&\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)\\ =&a_1^2\underline{b}^2a_2^{2\gamma-2\beta}+\sum_{k=1}^{\infty}\left(\frac{a_1\underline{b}(k+a_2)^\gamma}{(k+a_2)^\beta}\right)^2\\ \le&a_1^2\underline{b}^2a_2^{2\gamma-2\beta}+\int_0^{\infty}\left(\frac{a_1\underline{b}(x+a_2)^\gamma}{(x+a_2)^\beta}\right)^2\mathrm{d}x\\ \le&-\frac{a_1^2\underline{b}^2a_2^{2\gamma-2\beta+1}}{2\gamma-2\beta+1}+a_1^2\underline{b}^2a_2^{2\gamma-2\beta}\\ \le&\frac{m^\star(r^\star)^2N^2}{2\sum_{i\in\mathcal{V}}c_i^2}.\end{aligned}$$

This completes the proof. □

Under the time-varying noises, the predefined accuracy is ensured by properly selecting the step-size $\alpha(k)$ and the noise parameter $b(k)$, $k\in\mathbb{N}$. Besides, we can enhance the accuracy by optimizing $\sum_{k=0}^{\infty}\alpha^2(k)b^2(k)$.

C. Convergence rate

In this subsection, we analyze the mean-square and almost-sure convergence rates of Algorithm 1. Regarding the algorithm's step-size and noise parameter, we give a step-size form $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and the noise parameter $b(k)=O(k^\gamma)$. First, we give the mean-square convergence rate of Algorithm 1 with $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and $b(k)=O(k^\gamma)$.

Theorem 3.5: Suppose Assumptions 2.1 and 3.2 hold. Let the step-size $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$, $b(k)=O(k^\gamma)$, $\beta\in(0,1]$, $\gamma<\beta-\frac{1}{2}$, $a_1,a_2>0$. Then, the mean-square convergence rate of Algorithm 1 is given as follows.
When $\beta\in(0,1)$, for all $i\in\mathcal{V}$, we have

$$\mathbb{E}\left[x_i(k)-s_ix^\star\right]^2=O\left(k^{1+2\gamma-2\beta}\right). \tag{14}$$

When $\beta=1$, for all $i\in\mathcal{V}$, we have

$$\mathbb{E}\left[x_i(k)-s_ix^\star\right]^2=\begin{cases}O\left(k^{-2a_1\lambda_2(\mathcal{L})}\right), & \gamma+a_1\lambda_2(\mathcal{L})<1/2;\\ O\left(k^{2\gamma-1}\ln k\right), & \gamma+a_1\lambda_2(\mathcal{L})=1/2;\\ O\left(k^{2\gamma-1}\right), & \gamma+a_1\lambda_2(\mathcal{L})>1/2.\end{cases} \tag{15}$$

Proof: For analyzing the mean-square convergence rate of Algorithm 1, we do it in the following three steps.

Step 1: We give the mean-square convergence rate of $s_ix_i(k)-\frac{1}{N}\mathbf{1}_N^Tz(k)$. When $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and $b(k)=O(k^\gamma),\beta\in(0,1],\gamma<\beta-\frac{1}{2}$, there exists $\varrho>0$ such that

$$2\alpha^2(k)b^2(k)N\|\mathcal{A}\|^2\le\frac{\varrho}{(k+a_2)^{2\beta-2\gamma}}. \tag{16}$$

Note that there exists a sufficiently large $k_0>0$ such that $1-\frac{2a_1\lambda_2(\mathcal{L})}{(k+a_2)^\beta}>0$ for all $k>k_0$. Then, from (9) and (16) it follows that

$$\mathbb{E}V(k+1)\le\left(1-\frac{2a_1\lambda_2(\mathcal{L})}{(k+a_2)^\beta}+\frac{a_1^2\lambda_2^2(\mathcal{L})}{(k+a_2)^{2\beta}}\right)\mathbb{E}V(k)+\frac{\varrho}{(k+a_2)^{2\beta-2\gamma}},\quad\text{as}\quad k>k_0.$$

Iterating the above process gives

$$\begin{aligned}&\mathbb{E}V(k+1)\\&\le\prod_{t=k_0}^{k}\left(1-\frac{2a_1\lambda_2(\mathcal{L})}{(t+a_2)^\beta}+\frac{a_1^2\lambda_2^2(\mathcal{L})}{(t+a_2)^{2\beta}}\right)\mathbb{E}V(k_0)\\&\quad+\sum_{l=k_0}^{k-1}\prod_{t=l+1}^{k}\left(1-\frac{2a_1\lambda_2(\mathcal{L})}{(t+a_2)^\beta}+\frac{a_1^2\lambda_2^2(\mathcal{L})}{(t+a_2)^{2\beta}}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\\&\quad+\frac{\varrho}{(k+a_2)^{2\beta-2\gamma}}.\end{aligned} \tag{17}$$

When $\beta=1$, from Lemma A.2 and (17) it follows that

$$\begin{aligned}&\mathbb{E}V(k+1)\\&=O\left(\frac{1}{(k+a_2)^{2a_1\lambda_2(\mathcal{L})}}\right)+O\left(\frac{1}{(k+a_2)^{2-2\gamma}}\right)\\&\quad+O\left(\frac{1}{(k+a_2)^{2a_1\lambda_2(\mathcal{L})}}\sum_{l=k_0}^{k-1}\frac{1}{(l+a_2)^{2-2\gamma-2a_1\lambda_2(\mathcal{L})}}\right).\end{aligned}$$

Note that

$$\begin{aligned}&\sum_{l=k_0}^{k-1}\frac{1}{(l+a_2)^{2-2\gamma-2a_1\lambda_2(\mathcal{L})}}\\&\le\int_{k_0-1}^{k}\frac{1}{(x+a_2)^{2-2\gamma-2a_1\lambda_2(\mathcal{L})}}\mathrm{d}x.\end{aligned}$$

Then, we have

$$\begin{aligned}&\mathbb{E}V(k+1)\\&=\begin{cases}O\left((k+a_2)^{-2a_1\lambda_2(\mathcal{L})}\right), & \gamma+a_1\lambda_2(\mathcal{L})<1/2;\\ O\left((k+a_2)^{2\gamma-1}\ln k\right), & \gamma+a_1\lambda_2(\mathcal{L})=1/2;\\ O\left((k+a_2)^{2\gamma-1}\right), & \gamma+a_1\lambda_2(\mathcal{L})>1/2.\end{cases}\end{aligned} \tag{18}$$

When $0<\beta<1$, there exists a sufficiently large $k_1\ge k_0$ such that for all $k\ge k_1$, $-\frac{2a_1\lambda_2(\mathcal{L})}{(k+a_2)^\beta}+\frac{a_1^2\lambda_2^2(\mathcal{L})}{(k+a_2)^{2\beta}}\le-\frac{a_1\lambda_2(\mathcal{L})}{(k+a_2)^\beta}$. Note that $(1-\frac{a_1\lambda_2(\mathcal{L})}{(l+a_2)^\beta})^{-1}\le2$ for all $l\ge k_1$. Then, from Lemma A.2 and (17) it follows that

$$\begin{aligned}&\mathbb{E}V(k+1)\\&\le\prod_{t=k_1}^{k}\left(1-\frac{a_1\lambda_2(\mathcal{L})}{(t+a_2)^\beta}\right)\mathbb{E}V(k_1)+\frac{\varrho}{(k+a_2)^{2\beta-2\gamma}}\\&\quad+2\sum_{l=k_1}^{k-1}\prod_{t=l}^{k}\left(1-\frac{a_1\lambda_2(\mathcal{L})}{(t+a_2)^\beta}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\\&=O\left(\exp\left(-\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2+1)^{1-\beta}\right)\right)\\&\quad+O\left(\frac{1}{(k+a_2)^{2\beta-2\gamma}}\right)\\&\quad+O\left(\sum_{l=k_1}^{k-1}\exp\left(-\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2+1)^{1-\beta}\right)\right.\\&\quad\left.\cdot\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\right).\end{aligned} \tag{19}$$

Note that $\frac{\beta-2\gamma}{a_1\lambda_2(\mathcal{L})(l+a_2)^{1-\beta}}<\frac{1}{2}$ for all $l\ge k_1$. Then, we have

$$\begin{aligned}&\sum_{l=k_1}^{k-1}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\\&\le\int_{k_1}^{k}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\mathrm{d}l\\&=\frac{1}{a_1\lambda_2(\mathcal{L})}\int_{k_1}^{k}\frac{\varrho}{(l+a_2)^{\beta-2\gamma}}\mathrm{d}\left(\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\right)\\&\le\frac{1}{a_1\lambda_2(\mathcal{L})}\frac{\varrho}{(k+a_2)^{\beta-2\gamma}}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2)^{1-\beta}\right)\\&\quad+\frac{1}{2}\int_{k_1}^{k}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\mathrm{d}l.\end{aligned}$$

Furthermore, we have

$$\begin{aligned}&\sum_{l=k_1}^{k-1}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{\varrho}{(l+a_2)^{2\beta-2\gamma}}\\&=O\left(\frac{1}{(k+a_2)^{\beta-2\gamma}}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2)^{1-\beta}\right)\right).\end{aligned}$$

From (19) it follows that

$$\begin{aligned}&\mathbb{E}V(k+1)\\=&O\left(\frac{1}{(k+a_2)^{2\beta-2\gamma}}\right)\\&+O\left(\exp\left(-\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2+1)^{1-\beta}\right)\right.\\&\left.\cdot\frac{1}{(k+a_2)^{\beta-2\gamma}}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2)^{1-\beta}\right)\right).\\=&O\left((k+a_2)^{2\gamma-\beta}\right).\end{aligned}\tag{20}$$

