Differentially Private Average Consensus with Optimal Noise Selection

Erfan Nozari [*] **Pavankumar Tallapragada** [*] **Jorge Cortés** [*]

[*] *Department of Mechanical and Aerospace Engineering, University of California, San Diego, {enozari,ptallapragada,cortes}@ucsd.edu*

Abstract: This paper studies the problem of privacy-preserving average consensus in multi-agent systems. The network objective is to compute the average of the initial agent states while keeping these values differentially private against an adversary that has access to all inter-agent messages. We establish an impossibility result that shows that exact average consensus cannot be achieved by any algorithm that preserves differential privacy. This result motives our design of a differentially private discrete-time distributed algorithm that corrupts messages with Laplacian noise and is guaranteed to achieve average consensus in expectation. We examine how to optimally select the noise parameters in order to minimize the variance of the network convergence point for a desired level of privacy.

Keywords: average consensus, differential privacy, multi-agent systems

1. INTRODUCTION

Multi-agent average consensus is a basic distributed control problem where a group of agents seek to agree on the average of their individual values by only interchanging information with their neighbors. This problem has found numerous applications in sensor networks, synchronization, network management, and distributed computation and optimization. In many of these applications, guaranteeing the privacy of the individual agents is an important aspect that has not been sufficiently studied in the context of networked systems and cooperative strategies. An increasing number of works look at the notion of differential privacy, which specifies that the information of an agent has no significant effect in the aggregate output of the algorithm, and hence its data cannot be inferred by an adversary from its execution. This is a strong notion of privacy with a rigorous formulation and proven security properties, including resilience to post-processing and auxiliary information and independence from the model of the adversary. This paper is a contribution to this body of research where we focus our attention on gaining insight into the achievable trade-offs between privacy and performance in multi-agent average consensus.

Literature Review There is a large literature on the (average) consensus problem in networked systems and the interested reader is referred to (Bullo et al., 2009; Ren and Beard, 2008; Mesbahi and Egerstedt, 2010) and references therein for a comprehensive review. The notion of differential privacy, first introduced in (Dwork et al., 2006; Dwork, 2006), has been the subject of extensive research in the database literature over the past decade. A recent comprehensive text can be found in (Dwork and Roth, 2014). Recently, this notion has found its way into a number of areas pertaining networked systems including control (Huang et al., 2012, 2014; Wang et al., 2014), estimation (Ny and Pappas, 2014), and optimization (Han et al., 2014; Huang et al., 2015). Of particular relevance to our paper is the work of Huang et al. (2012), which considers the multi-agent average consensus problem and proposes an adjacency-based distributed algorithm with decaying Laplacian noise in the inter-agent messages. The algorithm is differentially private and agents asymptotically agree on a value that may not be the average of their initial states, even in expectation. Our present work improves upon (Huang et al., 2012) by providing a performance bound that sheds light on what can be achieved in terms of differential privacy for general average consensus dynamics and studying a stronger notion of convergence. Our results also allow individual agents to independently choose their level of privacy. Other works have looked at the average consensus problem employing different notions of privacy. Manitara and Hadjicostis (2013) improve upon (Kefayati et al., 2007) to propose a distributed algorithm where any agent has the option to add a zero-sum noise sequence with finite random length to its first set of transmitted messages. Since the sequence is zero-sum, agents converge to the true average. Privacy of a participating agent, understood as the property that different initial conditions produce the same transmitted messages, is preserved if the malicious nodes cannot listen to it and all its neighbors. The work of Mo and Murray (2014) adds infinite-length exponentially-decaying zero-sum noise sequences to inter-agent messages and formally defines privacy as the inability of a malicious node to perfectly recover the initial state of other nodes via maximum-likelihood estimation. The proposed algorithm is mean-square convergent to the true average and preserves the privacy of nodes whose messages and those of their neighbors are not listened to by the malicious nodes.

Statement of Contributions We consider the multi-agent average consensus problem with privacy preservation requirements on the initial agent states. Our main contributions pertain the understanding of the trade-offs between differential privacy and performance, and can be divided into three groups as follows. Our first contribution is a general result stating that any distributed coordination algorithm cannot simultaneously be differentially private and guarantee weak convergence of agents to the average of their initial states. Our second contribution is the design of a distributed algorithm that guarantees that the agents converge in expectation to the average of their initial states. Our design uses the classical discrete-time Laplacian-based linear stationary dynamics together with additive Laplacian noise processes. We establish the almost sure convergence, unbiasedness, bounded dispersion, and differential privacy of our design in successive results. Our final contribution pertains to the optimal tuning of the design parameters of the algorithm (specifically, the noise-to-state gain of the system and the amplitude and decay rate of the noise) to minimize the variance of the network convergence point for a desired level of privacy. Various simulations illustrate our results. Most of the proofs are omitted for space reasons and will appear elsewhere.

