



Research Repository UCD

Title	Resilient consensus for multi-agent systems subject to differential privacy requirements
Authors(s)	Fiore, Davide, Russo, Giovanni
Publication date	2019-08
Publication information	Fiore, Davide, and Giovanni Russo. "Resilient Consensus for Multi-Agent Systems Subject to Differential Privacy Requirements." Elsevier, August 2019. https://doi.org/10.1016/j.automatica.2019.04.029 .
Publisher	Elsevier
Item record/more information	http://hdl.handle.net/10197/12285
Publisher's statement	This is the author's version of a work that was accepted for publication in Automatica. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in Automatica (106, (2019)) https://doi.org/10.1016/j.automatica.2019.04.029
Publisher's version (DOI)	10.1016/j.automatica.2019.04.029

Downloaded 2025-12-04 22:43:41

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

Resilient consensus for multi-agent systems subject to differential privacy requirements [★]

Davide Fiore ^a, Giovanni Russo ^b

^a*Department of Electrical Engineering and Information Technology, University of Naples Federico II, Via Claudio 21, 80125 Naples, Italy*

^b*School of Electrical and Electronic Engineering, University College Dublin, Belfield, Dublin 4, Ireland*

Abstract

We consider multi-agent systems interacting over directed network topologies where a subset of agents is adversary/faulty and where the non-faulty agents have the goal of reaching consensus, while fulfilling a differential privacy requirement on their initial conditions. To address this problem, we develop an update law for the non-faulty agents. Specifically, we propose a modification of the so-called Mean-Subsequence-Reduced (MSR) algorithm, the Differentially Private MSR (DP-MSR) algorithm, and characterize three important properties of the algorithm: correctness, accuracy and differential privacy. We show that if the network topology is $(2f+1)$ -robust, then the algorithm allows the non-faulty agents to reach consensus despite the presence of up to f faulty agents and we characterize the accuracy of the algorithm. Furthermore, we also show in two important cases that our distributed algorithm can be tuned to guarantee differential privacy of the initial conditions and the differential privacy requirement is related to the maximum network degree. The results are illustrated via simulations.

Key words: Multi-agent systems, Resilient consensus, Differential privacy, Networked control systems

1 Introduction

The introduction of low-cost, high performance and connected devices lead, over the last few years, to a paradigm shift on how engineering systems are designed (di Bernardo et al., 2016; Albert et al., 2000; Fiore et al., 2017). The philosophy design for those systems is indeed transitioning from a centralized to a distributed paradigm and, in this context, two fundamental challenges are: (i) the fact that agents/nodes might be subject of failures or attacks; and (ii) privacy preservation of individual information. Power grids (Dörfler and Bullo, 2014) and smart transportation systems (Griggs et al., 2018) are examples of systems where control and coordination strategies cannot neglect these fundamental requirements of privacy preservation and fault tolerance. Motivated by this, we present an algorithm that: (i) allows a multi-agent system to reach consensus despite a subset of agents being faulty; and (ii) guarantees differential privacy of the initial conditions of the

non-faulty agents. We now briefly review some related works on Differential Privacy and Resilient Consensus.

Differential Privacy. The notion of differential privacy was originally introduced in (Dwork et al., 2006a,b) in relation to the design of privacy preserving mechanisms of databases. From the intuitive viewpoint, a database is differentially private if the removal or addition of a single database entry does not substantially affect the outcome of any analysis on the database data. This privacy definition is also particularly well-suited for networked multi-agent systems, where communication channels might not be fully secure and where agents might not trust each other. In the context of multi-agent systems, protecting initial information of each agent often implies protecting the current state of the network. This is the case, for example, of platoon systems, see e.g. (Monteil and Russo, 2017), where each vehicle in the platoon wishes to keep private its initial position. Based on such observations, (Huang et al., 2012) considered the problem of designing a distributed consensus algorithm for multi-agent systems (evolving over an undirected graph topology) guaranteeing differential privacy of the initial conditions of the agents. In particular, the paper shows that if each agent adds some decaying Laplacian noise to its averaging dynamics, then mean square convergence to a ran-

[★] This work was conducted while the authors were with IBM Research Ireland. Corresponding author: Giovanni Russo.

Email addresses: `dvd.fiore@gmail.com` (Davide Fiore), `giovanni.russo1@ucd.ie` (Giovanni Russo).

dom consensus value, whose expected value depends on the undirected network topology, can be proved and differential privacy of the initial conditions can be guaranteed. More recently, (Nozari et al., 2017) proposed an algorithm which guarantees both almost sure convergence to an unbiased estimate of the average of agents' initial conditions and differential privacy of their values, allowing individual agents, interacting over an undirected graph, to independently choose their level of privacy. The concept of differential privacy has been also successfully used in distributed estimation (Le Ny and Pappas, 2014) and optimization (Han et al., 2014; Huang et al., 2015; Nozari et al., 2018), while other notions of privacy, see (Mo and Murray, 2017) and references therein for further details, have been also investigated for consensus.