Step 2: We give the mean-square convergence rate of $\frac{1}{N}\mathbf{1}_N^Tz(k)-x^\star$. From (11) it follows that

$$\begin{aligned}\mathbb{E}\left[\frac{1}{N}\mathbf{1}_N^Tz(k)-x^\star\right]^2=&\frac{2\sum_{i\in\mathcal{V}}c_i^2}{N^2}\sum_{j=k+1}^{\infty}\alpha^2(j)b^2(j)\\=&O\left(\sum_{j=k+1}^{\infty}\frac{1}{(j+a_2)^{2\beta-2\gamma}}\right).\end{aligned}$$

Note that $\gamma<\beta-\frac{1}{2}$ and $\sum_{j=k+1}^{\infty}\frac{1}{(j+a_2)^{2\beta-2\gamma}}\leq\int_k^{\infty}\frac{1}{(x+a_2)^{2\beta-2\gamma}}\mathrm{d}x$. Then, we have

$$\mathbb{E}\left[\frac{1}{N}\mathbf{1}_N^Tz(k)-x^\star\right]^2=O\left((k+a_2)^{1+2\gamma-2\beta}\right).\tag{21}$$

Step 3: We give the mean-square convergence rate of Algorithm 1. Note that

$$\begin{aligned}&\mathbb{E}\left[x_i(k)-s_ix^\star\right]^2\\=&\mathbb{E}\left[s_ix_i(k)-\frac{1}{N}\mathbf{1}_N^Tz(k)+\frac{1}{N}\mathbf{1}_N^Tz(k)-x^\star\right]^2\\\leq&2\mathbb{E}\left[s_ix_i(k)-\frac{1}{N}\mathbf{1}_N^Tz(k)\right]^2+2\mathbb{E}\left[\frac{1}{N}\mathbf{1}_N^Tz(k)-x^\star\right]^2.\end{aligned}\tag{22}$$

Then, when $\beta=1$, from (18), (21) and (22) it follows that (15) holds; when $0<\beta<1$, from (20), (21) and (22) it follows that (14) holds. The proof is completed. □

In the following, we give the almost-sure convergence rate of Algorithm 1 with $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and $b(k)=O(k^\gamma)$.

Theorem 3.6: Suppose Assumptions 2.1 and 3.2 hold. Let the step-size $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$, $b(k)=O(k^\gamma)$, $\beta\in(0,1]$, $\gamma<\beta-1/2$, $a_1,a_2>0$. Then, the almost-sure convergence rate of Algorithm 1 is given as follows.

When $\beta\in(0,1)$, for any $\eta\in\left(\frac{1}{4},\frac{\beta/2-\gamma}{1-\beta}\right)$ and all $i,j\in\mathcal{V}$, we have

$$s_ix_i(k)-s_jx_j(k)=O\left(k^{\gamma+\eta-(\eta+1/2)\beta}\right),\quad\text{a.s.}\tag{23}$$

When $\beta=1$, for all $i,j\in\mathcal{V}$, we have

$$\begin{aligned}&s_ix_i(k)-s_jx_j(k)\\=&\begin{cases}O\left(k^{\gamma-1/2}\sqrt{\ln\ln k}\right), & a_1\lambda_2(\mathcal{L})+\gamma>1/2;\\O\left(k^{\gamma-1/2}\sqrt{\ln k\ln\ln\ln k}\right), & a_1\lambda_2(\mathcal{L})+\gamma=1/2;\\O\left(k^{-a_1\lambda_2(\mathcal{L})}\right), & a_1\lambda_2(\mathcal{L})+\gamma<1/2.\end{cases}\quad\text{a.s.}\end{aligned}\tag{24}$$

Proof: As clarified in Theorem 3.1, it is equivalent to calculate the convergence rate of $\delta(k)=z(k)-Jz(k)$, where $z(k)$ and J are defined in Theorem 3.1.

Note that

$$\begin{aligned}\|\delta(k)\|=&\sup_{\|v\|=1}|v^T\delta(k)|\\=&\sup_{\substack{p^2\|e\|^2+Nq^2=1\\e^T\mathbf{1}=0}}|(pe+q\mathbf{1})^T\delta(k)|\\=&\sup_{\substack{p^2\|e\|^2+Nq^2=1\\e^T\mathbf{1}=0}}|pe^Tz(k)|\\=&\sup_{\substack{\|e\|=1\\e^T\mathbf{1}=0}}|e^Tz(k)|,\end{aligned}$$

and there exists $e_1,\ldots,e_{N-1}$ such that $e_i^Te_j=0$ for $i\neq j$, $\|e_i\|=1$ and $e_i^T\mathbf{1}=0$ for $i=1,2,\ldots,N-1$. Then, we have

$$\begin{aligned}\sup_{\substack{\|e\|=1\\e^T\mathbf{1}=0}}|e^Tz(k)|=&\sup_{\sum_{i=1}^{N-1}p_i^2=1}\sum_{i=1}^{N-1}|p_i||e_i^Tz(k)|\\\leq&\sup_{\sum_{i=1}^{N-1}p_i^2=1}\sqrt{N\sum_{i=1}^{N-1}|p_i|^2}\max_i|e_i^Tz(k)|\\=&\sqrt{N}\max_i|e_i^Tz(k)|.\end{aligned}$$

Set

$$\tilde{\mathcal{D}}=\begin{bmatrix}e_1&\cdots&e_{N-1}\end{bmatrix}^T,\quad\mathcal{D}=\begin{bmatrix}e_1&\cdots&e_{N-1}&\frac{\mathbf{1}}{\sqrt{N}}\end{bmatrix}^T.$$

Then, to calculate the convergence rate of $\delta(k)$, it suffices to analyze that of $\tilde{\mathcal{D}}z(k)$.

From the properties of e_i, we have $\mathcal{D}^T\mathcal{D}=I_N$. Let $\tilde{\mathcal{L}}=\tilde{\mathcal{D}}\mathcal{L}_S\tilde{\mathcal{D}}^T$. Then, $\mathcal{D}\mathcal{L}_S\mathcal{D}^T=\mathrm{diag}(\tilde{\mathcal{L}},0)$ and $\lambda_{\min}(\tilde{\mathcal{L}})=\lambda_2(\mathcal{L})$.

From $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$ and (6) it follows that

$$\begin{aligned}&z(k+1)\\=&\left(I_N-\frac{a_1}{(k+a_2)^\beta}\mathcal{L}_S\right)z(k)+\frac{a_1b(k)}{(k+a_2)^\beta}S\mathcal{A}\frac{\omega(k)}{b(k)}.\end{aligned}$$

Hence, we have

$$\begin{aligned}&\mathcal{D}z(k+1)\\=&\left(I_N-\frac{a_1}{(k+a_2)^\beta}\mathcal{D}\mathcal{L}_S\mathcal{D}^T\right)\mathcal{D}z(k)\\&+\frac{a_1b(k)}{(k+a_2)^\beta}\mathcal{D}S\mathcal{A}\frac{\omega(k)}{b(k)}\\=&\begin{bmatrix}I_{N-1}-\frac{a_1}{(k+a_2)^\beta}\tilde{\mathcal{L}}&0\\0&1\end{bmatrix}\mathcal{D}z(k)+\frac{a_1b(k)}{(k+a_2)^\beta}\mathcal{D}S\mathcal{A}\frac{\omega(k)}{b(k)},\end{aligned}$$

which implies

$$\begin{aligned}&\tilde{\mathcal{D}}z(k+1)\\=&\left(I_{N-1}-\frac{a_1}{(k+a_2)^\beta}\tilde{\mathcal{L}}\right)\tilde{\mathcal{D}}z(k)+\frac{a_1b(k)}{(k+a_2)^\beta}\tilde{\mathcal{D}}S\mathcal{A}\frac{\omega(k)}{b(k)}.\end{aligned}$$

Iterating the above equation results in

$$\begin{aligned}&\tilde{\mathcal{D}}z(k)\\=&\sum_{l=0}^{k-1}\prod_{i=l+1}^{k-1}\left(I_{N-1}-\frac{a_1}{(i+a_2)^\beta}\tilde{\mathcal{L}}\right)\frac{a_1b(l)}{(l+a_2)^\beta}\tilde{\mathcal{D}}S\mathcal{A}\frac{w(l)}{b(l)}\\&+\prod_{i=0}^{k-1}\left(I_{N-1}-\frac{a_1}{(i+a_2)^\beta}\tilde{\mathcal{L}}\right)\tilde{\mathcal{D}}z(0).\end{aligned}\tag{25}$$

Note that when $0<\beta<1$, by Lemma A.2, we have

$$\begin{aligned}&\prod_{i=l+1}^{k-1}\left\|I_{N-1}-\frac{a_1}{(i+a_2)^\beta}\tilde{\mathcal{L}}\right\|\\\le&\prod_{i=l+1}^{k-1}\left(1-\frac{a_1\lambda_2(\mathcal{L})}{(i+a_2)^\beta}\right)\\=&O\left(\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}\left((l+a_2)^{1-\beta}-(k+a_2)^{1-\beta}\right)\right)\right).\end{aligned}\tag{26}$$

According to the Lemma 2 in [47] and Lemma A.3, for any $\eta>\frac{1}{4}$, we have

$$\begin{aligned}&\sum_{l=0}^{k-1}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{a_1b(l)}{(l+a_2)^\beta}\tilde{\mathcal{D}}S\mathcal{A}\frac{w(l)}{b(l)}\\=&O\left(\sum_{l=0}^{k-1}\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(l+a_2)^{1-\beta}\right)\frac{a_1\tilde{\mathcal{D}}S\mathcal{A}}{(l+a_2)^{\beta-\gamma}}\frac{w(l)}{b(l)}\right)\\=&O\left(\exp\left(\frac{a_1\lambda_2(\mathcal{L})}{1-\beta}(k+a_2)^{1-\beta}\right)(k+a_2)^{\gamma+\eta-(\eta+1/2)\beta}\right),\\&\qquad\qquad\text{a.s.}\end{aligned}\tag{27}$$

Substituting (26) and (27) into (25) gives $\tilde{\mathcal{D}}z(k)=O\left(k^{\gamma+\eta-(\eta+1/2)\beta}\right)$, a.s.