2. PRELIMINARIES

This section introduces notation and basic concepts. We denote the set of reals, positive reals, non-negative reals, positive integers, and nonnegative integers by $\mathbb{R}$, $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, $\mathbb{N}$, and $\mathbb{Z}_{\geq 0}$, respectively. We let $(\mathbb{R}^n)^{\mathbb{N}}$ denote the space of vector-valued sequences in the Euclidean space $\mathbb{R}^n$. Given n numbers $c_1, \dots, c_n \in \mathbb{R}$, $\operatorname{diag}(c_1, \dots, c_n) \in \mathbb{R}^{n\times n}$ denotes a diagonal matrix with $c_1, \dots, c_n$ on its diagonal. For any $\{x(k)\}_{k=0}^{\infty} \in (\mathbb{R}^n)^{\mathbb{N}}$, we define the shorthand notations $\mathbf{x} = \{x(k)\}_{k=0}^{\infty}$ and $\mathbf{x}_k = \{x(j)\}_{j=0}^{k}$. $I_n \in \mathbb{R}^{n\times n}$ and $\mathbf{1}_n \in \mathbb{R}^n$ denote the identity matrix and the vector of ones, respectively. For $x \in \mathbb{R}^n$, $\operatorname{Ave}(x) = \frac{1}{n}\mathbf{1}_n^T x$ denotes the average of its components. We let $\Pi_n = \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$. Note that Π_n is diagonalizable, has one eigenvalue equal to 1 associated with eigenspace

$$\mathcal{D}_n = \{x \in \mathbb{R}^n \mid x_i = \operatorname{Ave}(x),\, i \in \{1, \dots, n\}\},$$

while all other eigenvalues equal 0. For a vector space $V \subset \mathbb{R}^n$, we let $V^{\perp}$ denote the vector space orthogonal to V. We denote the Euclidean norm in $\mathbb{R}^n$ by $\|\cdot\|$. We say a matrix $A \in \mathbb{R}^{n\times n}$ is stable if all its eigenvalues have magnitude strictly less than 1. For $q \in (0, 1)$, the Euler function is given by

$$\varphi(q) = \prod_{k=1}^{\infty}(1 - q^k) > 0.$$

Note that

$$\lim_{k\to\infty} \prod_{j=k}^{\infty}(1 - q^j) = \lim_{k\to\infty} \frac{\varphi(q)}{\prod_{j=1}^{k-1}(1 - q^j)} = 1.$$

2.1 Graph Theory

We present some useful notions on algebraic graph theory following (Bullo et al., 2009). Let $\mathcal{G} = (V, E, A)$ denote a weighted undirected graph with vertex set V of cardinality n, edge set $E \subset V \times V$ and symmetric adjacency matrix $A \in \mathbb{R}_{\geq 0}^{n\times n}$. A path from i to j is a sequence of vertices starting from i and ending in j such that any pair of consecutive vertices is an edge of the graph. The set of neighbors of i, denoted $\mathcal{N}_i$, is the set of nodes j such that $(i, j) \in E$. The graph $\mathcal{G}$ is connected if for each node there exists a path to any other node. The weighted degree matrix of $\mathcal{G}$ is a diagonal matrix $D \in \mathbb{R}^{n\times n}$ whose ith diagonal element, $i \in \{1, \dots, n\}$, is the sum of the ith row of A. The Laplacian of $\mathcal{G}$ is the symmetric matrix

$$L = D - A,$$

and has the following properties:

- L is positive semi-definite;
- $L\mathbf{1}_n = 0$ and $\mathbf{1}_n^T L = 0$, i.e., 0 is an eigenvalue of L corresponding to the eigenspace $\mathcal{D}_n$;
- $\mathcal{G}$ is connected if and only if $\operatorname{rank}(L) = n - 1$, so 0 is a simple eigenvalue of L;
- All eigenvalues of L belong to $[0, 2d_{\max}]$, where $d_{\max}$ is the largest element of D.

For convenience, we define $L_{\text{cpt}} = I_n - \Pi_n$.

2.2 Probability Theory

Here we briefly review basic notions on probability following (Papoulis and Pillai, 2002; Durrett, 2010). Consider a probability space $(\Omega, \Sigma, \mathbb{P})$. If $E, F \in \Sigma$ are two events with $E \subseteq F$, then $\mathbb{P}\{E\} \leq \mathbb{P}\{F\}$. For simplicity, we may sometimes denote events of the type $E_p = \{\omega \in \Omega \mid p(\omega)\}$ by $\{p\}$ where p is a logical statement on the elements of Ω. Clearly, for two statements p and q,

$$(p \Rightarrow q) \Rightarrow (\mathbb{P}\{p\} \leq \mathbb{P}\{q\}). \tag{1}$$

A random variable is a function $X : \Omega \to \mathbb{R}$ such that the inverse image of any open set $B \subseteq \mathbb{R}$ belongs to Σ. For any $N \in \mathbb{R}_{>0}$ and any random variable X with finite expected value μ and finite nonzero variance σ^2, Chebyshev's inequality states that

$$\mathbb{P}\{|X - \mu| \geq N\sigma\} \leq \frac{1}{N^2}.$$

Let for a random variable X, $\mathbb{E}[X]$ and F_X denote its expectation and cumulative distribution function, respectively. Then, a sequence of random variables $\{X_k\}_{k\in\mathbb{Z}_{\geq 0}}$ converges to a random variable X

- almost surely (a.s.) or with probability one if
$$\mathbb{P}\{\lim_{k\to\infty} X_k = X\} = 1;$$
- in mean square (m.s.) if $\mathbb{E}[X_k^2], \mathbb{E}[X^2] < \infty$ for all $k \in \mathbb{Z}_{\geq 0}$ and
$$\lim_{k\to\infty} \mathbb{E}[(X_k - X)^2] = 0;$$
- in probability, if for any $\varepsilon > 0$,
$$\lim_{k\to\infty} \mathbb{P}\{|X_k - X| < \varepsilon\} = 1;$$
- in distribution, or weakly, if for any $x \in \mathbb{R}$ at which F_X is continuous,
$$\lim_{k\to\infty} F_{X_k}(x) = F_X(x).$$

Almost sure convergence and convergence in mean square imply convergence in probability, which itself implies convergence in distribution. Moreover, if $\mathbb{P}\{|X_k| \leq \bar{X}\} = 1$ for all $k \in \mathbb{Z}_{\geq 0}$ and some fixed random variable $\bar{X}$ with $\mathbb{E}[\bar{X}^2] < \infty$, then convergence in probability implies mean square convergence, and if X is a constant, then convergence in distribution implies convergence in probability.