Resilient Consensus. Resilient consensus has received much research attention and we refer the reader to (Tseng and Vaidya, 2016) and (LeBlanc et al., 2013) for a detailed literature review. The original formulation of the problem can be traced back to (Lamport et al., 1982; Dolev, 1982). In such works, the authors introduced the *Approximate Byzantine Consensus problem* where the agents have to agree on some legit value despite the fact that at most f agents may be adversary (also termed faulty or Byzantine). The networks considered in the paper were complete and, for such networks, the *Mean-Subsequence-Reduced (MSR)* algorithm was proposed, see also (Kieckhafer and Azadmanesh, 1994). Tight conditions for approximate consensus using the MSR algorithm in directed networks with f Byzantine agents were derived in (Vaidya et al., 2012), while, in (LeBlanc et al., 2013; Zhang and Sundaram, 2012) a modification of the MSR algorithm has been proposed, the Weighted-MSR. The conditions derived therein were stated in terms of the so-called r -robustness property that characterizes the local redundancy of information flow (Zhang et al., 2015). Recent works (LeBlanc and Koutsoukos, 2011, 2012) proposed a continuous-time variation of the MSR algorithm, while in (Su and Vaidya, 2017) the authors allowed agents to take as input not only the information from their neighbors but also the information provided by the agents that are up to ℓ hops away. In (Dibaji et al., 2018; Dibaji and Ishii, 2017), a modification of the MSR algorithm has been proposed to solve the resilient consensus problem for asynchronous multi-agent systems with quantized communication and delays. The works (Sundaram and Hadjicostis, 2011; Pasqualetti et al., 2012) proposed a different approach to detect and identify malicious agents, which is based on observability and reconstruction of the initial conditions. Other complementary approaches to the problem include adding trusted agents (Abbas et al., 2018) and designing self-triggered control strategies (Senejohnny et al., 2018).

1.1 Contributions of the paper

In this paper we combine differential privacy and resilient consensus in a multi-agent system with directed communication topology. For such systems, we introduce a distributed algorithm guaranteeing that consensus can be attained by honest agents in presence of a certain number of malicious agents along with differential privacy of their initial states. We characterize algorithm correctness and accuracy. Moreover, in two special cases, we show how differential privacy of the initial conditions of the non-faulty agents can be guaranteed. The Differentially Private MSR (DP-MSR) algorithm presented here is the main contribution of this paper and, to the best of our knowledge, this is the first algorithm that simultaneously handles differential privacy and fault tolerance requirements over directed graphs. Indeed:

- (i) in the literature on differentially private consensus (Huang et al., 2012; Nozari et al., 2017; Mo and Murray, 2017), it is assumed that the graph topology is undirected and that there are no malicious agents;
- (ii) vice versa, the literature on resilient consensus (see the related work section) considers agents evolving over directed graphs but, in such works, no differentially private mechanisms are considered.

2 Mathematical Preliminaries

Throughout the paper, we denote by $\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Z}_{\geq 0}$ the set of real numbers, positive real numbers and non-negative integers, respectively. Let $A \in \mathbb{R}^{n \times n}$ ($x \in \mathbb{R}^n$); we denote by $\|A\|$ ($\|x\|$) the Euclidean matrix (vector) norm and by $\|A\|_F = [\text{tr}(A^T A)]^{1/2}$ the Frobenius norm ($\text{tr}(\cdot)$ denotes the trace of a square matrix). Let \mathcal{A} and \mathcal{B} be two sets, then $|\mathcal{A}|$ is the cardinality of \mathcal{A} and we denote by $\mathcal{A} \cup \mathcal{B}, \mathcal{A} \cap \mathcal{B}$ and $\mathcal{A} \setminus \mathcal{B}$ the union, intersection and difference of the sets. Let $(\mathbb{R}^n)^{\mathbb{N}}$ denote the space of vector-valued sequences in \mathbb{R}^n . Then, for $\{x(k)\}_{k=0}^{\infty} \in (\mathbb{R}^n)^{\mathbb{N}}$, we set $\mathbf{x} := \{x(k)\}_{k=0}^{\infty}$ and $\mathbf{x}_k := \{x(h)\}_{h=0}^k$. For any topological space, say X , we denote by $\mathcal{B}(X)$ the set of Borel subsets of X . We denote by $\overline{\text{co}}\{\mathcal{A}\}$ the closed convex hull of the set \mathcal{A} . Finally, we denote by $\mathbf{1}$ the column vector of adequate dimension with all entries equal to 1, and by I_n the identity matrix of size n .