When $\beta=1$, by Lemma A.2, we have

$$\begin{aligned}\prod_{i=l+1}^{k-1}\left\|I_{N-1}-\frac{a_1}{i+a_2}\tilde{\mathcal{L}}\right\|\le&\prod_{i=l+1}^{k-1}\left(1-\frac{a_1\lambda_2(\mathcal{L})}{i+a_2}\right)\\=&O\left(\left(\frac{l+a_2}{k+a_2}\right)^{a_1\lambda_2(\mathcal{L})}\right).\end{aligned}\tag{28}$$

According to the Lemma 2 in [47], one can get

$$\begin{aligned}&\frac{1}{(k+a_2)^{a_1\lambda_2(\mathcal{L})}}\sum_{l=0}^{k-1}\frac{a_1b(l)}{(l+a_2)^{1-a_1\lambda_2(\mathcal{L})}}\tilde{\mathcal{D}}S\mathcal{A}\frac{w(l)}{b(l)}\\=&O\left(\frac{1}{(k+a_2)^{a_1\lambda_2(\mathcal{L})}}\sum_{l=0}^{k-1}\frac{a_1\tilde{\mathcal{D}}S\mathcal{A}}{(l+a_2)^{1-\gamma-a_1\lambda_2(\mathcal{L})}}\frac{w(l)}{b(l)}\right)\\=&\begin{cases}O\left(k^{\gamma-1/2}\sqrt{\ln\ln k}\right), & a_1\lambda_2(\mathcal{L})+\gamma>1/2;\\O\left(k^{\gamma-1/2}\sqrt{\ln k\ln\ln\ln k}\right), & a_1\lambda_2(\mathcal{L})+\gamma=1/2;\\O\left(k^{-a_1\lambda_2(\mathcal{L})}\right), & a_1\lambda_2(\mathcal{L})+\gamma<1/2,\end{cases}\text{ a.s.}\end{aligned}\tag{29}$$

Substituting (28) and (29) into (25) gives

$$\begin{aligned}&\tilde{\mathcal{D}}z(k)\\=&\begin{cases}O\left(k^{\gamma-1/2}\sqrt{\ln\ln k}\right), & a_1\lambda_2(\mathcal{L})+\gamma>1/2;\\O\left(k^{\gamma-1/2}\sqrt{\ln k\ln\ln\ln k}\right), & a_1\lambda_2(\mathcal{L})+\gamma=1/2;\\O\left(k^{-a_1\lambda_2(\mathcal{L})}\right), & a_1\lambda_2(\mathcal{L})+\gamma<1/2,\end{cases}\text{ a.s.}\end{aligned}$$

This completes the proof. □

Remark 3.7: Theorems 3.5 and 3.6 show that the algorithm's convergence rate will slow down when the privacy noise parameter γ increases. This is because the increase of the privacy noises enhances data randomness, and thus, worsens the convergence rate of the algorithm.

Remark 3.8: In distributed systems, communication imperfections can be modeled as communication noises [7], [9]–[11], and can be regarded as a special case of differential privacy-noises considered here. Therefore, Algorithm 1 can also be used to counteract communication imperfections in distributed computation. The proof techniques of mean-square and almost-sure convergence rates are fundamentally different from existing counterparts (e.g. [7], [9]–[11]) and are of independent interest in themselves. To the best of our knowledge, even without considering privacy protection, it is the first to rigorously characterize both the mean-square and almost-sure convergence rates of distributed consensus with increasing noises.

D. Privacy analysis

This subsection demonstrates that Algorithm 1 is ϵ-differentially private on dataset $D=\{x_i(0),i\in\mathcal{V}\}$. Before giving the privacy analysis, we first introduce the definition of sensitivity. For a private dataset D and an observation $O=\{y_i(k),i\in\mathcal{V}\}_{k=0}^T$, there exists a sequence of noises $\{\omega_i(k),i\in\mathcal{V}\}_{k=0}^T$ and trajectories $\rho(D,O)=\{x_i^{D,O}(k),i\in\mathcal{V}\}_{k=0}^T$. Below we first give the sensitivity of Algorithm 1.

Definition 3.1 (Sensitivity): The sensitivity with respect to a randomized mechanism $\mathcal{M}$ at time $k\ge0$ is given as follows.

$$S(k)=\sup_{D,D'\in\mathcal{D},O\in\mathcal{O}}\|\rho(D,O)(k)-\rho(D',O)(k)\|_1.$$

Sensitivity is a measure of the difference of two trajectories induced by changing the private dataset.

Theorem 3.7: Suppose Assumptions 2.1 and 3.2 hold. Then, the sensitivity of Algorithm 1 satisfies

$$S(k)\le\begin{cases}\delta,\ k=0;\\\prod_{l=0}^{k-1}(1-\alpha(l)c_{\min})\delta,\ k\ge1.\end{cases}\tag{30}$$

Proof: Denote $\mathcal{P}=\{\rho(D,O):O\in\mathcal{O}\}$ and $\mathcal{P}'=\{\rho(D',O):O\in\mathcal{O}\}$ as the sets of possible trajectories under the controller (3) w.r.t. D and D' in the observation set $\mathcal{O}$, and the trajectories subject to the probability density functions $f(D,\rho(D,O))$ and $f(D',\rho(D',O))$, respectively. Based on the controller (3), we have

$$\begin{aligned}x_i^{D,O}(k)=&(1-\alpha(k-1)c_i)x_i^{D,O}(k-1)\\&+\alpha(k-1)\sum_{j\in\mathcal{N}_i}|a_{ij}|\text{sgn}(a_{ij})y_j(k-1).\end{aligned}$$

Similarly, for D' we have

$$\begin{aligned} x_i^{D',O}(k) =&(1-\alpha(k-1)c_i)x_i^{D',O}(k-1) \\ &+\alpha(k-1)\sum_{j\in\mathcal{N}_i}|a_{ij}|\text{sgn}(a_{ij})y_j(k-1). \end{aligned}$$

Since observations $O=\{y_i(k-1), i\in\mathcal{V}\}$ for D and D' are the same, we have

$$\begin{aligned} &x_i^{D',O}(k)-x_i^{D,O}(k) \\ =&(1-\alpha(k-1)c_i)\left(x_i^{D',O}(k-1)-x_i^{D,O}(k-1)\right) \\ =&\prod_{l=0}^{k-1}(1-\alpha(l)c_i)\left(x_i^{D',O}(0)-x_i^{D,O}(0)\right). \end{aligned}$$

Thus, it follows that for $k=0$

$$\begin{aligned} &\|\rho(D,O)(k)-\rho(D',O)(k)\|_1 \\ =&\sum_{i\in\mathcal{V}}\left|x_i^{D',O}(0)-x_i^{D,O}(0)\right|\le\delta, \end{aligned} \tag{31}$$

which implies that $S(0)\le\delta$, and for $k\ge 1$

$$\begin{aligned} &\|\rho(D,O)(k)-\rho(D',O)(k)\|_1 \\ =&\sum_{i\in\mathcal{V}}\left|x_i^{D',O}(k)-x_i^{D,O}(k)\right| \\ =&\sum_{i\in\mathcal{V}}\left(\prod_{l=0}^{k-1}(1-\alpha(l)c_i)\right)\left|x_i^{D',O}(0)-x_i^{D,O}(0)\right| \\ \le&\left(\prod_{l=0}^{k-1}(1-\alpha(l)c_{\min})\right)\sum_{i\in\mathcal{V}}\left|x_i^{D',O}(0)-x_i^{D,O}(0)\right| \\ \le&\left(\prod_{l=0}^{k-1}(1-\alpha(l)c_{\min})\right)\delta. \end{aligned} \tag{32}$$

Thus, $S(k)\le\prod_{l=0}^{k-1}(1-\alpha(l)c_{\min})\delta$ for $k\ge 1$. This completes the proof. □

Next, we calculate the algorithm's differential privacy level ϵ.

Theorem 3.8: Suppose Assumptions 2.1 and 3.2 hold. Then, Algorithm 1 is ϵ-differentially private over the time horizon T with

$$\epsilon=\sum_{k=0}^{T}\frac{S(k)}{b(k)}. \tag{33}$$

Proof: Recall that $\mathcal{P}=\{\rho(D,O): O\in\mathcal{O}\}$ and $\mathcal{P}'=\{\rho(D',O): O\in\mathcal{O}\}$ are the sets of possible trajectories under the controller (3) w.r.t. D and D' in the observation set $\mathcal{O}$, and the trajectories subject to the probability density functions $f(D,\rho(D,O))$ and $f(D',\rho(D',O))$, respectively. Then, it is obtained that

$$\frac{\mathbb{P}[\mathcal{M}(D)\in\mathcal{O}]}{\mathbb{P}[\mathcal{M}(D')\in\mathcal{O}]}=\frac{\int_{\rho(D,O)\in\mathcal{P}}f\left(D,\rho(D,O)\right)\mathrm{d}\tau}{\int_{\rho(D',O)\in\mathcal{P}'}f\left(D',\rho(D',O)\right)\mathrm{d}\tau'}.$$

Let $\mathcal{T}=\{0,1,2,\ldots,T\}$ and $\mathcal{W}=\mathcal{V}\times\mathcal{T}$. Then, the probability density functions $f(D,\rho(D,O))$ over the time horizon T are expressed as

$$\begin{aligned} &f(D,\rho(D,O)) \\ =&\prod_{i\in\mathcal{V},\ k\in\mathcal{T}}f(D,\rho(D,O)_i(k-1)) \\ =&\prod_{(i,k)\in\mathcal{W}}\frac{1}{2b(k)}\exp\left(-\frac{|\rho(D,O)_i(k)-y_i(k)|}{b(k)}\right). \end{aligned} \tag{34}$$