A zero-mean random variable X has Laplace distribution with scale $b \in \mathbb{R}_{>0}$, denoted $X \sim \operatorname{Lap}(b)$, if the pdf of X is

$$f_X(x) = \mathcal{L}(x; b) \triangleq \frac{1}{2b} e^{-\frac{|x|}{b}},$$

for any $x \in \mathbb{R}$. It is easy to see that $|X|$ has an exponential distribution with rate $\lambda = \frac{1}{b}$.

2.3 Input-to-State Stability of Discrete-Time Systems

The material of this section are borrowed from (Jiang and Wang, 2001), which the interested reader may consult for details. Consider a discrete-time system of the form

$$x(k+1) = f(x(k), u(k)), \qquad (2)$$

where $u : \mathbb{Z}_{\geq 0} \to \mathbb{R}^m$ is a disturbance input, $x : \mathbb{Z}_{\geq 0} \to \mathbb{R}^n$ is the state, and $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is a vector field satisfying $f(0,0) = 0$. The system (2) is globally input-to-state stable (ISS) if there exists a class $\mathcal{KL}$ function β and a class $\mathcal{K}$ function γ such that, for any bounded input u, any initial condition $x_0 \in \mathbb{R}^n$, and all $k \in \mathbb{Z}_{\geq 0}$,

$$\|x(k)\| \leq \beta(\|x_0\|, k) + \gamma(\|u(.)\|_{L_\infty}),$$

where $\|u(.)\|_{L_\infty} = \sup\{\|u(k)\| \mid k \in \mathbb{Z}_{\geq 0}\}$. Moreover, the system (2) has a $\mathcal{K}$-asymptotic gain if there exists a class $\mathcal{K}$ function γ_a such that, for any initial condition $x_0 \in \mathbb{R}^n$,

$$\limsup_{k\to\infty} \|x(k)\| \leq \gamma_a \left(\limsup_{k\to\infty} \|u(k)\| \right).$$

If a system is ISS, then it has a $\mathcal{K}$-asymptotic gain. Furthermore, any LTI system $x(k+1) = Ax(k) + Bu(k)$ is ISS if A is stable.

3. PROBLEM STATEMENT

Consider a group of n agents whose interaction topology is described by an undirected connected graph $\mathcal{G}$. The group objective is to compute the average of the agents' initial states while preserving the privacy of these values against potential adversaries eavesdropping the network communications.

We follow the exposition in (Huang et al., 2012) to formally present the problem statement. The state of each agent $i \in \{1, \ldots, n\}$ is represented by $\theta_i \in \mathbb{R}$. The message that agent i shares with its neighbors about its current state is denoted by $x_i \in \mathbb{R}$. For convenience, the aggregated network state and the vector of transmitted messages are denoted by $\theta = (\theta_1, \ldots, \theta_n) \in \mathbb{R}^n$ and $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, respectively. Agents update their states in discrete time according to

$$\theta(k+1) = f(\theta(k), x(k)), \quad k \in \mathbb{Z}_{\geq 0}, \qquad (3)$$

where the state-transition function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is such that its ith element depends only on θ_i and $\{x_j\}_{j\in\mathcal{N}_i\cup\{i\}}$. The messages are calculated from

$$x(k) = h(\theta(k), \eta(k)), \quad k \in \mathbb{Z}_{\geq 0}, \qquad (4)$$

where $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ is such that its ith element depends only on θ_i and η_i. $\eta(k) \in \mathbb{R}^n$ is a vector random variable with $\eta_i(k)$ being the noise generated by agent i at time k from an arbitrary distribution. Then, the sequences $\boldsymbol{\theta}$ and $\mathbf{x}$ are, in general, random variables on the total sample space

$$\Omega = (\mathbb{R}^n)^{\mathbb{N}},$$

whose elements are noise sequences $\boldsymbol{\eta}$. Although one can choose h to be a function only of the argument θ, corrupting the messages by noise is necessary to preserve privacy.

In order to formulate the privacy requirements, we first introduce some definitions. For $\delta \in \mathbb{R}_{>0}$, a pair of initial network states $\theta_0^{(1)}$ and $\theta_0^{(2)}$ are called δ-adjacent, with δ called the "adjacency bound", if, for some $i_0 \in \{1, \ldots, n\}$,

$$\forall i \in \{1, \ldots, n\} \qquad |\theta_{0,i}^{(2)} - \theta_{0,i}^{(1)}| \leq \begin{cases} \delta & \text{if } i = i_0, \\ 0 & \text{if } i \neq i_0. \end{cases} \qquad (5)$$

Moreover, from (3) and (4), it is clear that for any fixed initial state θ_0, $\mathbf{x}$ is uniquely determined by $\boldsymbol{\eta}$. Therefore, the function $X_{\theta_0} : (\mathbb{R}^n)^{\mathbb{N}} \to (\mathbb{R}^n)^{\mathbb{N}}$ such that

$$X_{\theta_0}(\boldsymbol{\eta}) = \mathbf{x}$$

is well defined. Privacy is defined as follows.