2.1 Probability theory

Consider a probability space $(\Omega, \Sigma, \mathbb{P})$, where Ω is the sample space, Σ is the collection of all the events, and \mathbb{P} is the probability measure, respectively. A random variable is a measurable function $X : \omega \in \Omega \rightarrow \mathbb{R}$, and we denote its expected value by $\mathbb{E}[X]$. Chebyshev's inequality states that, for any $N \in \mathbb{R}_{>0}$ and any random variable X with finite expected value μ and finite nonzero variance σ^2 , $\mathbb{P}\{|X - \mu| \geq N\sigma\} \leq \frac{1}{N^2}$. We denote by $X \sim \text{Lap}(b)$ a zero mean random variable with

Laplace distribution and variance $2b^2$, and with probability density function $\mathcal{L}(x; b) = \frac{1}{2b} e^{-\frac{|x|}{b}}$ for any $x \in \mathbb{R}$. Finally, consider the random n -dimensional vector $X = [X_1, \dots, X_n]^T$. The expected value of X is a vector whose elements are the expected value of elements of X , that is $\mathbb{E}[X] = [\mathbb{E}[X_1], \dots, \mathbb{E}[X_n]]^T$, and the covariance matrix of X is $\text{Cov}(X) = \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^T]$. We denote with $\text{var}(X)$ the trace of the covariance matrix. That is, $\text{var}(X) = \text{tr}(\text{Cov}(X)) = \mathbb{E}[X^T X] - \mathbb{E}[X]^T \mathbb{E}[X]$.

2.2 Directed graphs and r -robustness

We consider agents interacting over a network topology modeled as a directed graph, or *digraph*. We denote a digraph by $\mathcal{G}(\mathcal{V}, \mathcal{E})$, where \mathcal{V} is the *node set*, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is the *directed edge set*. We denote the number of nodes in the graph by n and we consider graphs for which $n \geq 2$. A directed edge is an ordered pair of distinct vertices and we denote by $(j, i) \in \mathcal{E}$ an edge from node $j \in \mathcal{V}$ to node $i \in \mathcal{V}$. In the context of this paper, the directed edge (j, i) models the information flow from agent j to agent i . The set of *in-neighbors* of node i is denoted by $\mathcal{N}_i := \{j \in \mathcal{V} : (j, i) \in \mathcal{E}\}$, and the *in-degree* of node i , denoted by d_i , is the cardinality of this set. Likewise, the set of *out-neighbors* of node i is defined as $\mathcal{N}_i^{\text{out}} := \{j \in \mathcal{V} : (i, j) \in \mathcal{E}\}$ and the *out-degree* is denoted by $d_i^{\text{out}} := |\mathcal{N}_i^{\text{out}}|$. In the rest of the paper, d_{\max}^{out} is the maximum outgoing degree in the network, that is, $d_{\max}^{\text{out}} := \max_{i \in \mathcal{V}} \{d_i^{\text{out}}\}$. We also denote the minimum in-degree by $d_{\min}^{\text{in}} := \min_{i \in \mathcal{V}} \{d_i\}$. For a digraph $\mathcal{G}(\mathcal{V}, \mathcal{E})$, a nonempty subset $\mathcal{S} \subset \mathcal{V}$ is an *r -reachable set* if there exists a node $i \in \mathcal{S}$ such that $|\mathcal{N}_i \setminus \mathcal{S}| \geq r$. That is, the subset \mathcal{S} has at least r in-neighbors in $\mathcal{N}_i \setminus \mathcal{S}$ (LeBlanc et al., 2013; Vaidya et al., 2012). This leads to the following concept of *r -robustness* for the digraph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ that characterizes its local redundancy of information flow (LeBlanc et al., 2013; Vaidya et al., 2012; Zhang et al., 2015).

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

As in (LeBlanc et al., 2013), we say that a set $\mathcal{S} \subset \mathcal{V}$ is *f -total* if $|\mathcal{S}| \leq f$, $f \in \mathbb{Z}_{\geq 0}$. The set \mathcal{S} is said to be *f -totally bounded* if it is an f -total set.