As they have the same observation over the time horizon T, there exists a bijection $g(\cdot)\colon\mathcal{P}\to\mathcal{P}'$, such that for any pair of $\rho(D,O)\in\mathcal{P}$ and $\rho(D',O)\in\mathcal{P}'$, it has $g(\rho(D,O))=\rho(D',O)$. From the rationale of $y_i(k)=x_i(k)+\omega_i(k)$, $\omega_i(k)\sim\text{Lap}(0,b(k))$, and the observations $O=\{y_i(k), i\in\mathcal{V}\}_{k=0}^{T}$, by (34) we have

$$\begin{aligned} &\frac{\mathbb{P}[\mathcal{M}(D)\in\mathcal{O}]}{\mathbb{P}[\mathcal{M}(D')\in\mathcal{O}]} \\ =&\frac{\int_{\rho(D,O)\in\mathcal{P}}f\left(D,\rho(D,O)\right)\mathrm{d}\tau}{\int_{g(\rho(D,O))\in\mathcal{P}'}f\left(D',g(\rho(D,O))\right)\mathrm{d}\tau} \\ =&\frac{\int_{\rho(D,O)\in\mathcal{P}}f\left(D,\rho(D,O)\right)\mathrm{d}\tau}{\int_{\rho(D,O)\in\mathcal{P}}f\left(D',g(\rho(D,O))\right)\mathrm{d}\tau} \\ =&\prod_{(i,k)\in\mathcal{W}}\exp\left(-\frac{|\rho(D,O)_i(k)-y_i(k)|}{b(k)}\right. \\ &\left.+\frac{|\rho(D',O)_i(k)-y_i(k)|}{b(k)}\right) \\ \le&\prod_{(i,k)\in\mathcal{W}}\exp\left(\frac{\left|x_i^{D',O}(k)-x_i^{D,O}(k)\right|}{b(k)}\right), \end{aligned}$$

which together with (32) leads to

$$\begin{aligned} &\frac{\mathbb{P}[\mathcal{M}(D)\in\mathcal{O}]}{\mathbb{P}[\mathcal{M}(D')\in\mathcal{O}]} \\ =&\exp\left(\sum_{k\in\mathcal{T}}\frac{\sum_{i\in\mathcal{V}}\left|x_i^{D',O}(k)-x_i^{D,O}(k)\right|}{b(k)}\right) \\ \le&\exp\left(\sum_{k=0}^{T}\frac{S(k)}{b(k)}\right). \end{aligned}$$

Hence, we can obtain that $\epsilon=\sum_{k=0}^{T}\frac{S(k)}{b(k)}$. □

Remark 3.9: Theorem 3.8 shows that the differential privacy level ϵ is effected by the step-sizes $\alpha(k)$ and the noise parameter $b(k)$. According to (33), a larger $\alpha(k)$ implies a smaller ϵ, which further implies a stronger privacy-preserving ability. Similarly, a larger $b(k)$ implies a smaller ϵ, which further implies a stronger privacy-preserving ability.

Next, we focus on how to design the time-varying step-size $\alpha(k)$ and the noise parameter $b(k)$ to satisfy the predefined $\epsilon^\star$-differential privacy over the infinite time horizon.

Theorem 3.9: Suppose Assumptions 2.1 and 3.2 hold. Let the step-size $\alpha(k)=\frac{a_1}{(k+a_2)^\beta}$, $b(k)=\underline{b}(k+a_2)^\gamma$, $a_1,a_2,\underline{b}>0$, $a_1c_{\min}+\gamma>1$. Then, for any given $\epsilon^\star>0$, Algorithm 1 achieves the $\epsilon^\star$-differential privacy in the following four cases:

1) $\beta=1$, $\gamma\ge 0$, and

$$\frac{2\delta}{\underline{b}a_2^\gamma}+\frac{\delta a_2^{1-\gamma}}{\underline{b}(a_1c_{\min}+\gamma-1)}\le\epsilon^\star; \tag{35}$$

2) $\beta = 1$, $\gamma < 0$, and

$$\frac{2\delta}{\underline{b}(1+a_2)^{\gamma}} + \frac{\delta(1+a_2)^{-\gamma}a_2}{\underline{b}(a_1c_{\min}+\gamma-1)} \leq \epsilon^{\star}; \tag{36}$$

3) $0 < \beta < 1$, $\gamma \geq 0$, and

$$\frac{\delta}{\underline{b}a_2^{\gamma}} + \frac{\delta \exp\left(\frac{a_1c_{\min}}{1-\beta}a_2^{1-\beta}\right)}{\underline{b}(1-\beta)}\left(\frac{1-\beta}{a_1c_{\min}}\right)^{\frac{1-\gamma}{1-\beta}} \cdot \Gamma\left(\frac{1-\gamma}{1-\beta}, \frac{a_1c_{\min}a_2^{1-\beta}}{1-\beta}\right) \leq \epsilon^{\star}; \tag{37}$$

4) $0 < \beta < 1$, $\gamma < 0$, and

$$\frac{2\delta}{\underline{b}(1+a_2)^{\gamma}} + \frac{\delta \exp\left(\frac{a_1c_{\min}}{1-\beta}a_2^{1-\beta}\right)}{\underline{b}(1-\beta)}\left(\frac{1-\beta}{a_1c_{\min}}\right)^{\frac{1-\gamma}{1-\beta}} \cdot \Gamma\left(\frac{1-\gamma}{1-\beta}, \frac{a_1c_{\min}(1+a_2)^{1-\beta}}{1-\beta}\right) \leq \epsilon^{\star}. \tag{38}$$

Proof: From Theorem 3.7 and substituting $\alpha(k) = \frac{a_1}{(k+a_2)^{\beta}}$ into (30), we have

$$S(k) \leq \begin{cases} \delta, \ k = 0; \\ \prod_{l=0}^{k-1}(1 - \frac{a_1c_{\min}}{(l+a_2)^{\beta}})\delta, \ k \geq 1. \end{cases}$$

Notice that $S(0) = \delta$. Then, one can get

$$\epsilon = \sup_T \epsilon_T = \sum_{k=0}^{\infty}\frac{S(k)}{b(k)} = \frac{\delta}{\underline{b}a_2^{\gamma}} + \sum_{k=1}^{\infty}\frac{S(k)}{b(k)} \leq \epsilon^{\star}.$$

Thus, it suffices to analyze $\sum_{k=1}^{\infty}\frac{S(k)}{b(k)}$. The following analysis is undertaken according to four cases.

Case 1: $\beta = 1$ and $\gamma \geq 0$.

From (30) and Lemma A.2 it follows that

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{S(k)}{b(k)} =& \sum_{k=1}^{\infty}\frac{\prod_{l=0}^{k-1}(1-\frac{a_1c_{\min}}{(l+a_2)^{\beta}})\delta}{b(k)} \\ \leq& \sum_{k=1}^{\infty}\frac{\delta a_2^{a_1c_{\min}}}{b(k)(k-1+a_2)^{a_1c_{\min}}} \\ \leq& \frac{\delta}{\underline{b}(1+a_2)^{\gamma}} + \int_1^{\infty}\frac{\delta a_2^{a_1c_{\min}}}{\underline{b}(x-1+a_2)^{a_1c_{\min}+\gamma}}\mathrm{d}x \\ \leq& \frac{\delta}{\underline{b}(1+a_2)^{\gamma}} + \frac{\delta a_2^{1-\gamma}}{\underline{b}(a_1c_{\min}+\gamma-1)}.\end{aligned}$$

Case 2: $\beta = 1$ and $\gamma < 0$.

From (30) and Lemma A.2 it follows that

$$\sum_{k=1}^{\infty}\frac{S(k)}{b(k)} \leq \sum_{k=1}^{\infty}\frac{\delta a_2^{a_1c_{\min}}}{b(k)(k-1+a_2)^{a_1c_{\min}}}.$$

Note that $\sup_{x\in[1,\infty)}(\frac{x+a_2}{x-1+a_2})^{-\gamma} = (\frac{1+a_2}{a_2})^{-\gamma}$, for any $\gamma < 0$. Then,

$$\begin{aligned}&\sum_{k=1}^{\infty}\frac{S(k)}{b(k)} \\ \leq& \frac{\delta}{\underline{b}(1+a_2)^{\gamma}} + (\frac{1+a_2}{a_2})^{-\gamma}\int_1^{\infty}\frac{\delta a_2^{a_1c_{\min}}}{\underline{b}(x-1+a_2)^{a_1c_{\min}+\gamma}}\mathrm{d}x \\ \leq& \frac{\delta}{\underline{b}(1+a_2)^{\gamma}} + \frac{\delta(1+a_2)^{-\gamma}a_2}{\underline{b}(a_1c_{\min}+\gamma-1)}.\end{aligned}$$

Case 3: $0 < \beta < 1$ and $\gamma \geq 0$.

From Lemma A.2 it follows that

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{S(k)}{b(k)} \leq& \sum_{k=1}^{\infty}\frac{\delta}{\underline{b}}(k+a_2)^{-\gamma}\exp\left(\frac{a_1c_{\min}}{1-\beta}a_2^{1-\beta}\right. \\ &\left. - \frac{a_1c_{\min}}{1-\beta}(k+a_2)^{1-\beta}\right).\end{aligned}$$

Further, when $\gamma \geq 0$, by Lemma A.4, we have

$$\begin{aligned}&\sum_{k=1}^{\infty}\frac{S(k)}{b(k)} \\ \leq& \exp\left(\frac{a_1c_{\min}}{1-\beta}a_2^{1-\beta}\right)\int_1^{\infty}\frac{\delta}{\underline{b}}(x+a_2-1)^{-\gamma} \\ &\cdot\exp\left(-\frac{a_1c_{\min}}{1-\beta}(x+a_2-1)^{1-\beta}\right)\mathrm{d}x. \\ =& \frac{\delta\exp\left(\frac{a_1c_{\min}}{1-\beta}a_2^{1-\beta}\right)}{\underline{b}(1-\beta)}\left(\frac{1-\beta}{a_1c_{\min}}\right)^{\frac{1-\gamma}{1-\beta}} \\ &\cdot\Gamma\left(\frac{1-\gamma}{1-\beta}, \frac{a_1c_{\min}a_2^{1-\beta}}{1-\beta}\right).\end{aligned} \tag{39}$$

Case 4: $0 < \beta < 1$ and $\gamma < 0$.