Definition 3.1. (Differential Privacy). Given $\delta, \epsilon \in \mathbb{R}_{\geq 0}$, the dynamics (3), (4) is ϵ-differentially private if, for any pair $\theta_0^{(1)}$ and $\theta_0^{(2)}$ of δ-adjacent initial states and any set $\mathcal{O} \subset (\mathbb{R}^n)^{\mathbb{N}}$, one has

$$\mathbb{P}\{\boldsymbol{\eta} \in \Omega \,|\, X_{\theta_0^{(1)}}(\boldsymbol{\eta}) \in \mathcal{O}\} \leq e^{\epsilon}\mathbb{P}\{\boldsymbol{\eta} \in \Omega \mid X_{\theta_0^{(2)}}(\boldsymbol{\eta}) \in \mathcal{O}\}. \quad \bullet$$

Because of the presence of noise, the agents' states might converge under (3) to a neighborhood of $\mathrm{Ave}(\theta_0)$ instead of to $\mathrm{Ave}(\theta_0)$ itself, as captured by the notion of accuracy.

Definition 3.2. (Accuracy). For any $p \in [0, 1]$ and $r \in \mathbb{R}_{\geq 0}$, the dynamics (3), (4) is (p, r)-accurate if, starting from θ_0, the network state $\theta(k)$ converges to $\theta_\infty \in \mathbb{R}^n$ as $k \to \infty$,

$$\mathbb{E}[\theta_\infty] = \mathrm{Ave}(\theta_0)\mathbf{1}_n,$$

and

$$\mathbb{P}\{\|\theta_\infty - \mathrm{Ave}(\theta_0)\mathbf{1}_n\| \leq r\} \geq 1 - p. \quad \bullet$$

Note that, in Definition 3.2, the type of convergence of $\theta(k)$ to θ_∞ can be any of the four classes introduced in Section 2.2. Furthermore, for each notion of convergence, $(0, 0)$-accuracy is equivalent to the convergence of $\theta(k)$ to $\mathrm{Ave}(\theta_0)\mathbf{1}_n$.

We are now ready to formally state our problem as follows.

Problem 1. (Differentially Private Average Consensus): Design the dynamics (3), the inter-agent messages (4), and the distribution of noise sequences $\boldsymbol{\eta}$ such that asymptotic average consensus is guaranteed and ϵ-differential privacy and (p, r)-accuracy are achieved for (finite) ϵ, r, and $p \in \mathbb{R}_{\geq 0}$ as small as possible. $\bullet$

4. PERFORMANCE BOUND ON DIFFERENTIALLY PRIVATE AVERAGE CONSENSUS

In this section we establish the impossibility of solving Problem 1 with $(0, 0)$-accuracy, even if considering the weakest notion of convergence.

Theorem 4.1. (Impossibility Result). Consider a group of agents executing the distributed algorithm (3) with messages generated according to (4). Then, for any $\delta, \epsilon > 0$, agents cannot simultaneously converge to the average of their initial states in distribution and preserve ϵ-differential privacy of their initial states.

Proof. We reason by contradiction. Assume one such algorithm exists, achieving convergence in distribution to $\mathrm{Ave}(\theta_0)$, where θ_0 is the vector of agents' initial states, and preserving ϵ-differential privacy. Since the algorithm must preserve the privacy of *any* pair of δ-adjacent initial conditions, consider a specific pair satisfying

$$\theta_{0,i_0}^{(2)} = \theta_{0,i_0}^{(1)} + \delta,$$

for some $i_0 \in \{1, \ldots, n\}$ and $\theta_{0,i}^{(2)} = \theta_{0,i}^{(1)}$ for all $i \neq i_0$. Since $\mathrm{Ave}(\theta_0)$ is fixed, any $\theta_i(k), i \in \{1, \ldots, n\}$ converges to $\mathrm{Ave}(\theta_0)$ in probability. Thus, for any $i \in \{1, \ldots, n\}$ and any $\varepsilon > 0$,

$$\lim_{k\to\infty} \mathbb{P}\{|\theta_i^{(\ell)}(k) - \mathrm{Ave}(\theta_0^{(\ell)})| < \varepsilon\} = 1, \quad \ell = 1, 2.$$

Therefore, for any $\varepsilon' > 0$, there exists $k \in \mathbb{Z}_{\geq 0}$ such that for all $i \in \{1, \ldots, n\}$,

$$\mathbb{P}\{|\theta_i^{(\ell)}(k) - \mathrm{Ave}(\theta_0^{(\ell)})| < \varepsilon\} > 1 - \varepsilon', \quad \ell = 1, 2. \qquad (6)$$