From Lemma A.2 it follows that

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{S(k)}{b(k)} \leq& \sum_{k=1}^{\infty}\frac{\delta}{\underline{b}}(k+a_2)^{-\gamma}\exp\left(\frac{a_1c_{\min}}{1-\beta}a_2^{1-\beta}\right. \\ &\left. - \frac{a_1c_{\min}}{1-\beta}(k+a_2)^{1-\beta}\right).\end{aligned}$$

Set $g(x) = (x+a_2)^{-\gamma}\exp\left(-\frac{a_1c_{\min}}{1-\beta}(x+a_2)^{1-\beta}\right)$. Then, we have

$$\begin{aligned}g'(x) =& (x+a_2)^{-\gamma-1}\exp\left(-\frac{a_1c_{\min}}{1-\beta}(x+a_2)^{1-\beta}\right) \\ &\cdot\left(-\gamma - a_1c_{\min}(x+a_2)^{1-\beta}\right).\end{aligned}$$

Note that $a_1c_{\min}+\gamma > 1$. Then, $-\gamma - a_1c_{\min}(x+a_2)^{1-\beta} < -1$, for any $x \geq 1$, which implies that $g(x)$ decreases monotonically in $[1,\infty)$. Thus, by Lemma A.4, we can get

$$\begin{aligned}&\sum_{k=2}^{\infty}(k+a_2)^{-\gamma}\exp\left(-\frac{a_1c_{\min}}{1-\beta}(k+a_2)^{1-\beta}\right) \\ \leq& \int_1^{\infty}(x+a_2)^{-\gamma}\exp\left(-\frac{a_1c_{\min}}{1-\beta}(x+a_2)^{1-\beta}\right)\mathrm{d}x \\ =& \frac{1}{(1-\beta)}\left(\frac{1-\beta}{a_1c_{\min}}\right)^{\frac{1-\gamma}{1-\beta}}\Gamma\left(\frac{1-\gamma}{1-\beta}, \frac{a_1c_{\min}(1+a_2)^{1-\beta}}{1-\beta}\right).\end{aligned}$$

Thus, (38) holds. This completes the proof. □

Remark 3.10: Theorem 3.9 provides an upper bound of the differential privacy level ϵ when the step-size $\alpha(k)$ and the noise parameter $b(k)$ are designed in a certain form. From (30), (33) and (35)-(38), increasing β has the same effect as decreasing γ on both the differential privacy level ϵ and the obtained boundary. Moreover, it is known that ϵ decreases as γ increases (or $\underline{b}$, a_2 increase). Similarly, the obtained boundary decreases as γ increases (or $\underline{b}$, a_2 increase).

Remark 3.11: Since the step-size does not change $S(0)$, $\epsilon > \frac{\delta}{\underline{b}a_2^{\gamma}}$ is required regardless of the step-size. For any given $\epsilon^{\star} >$

$\frac{\delta}{\underline{b}a_2^\gamma}$, as long as a_1 or $\underline{b}$ is sufficiently large, it can always be $\epsilon \leq \epsilon^\star$.

E. Trade-off between accuracy and privacy

From Theorems 3.5, 3.6 and 3.9, we observe that the influence of the privacy noises on the convergence rate and the privacy level of Algorithm 1 is different. Specifically, when the privacy noise parameter γ increases, the convergence rate of the algorithm will slow down, but the privacy of the algorithm will be enhanced. This is because the increase of the privacy noises enhances data randomness, leading to a worse convergence rate and more robust privacy of the algorithm. In the following, we give sufficient conditions for the mean-square average bipartite consensus and differential privacy with a finite privacy level ϵ over the infinite time horizon simultaneously.

Corollary 3.2: Suppose Assumptions 2.1 and 3.2 hold. Let the step-size $\alpha(k) = \frac{a_1}{(k+a_2)^\beta}$, $b(k) = \underline{b}(k+a_2)^\gamma$, $a_1, a_2, \underline{b} > 0$. If $\beta \in (0,1]$, $\gamma < \beta - \frac{1}{2}$ and $a_1 c_{\min} + \gamma > 1$, then the mean-square average bipartite consensus and differential privacy with a finite privacy level ϵ over the infinite time horizon can be established simultaneously.

Proof: From Lemma A.2 it follows that

$$\begin{aligned}\epsilon =& \frac{\delta}{\underline{b}a_2^\gamma} + \sum_{k=1}^{\infty} \frac{\prod_{l=0}^{k-1}(1-\alpha(l)c_{\min})\delta}{b(k)} \\ =& \frac{\delta}{\underline{b}a_2^\gamma} + \sum_{k=1}^{\infty} \frac{\prod_{l=0}^{k-1}(1-\frac{a_1 c_{\min}}{(k+a_2)^\beta})\delta}{\underline{b}(k+a_2)^\gamma}. \end{aligned} \tag{40}$$

If $\beta \in (0,1)$, then by Lemma A.2, $\prod_{l=0}^{k-1}(1-\frac{a_1 c_{\min}}{(k+a_2)^\beta})$ converges to 0 exponentially, which implies that ϵ is finite. If $\beta = 1$, then by $a_1 c_{\min} + \gamma > 1$ and Lemma A.2, ϵ is also finite. This together with Theorem 3.5 proves the corollary. □

Remark 3.12: For any given $(m^\star, r^\star, \epsilon^\star)$, Theorems 3.4 and 3.9 provide a way to design the step-size $\alpha(k)$ and the noise parameter $b(k)$. From Corollary 3.2, as long as the parameters β and γ satisfy $\beta \in (0,1]$, $\gamma < \beta - \frac{1}{2}$ and $a_1 c_{\min} + \gamma > 1$, the feasible domain of the triplet ($m^\star$, $r^\star$, $\epsilon^\star$) always exists. But if both $m^\star$, $r^\star$, and $\epsilon^\star$ are required to be sufficiently small, then there is no such β and γ because there is a trade-off between the accuracy and the privacy of Algorithm 1. This is consistent with the current literature's results on differentially private algorithms.

Remark 3.13: By (40), ϵ is inversely proportional to $\underline{b}$. Note that ϵ can be arbitrarily small if $\underline{b}$ is sufficiently large. Then, any desired $\epsilon^\star$ can be obtained by adjusting $\underline{b}$.

Remark 3.14: From Corollary 3.2, even if the variances of the added noises increase, the mean-square average bipartite consensus and differential privacy with a finite privacy level ϵ over the infinite time horizon can still be established simultaneously. Hence, Algorithm 1 is effective for protecting the infinite time sequences of the state with guaranteed convergence, which is superior to the algorithms in [32], [39], [40], [42]–[44].

F. Extension to local differential privacy

In practice, each agent wants to set its own privacy level. In this scenario, the private dataset becomes $D_i = x_i(0)$ for any i. To achieve this goal, the different privacy noise parameter $b_i(k)$ can be chosen. In this subsection, we give the convergence and privacy analysis of Algorithm 1 with the different privacy noise parameter. We first give the following assumption on the step-size $\alpha(k)$ and the different noise parameter $b_i(k)$.

Assumption 3.3: The step-size $\alpha(k)$ and the different noise parameter $b_i(k)$ are positive and satisfy

$$\sup_k \alpha(k) \leq \frac{1}{\lambda_N(\mathcal{L})}, \sum_{k=0}^{\infty} \alpha(k) = \infty, \sum_{k=0}^{\infty} \alpha^2(k) b_i^2(k) < \infty.$$

Theorem 3.10: Suppose Assumptions 2.1, 3.2, and 3.3 hold. Then, Algorithm 1 achieves the mean-square and almost-sure average bipartite consensus with $\mathrm{Var}(x^\star) = \frac{2}{N^2}\sum_{k=0}^{\infty}\sum_{i\in\mathcal{V}} \alpha^2(k) c_i^2 b_i^2(k)$. Furthermore, if

$$\sum_{k=0}^{\infty}\sum_{i\in\mathcal{V}} \alpha^2(k) c_i^2 b_i^2(k) \leq \frac{m^\star (r^\star)^2 N^2}{2},$$

then the $(m^\star, r^\star)$-accuracy is ensured.

Proof: The proof is similar to that of Theorems 3.1-3.3. And thus, here we only present the main different parts as follows: We replace (8) by

$$\begin{aligned}&\mathbb{E}[V(k+1)|\mathcal{F}_k^\omega] \\ \leq& \big[1-\alpha(k)\lambda_2(\mathcal{L})\big]^2 V(k) + 2\sum_{i=1}^{N} \alpha^2(k) b_i^2(k) \|\mathcal{A}\|^2.\end{aligned}$$

By Lemma A.1, the mean-square and almost-sure average bipartite consensus are proved. Further, we replace (12) by $\mathrm{Var}(x^\star) = \frac{2}{N^2}\sum_{k=0}^{\infty}\sum_{i\in\mathcal{V}} \alpha^2(k) c_i^2 b_i^2(k)$. The proof of the $(m^\star, r^\star)$-accuracy is similar to Theorem 3.4. □

Remark 3.15: The same step-size $\alpha(k)$ is chosen for all agents to achieve the average bipartite consensus. If different step-sizes are chosen, then $\mathbb{E}[\mathbf{1}_N^T z(k)] = \mathbb{E}[\mathbf{1}_N^T z(k-1)]$ cannot hold in (10). In this case, it is difficult to ensure that $\mathbb{E}x^\star = \frac{1}{N}\sum_{i=1}^{N} s_i x_i(0)$, which is necessary for the mean-square and almost-sure average bipartite consensus.

Remark 3.16: Theorem 3.10 shows that the mean-square and almost-sure average bipartite consensus of Algorithm 1 still holds under the appropriate assumptions on the different privacy noise parameter. Different from the same privacy noise parameter $b(k)$ for all agents, the accuracy that depends on each agent has changed. Even so, each agent cannot arbitrarily choose its own $m^\star$ and $r^\star$, because $m^\star$ and $r^\star$ are global parameters. By Theorem 3.10, if the high accuracy is desired for some agents while the high privacy is desired for others, then the requirement for the high accuracy may not be met.