Now, consider (3), (4). It is clear that for any fixed initial state θ_0 and any $k \in \mathbb{Z}_{\geq 0}$, $\mathbf{x}_k$ is uniquely determined by $\boldsymbol{\eta}_k$ and $\boldsymbol{\theta}_k$ is uniquely determined by $\mathbf{x}_k$. Therefore, the functions $X_{k,\theta_0}, \Theta_{k,\theta_0} : \mathbb{R}^{n(k+1)} \to \mathbb{R}^{n(k+1)}$ such that

$$X_{k,\theta_0}(\boldsymbol{\eta}_k) = \mathbf{x}_k, \quad \Theta_{k,\theta_0}(\mathbf{x}_k) = \boldsymbol{\theta}_k \qquad (7)$$

are well defined. Next, define

$$R_k^{(1)} = \{\boldsymbol{\eta}_k \in \Omega_k \mid \forall i \in \{1, \ldots, n\}\ |\theta_i^{(1)}(k) - \mathrm{Ave}(\theta_0^{(1)})| < \varepsilon, \quad \boldsymbol{\theta}_k^{(1)} = \Theta_{k,\theta_0^{(1)}}(X_{k,\theta_0^{(1)}}(\boldsymbol{\eta}_k))\}, \qquad (8)$$

where $\Omega_k = \mathbb{R}^{n(k+1)}$ is the sample space up to time k. Note that $R_k^{(1)} \neq \emptyset$ because by (6),

$$\mathbb{P}(R_k^{(1)}) > 1 - \varepsilon'. \qquad (9)$$

Therefore,

$$\begin{aligned} \mathcal{O}_k &= X_{k,\theta_0^{(1)}}(R_k^{(1)}) && (10) \\ &= \{\mathbf{x}_k \in \mathbb{R}^{n(k+1)} \mid \exists \boldsymbol{\eta}_k \in R_k^{(1)} \quad \mathbf{x}_k = X_{k,\theta_0^{(1)}}(\boldsymbol{\eta}_k)\} \end{aligned}$$

is nonempty. Define

$$R_k^{(2)} = X_{k,\theta_0^{(2)}}^{-1}(\mathcal{O}_k) = \{\boldsymbol{\eta}_k \in \Omega_k \mid X_{k,\theta_0^{(2)}}(\boldsymbol{\eta}_k) \in \mathcal{O}_k\}. \qquad (11)$$

To reach a contradiction, we next show that $\mathbb{P}(R_k^{(2)})$ can be made arbitrarily small. To do this, pick any $\bar{\boldsymbol{\eta}}_k \in R_k^{(2)}$ and let $\bar{\mathbf{x}}_k = X_{k,\theta_0^{(2)}}(\bar{\boldsymbol{\eta}}_k) \in \mathcal{O}_k$ and $\bar{\boldsymbol{\theta}}_k^{(\ell)} = \Theta_{k,\theta_0^{(\ell)}}(\bar{\mathbf{x}}_k)$ for $\ell = 1, 2$. By (10),

$$\exists \bar{\boldsymbol{\eta}}'_k \in R_k^{(1)} \qquad \bar{\mathbf{x}}_k = X_{k,\theta_0^{(1)}}(\bar{\boldsymbol{\eta}}'_k).$$

Therefore, by (8),

$$\forall i \in \{1, \ldots, n\} \qquad |\bar{\theta}_i^{(1)}(k) - \mathrm{Ave}(\theta_0^{(1)})| < \varepsilon. \qquad (12)$$

Recall that in (3), we restricted f to be such that the next state of each agent only depends on its current state and the messages it receives. Hence, since for all $i \neq i_0$, $\theta_{0,i}^{(2)} = \theta_{0,i}^{(1)}$ and both $\bar{\boldsymbol{\theta}}_k^{(2)}$ and $\bar{\boldsymbol{\theta}}_k^{(1)}$ are constructed from $\bar{\mathbf{x}}_k$,

$$\forall i \neq i_0 \qquad \bar{\theta}_i^{(2)}(k) = \bar{\theta}_i^{(1)}(k) \xRightarrow{(12)} |\bar{\theta}_i^{(2)}(k) - \mathrm{Ave}(\theta_0^{(1)})| < \varepsilon.$$

This, together with the fact that $\mathrm{Ave}(\theta_0^{(2)}) = \mathrm{Ave}(\theta_0^{(1)}) + \frac{\delta}{n}$, shows that by taking $\varepsilon < \frac{\delta}{2n}$, the two events $R_k^{(2)}$ and $R_k^{'(2)}$ are disjoint, where

$$R_k^{'(2)} = \{\boldsymbol{\eta}_k \in \Omega_k \mid \forall i \in \{1, \ldots, n\}\ |\theta_i^{(2)}(k) - \mathrm{Ave}(\theta_0^{(2)})| < \varepsilon, \quad \boldsymbol{\theta}_k^{(2)} = \Theta_{k,\theta_0^{(2)}}(X_{k,\theta_0^{(2)}}(\boldsymbol{\eta}_k))\},$$

However, by (6), $\mathbb{P}(R_k^{'(2)}) > 1 - \varepsilon'$ so

$$\mathbb{P}(R_k^{(2)}) < \varepsilon'. \qquad (13)$$

Now, for $\mathcal{O}_k$ defined in (10), define a particular $\mathcal{O} \in (\mathbb{R}^n)^{\mathbb{N}}$ as the set of all $\mathbf{x}$ whose first subsequence of length $k+1$ belongs to $\mathcal{O}_k$. Therefore, for any initial condition θ_0,

$$\begin{aligned} \mathbb{P}\{\boldsymbol{\eta} \in \Omega \mid X_{\theta_0}(\boldsymbol{\eta}) \in &\mathcal{O}\} \\ &= \mathbb{P}\{\boldsymbol{\eta}_k \in \Omega_k \mid X_{k,\theta_0}(\boldsymbol{\eta}_k) \in \mathcal{O}_k\}. \end{aligned}$$