Next, we calculate the algorithm's local differential privacy level when each agent wants to set its own privacy level.

Theorem 3.11: Suppose Assumptions 2.1 and 3.2 hold. Then, Algorithm 1 is ϵ_i-locally differentially private over the time horizon T with $\epsilon_i = \sum_{k=0}^{T} \frac{S(k)}{b_i(k)}$.

Proof: The proof is similar to that of Theorem 3.8. And thus, here we only present the main different parts

as follows: We replace $\sum_{i\in\mathcal{V}}\left|x_i^{D',O}(0)-x_i^{D,O}(0)\right|$ by $\left|x_i^{D',O}(0)-x_i^{D,O}(0)\right|$ in (31) and (32), $f(D,\rho(D,O))=\prod_{i\in\mathcal{V},\ k\in\mathcal{T}} f(D,\rho(D,O)_i(k))$ by $f_i(D,\rho(D,O))=\prod_{k\in\mathcal{T}} f(D,\rho(D,O)_i(k))$ in (34). □

Remark 3.17: Based on Theorem 3.11, each agent can set its own privacy level ϵ_i by properly choosing the privacy noise parameter $b_i(k)$. Specifically, if a stronger privacy-preserving ability is desired for each agent, then a larger $b_i(k)$ should be chosen, which further implies the worse accuracy of the algorithm. This reveals the trade-off between accuracy and privacy.

IV. Numerical Example

This section considers discrete-time MASs of five agents coupled by the communication graph illustrated in Fig. 1. In this example, we set $\delta=0.1$ and aim to achieve bipartite consensus with the $(m^\star,r^\star)$-accuracy and $\epsilon^\star$-differential privacy, where $m^\star=0.44$, $r^\star=3$, and $\epsilon^\star=1.2$.

First, for the communication topology (a) in Fig. 1, we set the step-size as $\alpha(k)=1/(k+1)$. Compared with the decaying variance case with $b(k)=0.9^k$ in [43], [44], we employ the proposed controller (3) using the increasing variances of the privacy noises (2) with $b(k)=(k+1)^{0.1}$. The corresponding results are depicted in Fig. 2 (a) and (b), respectively, where the trajectories of $x(k)$ and $y(k)$ are displayed. The former figures in Fig. 2 (a) and (b) reveals that both algorithms of this paper and [43], [44] converge. The latter figures in Fig. 2 (a) and (b) reveals that $y(k)$ utilizing Algorithm 1 is random, while the corresponding $y(k)$ utilizing the algorithm of [43], [44] converges.

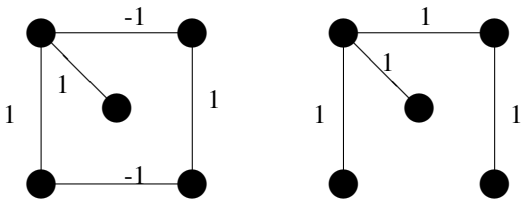
Second, for distributed consensus with unsigned graph ($S=I$), i.e., the communication topology (b) in Fig. 1, the comparison between Algorithm 1 and [32], [39] is illustrated in Fig. 3 (a) and (b), respectively. The former figure in Fig. 3 (a) and (b) also reveals that both algorithms of this paper and [32], [39] converge. The latter figure in Fig. 3 (a) and (b) also reveals that $y(k)$ utilizing Algorithm 1 is random, while the corresponding $y(k)$ utilizing the algorithm of [32], [39] converges.

Based on the above analysis, Fig. 2 and 3 highlight that Algorithm 1 has better privacy protection with guaranteed convergence compared with [32], [39], [43], [44].

Finally, we set $a_1=a_2=1$ and use the communication topology (a) in Fig. 1 with $c_{\min}=1$. The relationship of ϵ and γ, β is given in Fig. 4, which shows that the larger γ is, the smaller ϵ is. Based on Theorem 3.5, we set $\beta=0.8$. The mean-square convergence rate with different γ is given in Fig. 5, which shows that the larger γ is, the slower the algorithm's convergence rate is. This is consistent with the theoretical analysis.

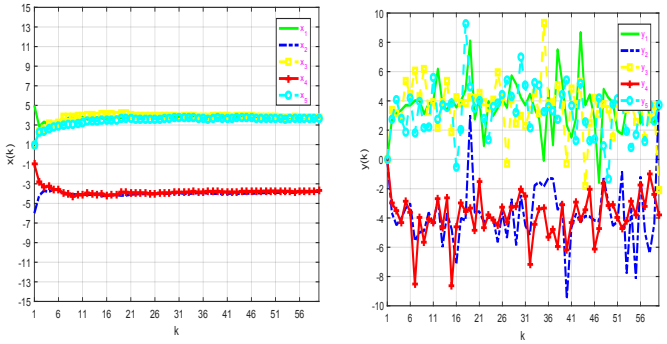
V. Conclusion

This paper develops a new differentially private bipartite consensus algorithm over signed networks. We relax the selection of privacy noises in the existing mechanisms, such that the variances of the privacy noises are time-varying and allowed to increase with time. By using the stochastic approximation method, the proposed algorithm achieves the mean-square and almost-sure average bipartite consensus, and at the same time, protects the initial value of each agent. Furthermore, we develop a method to design the time-varying step-size and the noise parameter to guarantee the desired consensus accuracy and predefined differential privacy level. We also give the mean-square and almost-sure convergence rates of the algorithm. Finally, we reveal the trade-off between the accuracy and privacy of the algorithm, and extend the results to local differential privacy. It is worth mentioning that many interesting topics deserve further investigation, including differentially private consensus-based optimization over signed networks and realizing privacy security for MASs under active adversaries.

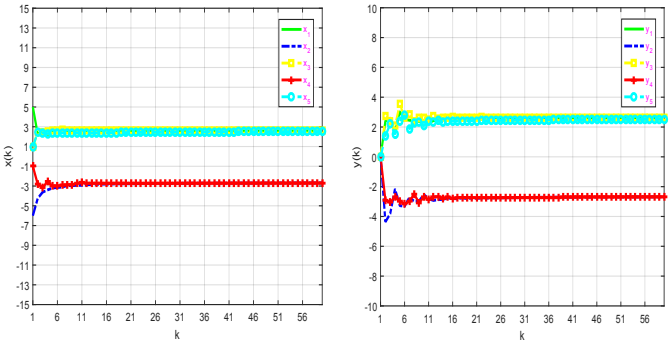


(a) Structural balance (b) Unsigned graph

Fig. 1: Communication topology

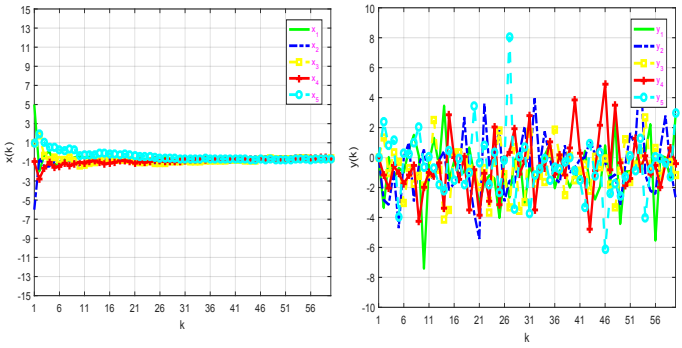


(a) Algorithm 1: $\alpha(k)=1/(k+1)$, $b(k)=(k+1)^{0.1}$

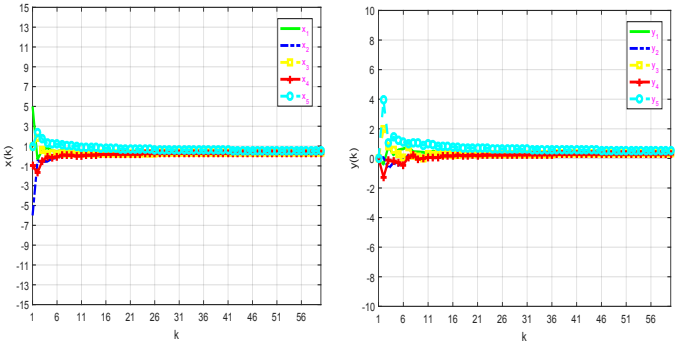


(b) [43], [44]: $b(k)=0.9^k$

Fig. 2: Comparison of Algorithm 1 with the existing differential privacy approach in [43], [44]



(a) Algorithm 1: $\alpha(k) = 1/(k+1)$, $b(k) = (k+1)^{0.1}$



(b) [32], [39]: $b(k) = 0.9^k$

Fig. 3: Comparison of Algorithm 1 with the existing differential privacy approach in [32], [39]

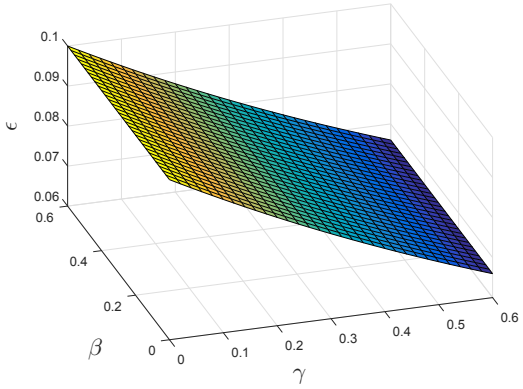


Fig. 4: The relationship of ϵ and γ, β

APPENDIX A. LEMMAS

Lemma A.1 ([48]): Let V_k, u_k, β_k, ζ_k be non-negative random variables. If $\sum_{k=0}^{\infty} u_k < \infty$,$\sum_{k=0}^{\infty} \beta_k < \infty$, and $\mathbb{E}[V_{k+1}|\mathcal{F}_k] \leq (1+u_k)V_k - \zeta_k + \beta_k$ for all $k \geq 0$, then V_k converges almost surely and $\sum_{k=0}^{\infty} \zeta_k < \infty$ almost surely. Here $\mathbb{E}[V_{k+1}|\mathcal{F}_k]$ denotes the conditional mathematical expectation for the given $V_0, \ldots, V_k$, $u_0, \ldots, u_k$, $\beta_0, \ldots, \beta_k$, $\zeta_0, \ldots, \zeta_k$.