Therefore, since the algorithm is ϵ-differentially private,

$$\begin{aligned} \mathbb{P}\{\boldsymbol{\eta}_k \in \Omega_k \mid &X_{k,\theta_0^{(1)}}(\boldsymbol{\eta}_k) \in \mathcal{O}_k\} \\ &\leq e^{\epsilon}\mathbb{P}\{\boldsymbol{\eta}_k \in \Omega_k \mid X_{k,\theta_0^{(2)}}(\boldsymbol{\eta}_k) \in \mathcal{O}_k\}, \end{aligned}$$

but from (8) and (11), this is equivalent to

$$\mathbb{P}(R_k^{(1)}) \leq e^{\epsilon}\mathbb{P}(R_k^{(2)}).$$

Thus, using (9) and (13), we have for all $\varepsilon' > 0$,

$$1 - \varepsilon' < e^{\epsilon}\varepsilon' \Rightarrow \frac{1}{1 + e^{\epsilon}} < \varepsilon',$$

that is clearly a contradiction because ϵ is finite. This completes the proof. □

Since convergence in distribution is the weakest notion of convergence, Theorem 4.1 implies that a differentially private algorithm cannot guarantee any type of convergence to the true average. Therefore, next, we relax the exact convergence requirement and allow for convergence to a random variable centered at the true average.

5. DYNAMICS DESIGN AND ANALYSIS

In this section, we develop a solution to Problem 1. Consider the following linear distributed dynamics,

$$\theta(k+1) = \theta(k) - hLx(k) + S\eta(k), \quad k \in \mathbb{Z}_{\geq 0} \qquad (14)$$

with the messages generated according to

$$x(k) = \theta(k) + \eta(k), \quad k \in \mathbb{Z}_{\geq 0}, \qquad (15)$$

where $h < (d_{\max})^{-1}$ is the step size. Here, $d_{\max}$ is the largest entry of D and $S = \mathrm{diag}(s_1, \ldots, s_n) \in \mathbb{R}^n$ is a matrix of design parameters independently chosen by each agent. Note that (14) is a special case of (3) (since $\eta(k) = x(k) - \theta(k)$) and (15) a special case of (4). Also note that without the term $S\eta(k)$, the average of the agents' initial states would be preserved throughout the evolution.

The next result establishes the convergence of our design.

Theorem 5.1. (Asymptotic Convergence). Consider a network of n agents executing the distributed dynamics (14), (15). For each $i \in \{1, \ldots, n\}$, let $s_i \in (0, 2)$ and assume its messages are corrupted by Laplacian noise $\eta_i(k) \sim \mathrm{Lap}(b_i(k))$ at time $k \in \mathbb{Z}_{\geq 0}$ with

$$b_i(k) = c_i q_i^k, \qquad c_i \in \mathbb{R}_{>0},\ q_i \in (|s_i - 1|, 1). \qquad (16)$$

Then, the states of all agents converge almost surely to a the random variable θ_∞ defined as

$$\theta_\infty = \mathrm{Ave}(\theta_0) + \sum_{i=1}^{n} \frac{s_i}{n} \sum_{j=0}^{\infty} \eta_i(j) \qquad (17)$$

if this series converges and $\theta_\infty = \mathrm{Ave}(\theta_0)$, otherwise.

Remark 5.2. (Laplacian Noise Distribution). Even though Laplacian noise is not the only choice for achieving differential privacy, it is the predominant choice in the literature (Dwork et al., 2006; Dwork, 2006). The work (Wang et al., 2014) shows that Laplacian noise is optimal in the sense that it minimizes the entropy of the transmitted messages while preserving differential privacy.

The next result elaborates on the statistical properties of the agreement value of the algorithm.

Corollary 5.3. (Accuracy). The convergence point θ_∞ of the agents' states given in (17) is an unbiased estimate of $\mathrm{Ave}(\theta_0)$ with bounded dispersion given by

$$\mathrm{var}\{\theta_\infty\} = \frac{2}{n^2} \sum_{i=1}^{n} \frac{s_i^2 c_i^2}{1 - q_i^2}. \qquad (18)$$

As a result, the algorithm (14)-(16) is $\left(p, \frac{1}{n}\sqrt{\frac{2}{p}\sum_{i=1}^{n} \frac{s_i^2 c_i^2}{1 - q_i^2}}\right)$-accurate for any $p \in (0, 1)$.

Theorem 5.1 and Corollary 5.3 establish almost sure convergence, with the expected value of convergence being the average of the agents' initial states. In contrast, the results in (Huang et al., 2012) establish convergence in mean square, and the expected value of convergence depends on the network topology. In both cases, the accuracy radius r decreases with the number of agents as $O(1/\sqrt{n})$.

The expression for the (p, r)-accuracy, cf. Corollary 5.3, shows that one cannot get the ideal case of $(0, 0)$-accuracy, as expected, and that r is a decreasing function of p with $r \to \infty$ as $p \to 0$. This is a consequence of the lack of preservation of the average by (14) due to the term $S\eta(k)$. In turn, the presence of this expression helps establish the differential privacy of the algorithm with bounded, asymptotically vanishing noise, as we show next.

Theorem 5.4. (Differential Privacy). Consider a network of n agents executing the distributed dynamics (14), (15) with $s_i \in (0, 2)$ for $i \in \{1, \dots, n\}$ and messages corrupted by Laplacian noise according to (16). For each $i \in \{1, \dots, n\}$, let

$$\epsilon_i = \delta \frac{q_i}{c_i(q_i + s_i - 1)}, \tag{19}$$

where δ is the adjacency bound as in (5). Then, the algorithm preserves the ϵ_i-differential privacy of agent i's initial state. Consequently, the algorithm is ϵ-differential private with $\epsilon = \max_i \epsilon_i$.

Note that Theorem 5.4 implies that each agent can choose its own level of privacy, and even opt not to add any noise to its messages, without affecting the privacy of other agents. In contrast, in (Huang et al., 2012), agents need to agree on the level of privacy before executing the algorithm. In both cases, privacy is achieved against an adversary that can hear everything and independently of how it processes the information. In contrast, the algorithm in (Mo and Murray, 2014) assumes the adversary uses maximum likelihood estimation and only preserves the differential privacy of those agents who are sufficiently "far" from it in the graph (i.e., the adversary cannot listen to an agent and all its neighbors).

6. OPTIMAL NOISE SELECTION

In this section, we discuss the effect of the free parameters present in our design on its performance. Assuming the privacy levels $\{\epsilon_i\}_{i=1}^n$ are fixed according to the agents' privacy requirements, the free parameters that each agent $i \in \{1, \dots, n\}$ gets to select are s_i, c_i, and q_i, which together determine the amount of noise introduced into the dynamics. Given the overall network objective, we consider the cost function as the variance of the agents' convergence point around $\mathrm{Ave}(\theta_0)$, i.e.,

$$J\left(\{s_i, c_i, q_i\}_{i=1}^n\right) = \frac{2}{n^2} \sum_{i=1}^n \frac{s_i^2 c_i^2}{1 - q_i^2}. \tag{20}$$

where $(s_i, c_i, q_i) \in \mathcal{P} = \{(s, c, q) \mid s \in (0, 2), c > 0, q \in (|s - 1|, 1)\}$, for each $i \in \{1, \dots, n\}$. The next result characterizes the global minimization of J.

Proposition 6.1. (Optimal Parameters). For a given adjacency bound $\delta > 0$ and set of privacy levels $\{\epsilon_i\}_{i=1}^n$, one has

$$\inf_{\{s_i, c_i, q_i\}_{i=1}^n \in \mathcal{P}^n} J = \frac{\delta^2}{2n^2} \sum_{i=1}^n \frac{1}{\epsilon_i^2}.$$

Furthermore, the infimum is not attained over $\mathcal{P}^n$ but approached if

$$c_i = \delta \frac{q_i}{\epsilon_i(q_i + s_i - 1)}, \tag{21}$$

and $(s_i, q_i) \to (1, 0)$ along $\gamma = \{(1 + t, t + t^2) \mid t > 0\}$ for all $i \in \{1, \dots, n\}$.

Based on Proposition 6.1, we have each agent $i \in \{1, \dots, n\}$ independently choose its noise parameters according to $(s_i, q_i) = (1 + \varepsilon, \varepsilon + \varepsilon^2)$, with $\varepsilon \ll 1$.

7. SIMULATIONS

In this section, we report simulation results of the distributed dynamics (14) on a network of $n = 50$ agents. The network topology and agents' initial conditions are chosen randomly, the latter from $\mathcal{N}(50, 100)$ for all agents. As can be seen from (18) and (19), neither privacy nor accuracy depend on the communication topology or the initial values although the convergence speed might. We set $\delta = 1$, $\epsilon_i = 0.1$, $s_i = s$, and $\alpha_i \triangleq \frac{q_i - |s_i - 1|}{1 - |s_i - |} = 10^{-6}$ for all $i \in \{1, \dots, n\}$.

Figure 1 depicts the result of performing simulations sweeping s over $[0.97, 1.9]$ with logarithmic step size. For each value of s, we repeated the simulation 10^4 times. For each run, the graph topology and initial conditions are the same and only noise realizations change between different runs in order to capture the statistical properties of the convergence point. Figure 1(a) shows the empirical (sample) standard deviation of the convergence point as a function of s. In particular, notice the sensitivity of the accuracy to s for $s < 1$. Figure 1(b) shows the number of rounds until convergence (measured by a tolerance of 10^{-2}) as a function of s. The optimal value for s in this case is different from 1 and in general depends on the network connectivity. The results obtained from different random choices of initial conditions and network topologies show very consistent trends as those in Figure 1.

Figure 2 shows the histogram of convergence points for 10^5 runs of the algorithm with $s = 1 + 10^{-6}$ (near optimal accuracy). The distribution of the convergence point is a bell-shaped curve with mean exactly at the true average, in accordance with Corollary 5.3.