Lemma A.2: For $0 < \beta \leq 1$, $\alpha > 0$, $k_0 \geq 0$ and

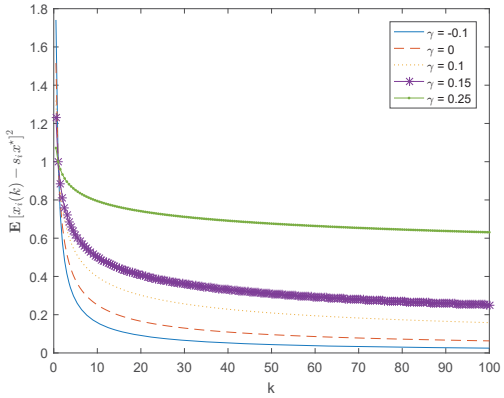


Fig. 5: The mean-square convergence rate with $\beta = 0.8$ and different γ

sufficiently large l, we have

$$\prod_{i=l}^{k}\left(1-\frac{\alpha}{(i+k_0)^{\beta}}\right) \leq \begin{cases} \left(\frac{l+k_0}{k+k_0}\right)^{\alpha}, & \beta = 1; \\ \exp\left(\frac{\alpha}{1-\beta}\left((l+k_0)^{1-\beta}-(k+k_0+1)^{1-\beta}\right)\right), & \beta \in (0,1). \end{cases} \tag{A.1}$$

If we further assume that $\beta > 1/2$, then for any $\gamma > 0$, we have

$$\prod_{i=l}^{k}\left(1-\frac{\alpha}{(i+k_0)^{\beta}}+\frac{\gamma}{(i+k_0)^{2\beta}}\right) = \begin{cases} O\left(\left(\frac{l+k_0}{k+k_0}\right)^{\alpha}\right), & \beta = 1; \\ O\left(\exp\left(\frac{\alpha}{1-\beta}\left((l+k_0)^{1-\beta}-(k+k_0+1)^{1-\beta}\right)\right)\right), & \beta \in (1/2,1). \end{cases} \tag{A.2}$$

Proof: By $\ln(1-x) \leq -x$, $\forall x \in (0,1)$, for sufficiently large l, we have

$$\prod_{i=l}^{k}\left(1-\frac{\alpha}{(i+k_0)^{\beta}}\right) = \exp\left(\sum_{i=l}^{k}\ln\left(1-\frac{\alpha}{(i+k_0)^{\beta}}\right)\right) \leq \exp\left(-\sum_{i=l}^{k}\frac{\alpha}{(i+k_0)^{\beta}}\right).$$

Note that $f(x) = \frac{\alpha}{x+k_0}$ with $\alpha > 0$ is a strictly decreasing function for $x > 0$. Then, when $\beta = 1$, we have

$$\exp\left(-\sum_{i=l}^{k}\frac{\alpha}{i+k_0}\right) \leq \exp\left(-\int_{l}^{k}\frac{\alpha}{x+k_0}\mathrm{d}x\right) = \exp\left(\alpha\ln(l+k_0)-\alpha\ln(k+k_0)\right) = \left(\frac{l+k_0}{k+k_0}\right)^{\alpha}.$$

When $\beta < 1$, from (35) in [41] it follows that

$$\exp\left(-\sum_{i=l}^{k}\frac{\alpha}{(i+k_0)^{\beta}}\right) \leq \exp\left(\frac{\alpha}{1-\beta}\left((l+k_0)^{1-\beta}-(k+k_0+1)^{1-\beta}\right)\right).$$

This completes the proof of (A.1).

Note that

$$\prod_{i=l}^{k}\left(1-\frac{\alpha}{(i+k_0)^{\beta}}+\frac{\gamma}{(i+k_0)^{2\beta}}\right)$$
$$=\prod_{i=l}^{k}\left(1-\frac{\alpha}{(i+k_0)^{\beta}}\right)\prod_{i=l}^{k}\left(1+O\left(\frac{1}{(i+k_0)^{2\beta}}\right)\right). \quad \text{(A.3)}$$

Since $\beta > 1/2$, by Theorem 2.1.3 of [49], we have $\sup_{l,k}\prod_{i=l}^{k}\left(1+O\left(\frac{1}{(i+k_0)^{2\beta}}\right)\right)<\infty$, which together with (A.1) and (A.3) implies (A.2). □

Lemma A.3: For any given $c,\ k_0\geq 0,\ 0<p\leq 1$, and $q\in\mathbb{R}$, we have $\sum_{l=1}^{k}\frac{\exp(c(l+k_0)^p)}{(l+k_0)^q}=O\left(\frac{\exp(c(k+k_0)^p)}{(k+k_0)^{p+q-1}}\right)$.

Proof: Note that

$$\sum_{l=1}^{k}(l+k_0)^{p-1}\exp\left(c(l+k_0)^p\right)$$
$$=O\left(\int_{1+k_0}^{k+k_0}t^{p-1}\exp(ct^p)\mathrm{d}t\right)$$
$$=O\left(\exp\left(c(k+k_0)^p\right)\right).$$

Then, using the Abel's transformation (see (6.29) in [50]), we have

$$\sum_{l=1}^{k}\frac{\exp\left(c(l+k_0)^p\right)}{(l+k_0)^q}$$
$$=\left(\sum_{i=1}^{k}\frac{\exp\left(c(i+k_0)^p\right)}{(i+k_0)^{1-p}}\right)\frac{1}{(k+k_0)^{p+q-1}}$$
$$+\sum_{l=1}^{k-1}\left(\sum_{i=1}^{l}\frac{\exp\left(c(i+k_0)^p\right)}{(i+k_0)^{1-p}}\right)$$
$$\cdot\left(\frac{1}{(l+k_0)^{p+q-1}}-\frac{1}{(l+k_0+1)^{p+q-1}}\right)$$
$$=O\left(\frac{\exp\left(c(k+k_0)^p\right)}{(k+k_0)^{p+q-1}}\right)+O\left(\sum_{l=1}^{k}\frac{\exp\left(c(l+k_0)^p\right)}{(l+k_0)^{p+q}}\right),$$

which together with

$$O\left(\sum_{l=1}^{k}\frac{\exp\left(c(l+k_0)^p\right)}{(l+k_0)^{p+q}}\right)=o\left(\sum_{l=1}^{k}\frac{\exp\left(c(l+k_0)^p\right)}{(l+k_0)^{q}}\right)$$

implies the lemma. □

Lemma A.4: For $\gamma<1,\ 0<\beta<1,\ \nu>0$, we have

$$\int_1^{\infty}x^{-\gamma}\exp\left(-\nu x^{1-\beta}\right)\mathrm{d}x=\frac{\nu^{-\frac{1-\gamma}{1-\beta}}}{1-\beta}\Gamma\left(\frac{1-\gamma}{1-\beta},\nu\right),$$

where $\Gamma(\cdot,\cdot)$ is the upper incomplete gamma function.

Proof: Denote $t=\nu x^{1-\beta}$. Then, we have $\mathrm{d}t=\nu(1-\beta)x^{-\beta}\mathrm{d}x$, and

$$\int_1^{\infty}x^{-\gamma}\exp\left(-\nu x^{1-\beta}\right)\mathrm{d}x$$
$$=\int_{\nu}^{\infty}\frac{1}{\nu(1-\beta)}\left(\frac{t}{\nu}\right)^{\frac{1-\gamma}{1-\beta}-1}e^{-t}\mathrm{d}t$$
$$=\frac{\nu^{-\frac{1-\gamma}{1-\beta}}}{1-\beta}\Gamma\left(\frac{1-\gamma}{1-\beta},\nu\right).$$

This proves the lemma. □

REFERENCES

[1] Z. Zhang and M. Y. Chow, "Convergence analysis of the incremental cost consensus algorithm under different communication network topologies in a smart grid," *IEEE Trans. Power Systems,* vol. 27, no. 4, pp. 1761-1768, 2012.

[2] W. Chen and T. Li, "Distributed economic dispatch for energy internet based on multiagent consensus control," *IEEE Trans. Automatic Control,* vol. 66, no. 1, pp. 137-152, 2021.

[3] S. Kang, J. Wang, G. Li, J. Shan, and I. R. Petersen, "Optimal cooperative guidance law for salvo attack: An MPC-based consensus perspective," *IEEE Trans. Aerospace and Electronic Systems,* vol. 54, no. 5, pp. 2397-2410, 2018.

[4] L. F. Wang, Y. G. Hong, G. D. Shi, and C. Altafini, "Signed social networks with biased assimilation," *IEEE Trans. Automatic Control,* vol. 67, no. 10, pp. 5134-5149, 2022.

[5] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control,* vol. 49, no. 9, pp. 1520-1533, 2004.

[6] T. Li and J. F. Zhang, "Mean square average-consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, pp. 1929-1936, 2009.

[7] T. Li and J. F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Trans. Automatic Control,* vol. 55, no. 9, pp. 2043-2057, 2010.

[8] A. Y. Kibangou, "Finite-time average consensus based protocol for distributed estimation over AWGN channels," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 5595-5600, 2011.

[9] M. Y. Huang, S. Dey, G. N. Nair, and J. H. Manton, "Stochastic consensus over noisy networks with Markovian and arbitrary switches," *Automatica,* vol. 46, pp. 1571-1583, 2010.

[10] M. Y. Huang, "Stochastic approximation for consensus: a new approach via ergodic backward products," *IEEE Trans. Automatic Control,* vol. 57, no. 12, pp. 2994-3008, 2012.

[11] M. Y. Huang, T. Li, and J. F. Zhang, "Stochastic approximation based consensus dynamics over Markovian networks," *SIAM J. Control Optim.,* vol. 53, no. 6, pp. 3339-3363, 2015.