8. CONCLUSIONS

We have considered the multi-agent average consensus problem with privacy preservation requirements on the initial agent states. We have formally proved that, under some mild conditions, no coordination algorithm can guarantee convergence to the true average of the agents' initial states while preserving differential privacy against an adversary that has full access to the network communications. As a consequence, the most one can expect of a differentially private algorithm is to guarantee convergence to the true average in expectation and minimize the dispersion around it. We have achieved this objective by designing a noise-optimal linear dynamics with almost sure guaranteed convergence. Future work will include further investigation of the benefits and costs of differential privacy for multi-agent systems, the extension of the results to distributed optimization, filtering, and estimation, and the design of algorithms for privacy preservation of the

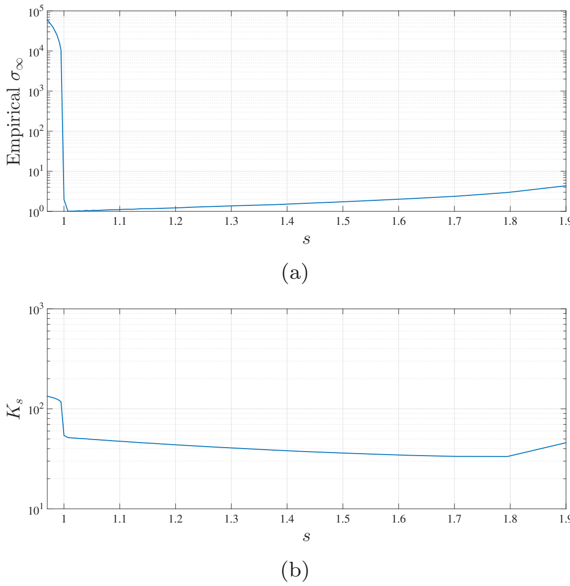

Fig. 1. (a) Empirical standard deviation of the convergence point of the algorithm (14) and (b) its settling time (K_s) in rounds for $\alpha = 10^{-6}$ and random topology and initial conditions. The trend in (a) validates the results of Section 6 while (b) shows the trade-off between accuracy and convergence speed.

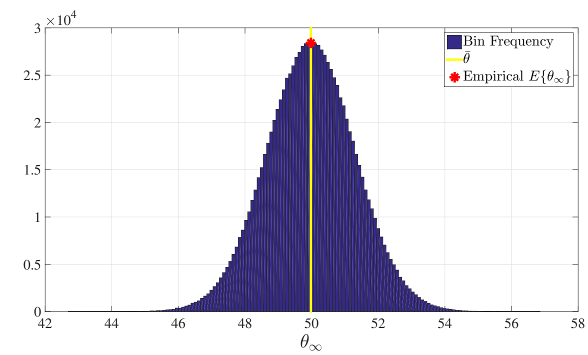

Fig. 2. Statistical distribution of the convergence point. The sample mean (starred) matches the true average (yellow vertical line).

network structure and other parameters relevant for the execution of cooperative strategies.

ACKNOWLEDGMENTS

This work was supported by NSF Award CNS-1329619.

REFERENCES

Bullo, F., Cortés, J., and Martínez, S. (2009). *Distributed Control of Robotic Networks*. Applied Mathematics Series. Princeton University Press. Electronically available at http://coordinationbook.info.

Durrett, R. (2010). *Probability: Theory and Examples*. Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 4th edition.

Dwork, C. (2006). Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP)*, 1–12. Venice, Italy.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In *Proceedings of the 3rd Theory of Cryptography Conference*, 265–284. New York, NY.

Dwork, C. and Roth, A. (2014). The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4), 211–407.

Han, S., Topcu, U., and Pappas, G.J. (2014). Differentially private convex optimization with piecewise affine objectives. In *IEEE Conf. on Decision and Control*, 2160–2166. Los Angeles, CA.

Huang, Z., Mitra, S., and Dullerud, G. (2012). Differentially private iterative synchronous consensus. In *Proceedings of the 2012 ACM workshop on Privacy in the electronic society*, 81–90. New York, NY.

Huang, Z., Mitra, S., and Vaidya, N. (2015). Differentially private distributed optimization. In *Proceedings of the 2015 International Conference on Distributed Computing and Networking*. Pilani, India.

Huang, Z., Wang, Y., Mitra, S., and Dullerud, G.E. (2014). On the cost of differential privacy in distributed control systems. In *Proceedings of the 3rd International Conference on High Confidence Networked Systems (HiCoNS)*, 105–114. Berlin, Germany.

Jiang, Z.P. and Wang, Y. (2001). Input-to-state stability for discrete-time nonlinear systems. *Automatica*, 37(6), 857–869.

Kefayati, M., Talebi, M.S., Khalaj, B.H., and Rabiee, H.R. (2007). Secure consensus averaging in sensor networks using random offsets. In *IEEE Intern. Conf. on Telec., and Malaysia Intern. Conf. on Communications*, 556–560. Penang.

Manitara, N.E. and Hadjicostis, C.N. (2013). Privacy-preserving asymptotic average consensus. In *European Control Conference*, 760–765. Zurich.

Mesbahi, M. and Egerstedt, M. (2010). *Graph Theoretic Methods in Multiagent Networks*. Applied Mathematics Series. Princeton University Press.

Mo, Y. and Murray, R.M. (2014). Privacy preserving average consensus. In *IEEE Conf. on Decision and Control*, 2154–2159. Los Angeles, CA.

Ny, J.L. and Pappas, G.J. (2014). Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2), 341–354.

Papoulis, A. and Pillai, S.U. (eds.) (2002). *Probability, Random Variables and Stochastic Processes*. McGraw-Hill.

Ren, W. and Beard, R.W. (2008). *Distributed Consensus in Multi-Vehicle Cooperative Control*. Communications and Control Engineering. Springer.

Wang, Y., Huang, Z., Mitra, S., and Dullerud, G.E. (2014). Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems. In *IEEE Conf. on Decision and Control*, 2130–2135. Los Angeles, CA.