[12] B. M. Nejad, S. A. Attia, and J. Raisch, "Max-consensus in a max-plus algebraic setting: The case of fixed communication topologies," in *Proceedings of the 22nd International Symposium on Information, Communication and Automation Technologies*, pp. 1-7, 2009.

[13] G. Chen, X. Duan, W. Mei, and F. Bullo, "Linear stochastic approximation algorithms and group consensus over random signed networks," *IEEE Trans. Automatic Control,* vol. 64, no. 5, pp. 1874-1889, 2019.

[14] C. Li and X. F. Zong, "Group consensus of multi-agent systems with additive noises," *Sci. China-Inf. Sci.*, vol. 65, pp. 202205:1-202205:14, 2022.

[15] C. Altafini, "Consensus problems on networks with antagonistic interactions," *IEEE Trans. Automatic Control,* vol. 58, no. 4, pp. 935-946, 2013.

[16] J. Liu, X. Chen, T. Başar, and M. Ali Belabbas, "Exponential convergence of the discrete- and continuous-time Altafini models," *IEEE Trans. Automatic Control,* vol. 62, no. 12, pp. 6168-6182, 2017.

[17] J. Hu, Y. Wu, T. Li, and B. K. Ghosh, "Consensus control of general linear multiagent systems with antagonistic interactions and communication noises," *IEEE Trans. Automatic Control,* vol. 64, no. 5, pp. 2122-2127, 2019.

[18] Y. Chen, Z. Zuo, and Y. Wang, "Bipartite consensus for a network of wave pdes over a signed directed graph," *Automatica,* vol. 129, 109640, 2021.

[19] A. Fontan, L. F. Wang, Y. G. Hong, G. D. Shi, and C. Altafini, "Multiagent consensus over time-invariant and time-varying signed digraphs via eventual positivity," *IEEE Trans. Automatic Control,* vol. 68, no. 9, pp. 5429-5444, 2023.

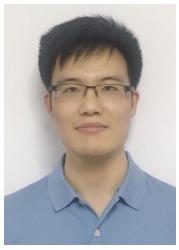
[20] G. Shi, A. Proutiere, M. Johansson, J. S. Baras, and K. H. Johansson, "The evolution of beliefs over signed social networks," *Operations Research,* vol. 64, no. 3, pp. 585-604, 2016.

[21] G. Shi, C. Altafini, and J. S. Baras, "Dynamics over signed networks," *SIAM Review,* vol. 61, no. 2, pp. 229-257, 2019.

[22] V. Amelkin, F. Bullo, and A. K. Singh, "Polar opinion dynamics in social networks," *IEEE Trans. Automatic Control,* vol. 62, no. 11, pp. 5650-5665, 2017.

[23] M. Ruan, H. Gao, and Y. Q. Wang, "Secure and privacy-preserving consensus," *IEEE Trans. Automatic Control,* vol. 64, no. 10, pp. 4035-4049, 2019.

[24] Y. Lu and M. H. Zhu, "Privacy preserving distributed optimization using homomorphic encryption," *Automatica,* vol. 96, pp. 314-325, 2018.
[25] Y. Mo and R. M. Murray, "Privacy preserving average consensus," *IEEE Trans. Automatic Control,* vol. 62, no. 2, pp. 753-765, 2017.
[26] J. He, L. Cai, P. Cheng, J. Pan, and L. Shi, "Consensus-based data-privacy preserving data aggregation," *IEEE Trans. Automatic Control,* vol. 6, no. 12, pp. 5222-5229, 2019.
[27] J. He, L. Cai, C. Zhao, P. Cheng, and X. Guan, "Privacy-preserving average consensus: Privacy analysis and algorithm design," *IEEE Trans. Signal and Information Processing over Networks,* vol. 5, no. 1, pp. 127-138, 2019.
[28] C. Altafini, "A system-theoretic framework for privacy preservation in continuous-time multiagent dynamics," *Automatica,* vol. 122, 109253, 2020.
[29] Y. Q. Wang and H. Vincent Poor, "Decentralized stochastic optimization with inherent privacy protection," *IEEE Trans. Automatic Control*, vol. 68, no. 4, pp. 2293-2308, 2023.
[30] Y. Q. Wang, "Privacy-preserving average consensus via state decomposition," *IEEE Trans. Automatic Control*, vol. 64, no. 11, pp. 4711-4716, 2019.
[31] C. Dwork, "Differential privacy," in *Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (Eds.), Automata, Languages and Programming. The Organization, Springer Berlin Heidelberg*, pp. 1-12, 2006.
[32] Z. Huang, S. Mitra, and G. Dullerud, "Differentially private iterative synchronous consensus," in *Proceedings of the ACM Workshop on Privacy in the Electronic Society*, pp. 81-90, 2012.
[33] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Trans. Automatic Control,* vol. 62, pp. 50-64, 2017.
[34] Y. Q. Wang and A. Nedic. "Tailoring gradient methods for differentially-private distributed optimization," *IEEE Trans. Automatic Control,* DOI: 10.1109/TAC.2023.3272968, 2023.
[35] J. M. Wang, J. F. Zhang, and X. K. He, "Differentially private distributed algorithms for stochastic aggregative games," *Automatica*, vol. 142, 110440, 2022.
[36] J. Le Ny and G. J. Pappas, "Differentially private filtering," *IEEE Trans. Automatic Control,* vol. 59, no. 2, pp. 341-354, 2014.
[37] J. Le Ny and M. Mohammady, "Differentially private MIMO filtering for event streams," *IEEE Trans. Automatic Control,* vol. 63, no. 1, pp. 145-157, 2018.
[38] J. F. Zhang, J. W. Tan, and J. M. Wang, "Privacy security in control systems," *Sci. China-Inf. Sci.*, vol. 64, pp. 176201:1-176201:3, 2021.
[39] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design," *Automatica,* vol. 81, pp. 221-231, 2017.
[40] D. Fiore and G. Russo, "Resilient consensus for multi-agent systems subject to differential privacy requirements," *Automatica,* vol. 106, pp. 18-26, 2019.
[41] X. K. Liu, J. F. Zhang, and J. M. Wang, "Differentially private consensus algorithm for continuous-time heterogeneous multi-agent systems," *Automatica*, vol. 122, 109283, 2020.
[42] L. Gao, S. Deng, and W. Ren, "Differentially private consensus with an event-triggered mechanism," *IEEE Trans. Control of Network Systems,* vol. 6, no. 1, pp. 60-71, 2019.
[43] Z. Q. Zuo, R. Tian, and Y. J. Wang, "Bipartite consensus for multi-agent systems with differential privacy constraint," in *Proceedings of the 40th Chinese Control Conference*, pp. 5062-5067, 2021.
[44] Z. Q. Zuo, R. Tian, Q. N. Han, Y. J. Wang, and W. T. Zhang, "Differential privacy for bipartite consensus over signed digraph," *Neurocomputing*, vol. 468, pp. 11-21, 2022.
[45] A. S. Leong, D. E. Quevedo, D. Dolz, and S. Dey, "Transmission scheduling for remote state estimation over packet dropping links in the presence of an eavesdropper," *IEEE Trans. Automatic Control*, vol. 64, no. 9, pp. 3732-3739, 2019.
[46] R. B. Ash, *Real Analysis and Probability.* New York: Academic Press, 1972.
[47] C. Z. Wei, "Asymptotic properties of least-squares estimates in stochastic regression models," *The Annals of Statistics,* vol. 13, no. 4, pp. 1498-1508, 1985.
[48] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control.* Englewood Cliffs, NJ: Prentice-Hall, 1984.
[49] C. D. Pan and X. Y. Yu, *Foundation of Order Estimation.* Beijing, China: Higher Education Press, 2015.
[50] V. A. Zorich, *Mathematical Analysis I (2 ed.)* Berlin, German: Springer-Verlag, 2016.



Jimin Wang received the B.S. degree in mathematics from Shandong Normal University, China, in 2012 and the Ph.D. degree from the School of Mathematics, Shandong University, China, in 2018. From May 2017 to May 2018, he was a joint Ph.D. student with the School of Electrical Engineering and Computing, The University of Newcastle, Australia. From July 2018 to December 2020, he was a postdoctoral researcher in the Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), China. He is currently an associate professor in the School of Automation and Electrical Engineering, University of Science and Technology Beijing. His current research interests include privacy and security in cyber-physical systems, stochastic systems and networked control systems. He was a recipient of Shandong University's excellent doctoral dissertation.

Jieming Ke received the B.S. degree in Mathematics from University of Chinese Academy of Science, Beijing, China, in 2020. He is currently working toward the Ph.D. degree majoring in system theory at Academy of Mathematics and Systems Science, Chinese Academy of Science, Beijing, China. His research interests include privacy and security in stochastic systems, and identification and control of quantized systems.

Ji-Feng Zhang received the B.S. degree in mathematics from Shandong University, China, in 1985 and the Ph.D. degree from the Institute of Systems Science (ISS), Chinese Academy of Sciences (CAS), China, in 1991. He is now with the ISS, Academy of Mathematics and Systems Science, CAS. His current research interests include system modeling, adaptive control, stochastic systems, and multi-agent systems.

He is an IEEE Fellow, IFAC Fellow, CAA Fellow, CSIAM Fellow, member of the European Academy of Sciences and Arts, and Academician of the International Academy for Systems and Cybernetic Sciences. He received the second prize of the State Natural Science Award of China in 2010 and 2015, respectively. He was a Vice-Chair of the IFAC Technical Board, member of the Board of Governors, IEEE Control Systems Society; Convenor of Systems Science Discipline, Academic Degree Committee of the State Council of China; Vice-President of the Systems Engineering Society of China, the Chinese Mathematical Society, and the Chinese Association of Automation. He has served as Editor-in-Chief, Deputy Editor-in-Chief, Senior Editor or Associate Editor for more than 10 journals, including *Science China Information Sciences*, *National Science Review*, *IEEE Transactions on Automatic Control*, and *SIAM Journal on Control and Optimization* etc.