2.3 Stochastic consensus

In (Huang et al., 2010; Huang, 2012) consensus algorithms were considered, which can be written as

$$\theta(k+1) = A(k)\theta(k) + H(k)v(k), \quad k \geq 0. \quad (1)$$

In the above equation: (i) $\theta(k) = [\theta_1(k) \dots \theta_n(k)]^T \in \mathbb{R}^n$ denotes the states of n agents at time k ; (ii) the initial condition is $\theta_0 = [\theta_{0,1}, \dots, \theta_{0,n}]^T$, with $\mathbb{E}[\|\theta_0\|^2] <$

∞ ; (iii) $v(k) \in \mathbb{R}^m$ is a noise vector, with $\{v(k), k \geq 0\}$ being a sequence of independent random vectors of zero mean such that $\sup_{k \geq 0} \mathbb{E}[\|v(k)\|^2] < \infty$. Moreover, $\{A(k), k \geq 0\}$ and $\{H(k), k \geq 0\}$ are two sequences of random matrices of appropriate dimensions. For each fixed $\omega \in \Omega$, $A(k, \omega)$ is a row-stochastic matrix for all $k \geq 0$. We can now give the following from (Huang and Manton, 2010; Huang et al., 2010; Huang, 2012).

Definition 2 The agents of a network are said to reach mean square consensus if $\mathbb{E}[\|\theta_i(k)\|^2] < \infty$, $\forall k \geq 0$, $i \in \mathcal{V}$, and there exists a random variable θ_∞ such that $\lim_{k \rightarrow \infty} \mathbb{E}[\|\theta_i(k) - \theta_\infty\|^2] = 0$ for all $i \in \mathcal{V}$.

In the above definition, convergence of $\theta_i(k)$ to θ_∞ is meant as convergence in mean square and in the rest of the paper we simply say that $\theta_i(k)$ converges to θ_∞ . As in (Huang, 2012), we define the backward product

$$\Psi_{k,s} := A(k-1) \dots A(s), \quad \text{for } k > s \geq 0, \quad \Psi_{s,s} := I.$$

Let $\Psi_{k,s}(i, j)$ be the (i, j) -th element of $\Psi_{k,s}$. We say that the backward product of the sequence of stochastic matrices $\{A(k), k \geq 0\}$ is *ergodic* if $\lim_{k \rightarrow \infty} |\Psi_{k,s}(i_1, j) - \Psi_{k,s}(i_2, j)| = 0$, for any given $s \geq 0$, and i_1, i_2, j . The following result from (Huang, 2012, Theorem 3) is also used in this paper.

Theorem 1 Consider (1) and assume that:

- (i) $\{v(k), k \geq 0\}$ is a sequence of independent random vectors of zero mean with bounded variance;
- (ii) $\sum_{k=0}^{\infty} \mathbb{E}[\|H(k)\|_F^2] \mathbb{E}[v(k)^T v(k)] < \infty$;
- (iii) $\{A(k), k \geq 0\}$ has ergodic backward product with probability 1.

Then, for any initial state θ_0 with bounded variance, system (1) achieves mean square consensus.

3 Problem Formulation

We now formalize the problem statement and, in order to do so, we first introduce the notions of adversary agent, resilient consensus and differential privacy. To this aim, consider a multi-agent system of n agents evolving on a digraph $\mathcal{G}(\mathcal{V}, \mathcal{E})$. The internal state of the i -th agent at time $k \in \mathbb{Z}_{\geq 0}$ is denoted by $\theta_i(k) \in \mathbb{R}$ and by $\theta_{i,0} := \theta_i(0)$ its initial condition. The message sent at time k from agent i is $x_i(k) \in \mathbb{R}$, while the messages received by agent i at time k from its in-neighbors are denoted by $\{x_j(k), j \in \mathcal{N}_i\}$. The internal state of agent $i \in \mathcal{V}$ is updated according to some update rule of the form

$$\theta_i(k+1) = g_i(k, \theta_i(k), \{x_j(k), j \in \mathcal{N}_i\}), \quad k \in \mathbb{Z}_{\geq 0}, \quad (2)$$

while the message sent by the i -th agent at time k is generated as

$$x_i(k) = h(\theta_i(k), \eta_i(k)), \quad k \in \mathbb{Z}_{\geq 0}, \quad (3)$$

with $\eta_i(k) \in \mathbb{R}$ being some noise injected by agent i at time k with some arbitrary distribution. We remark that the update rule $g_i(\cdot)$ may be different from agent to agent, while the function $h(\cdot)$ is the same for all the agents, i.e. agents use the same communication policy.

3.1 Adversary model and resilient consensus

Suppose that the network node set \mathcal{V} is partitioned into a set of normal or *non-faulty* nodes, evolving according to (2) and (3), and a set of adversary or *faulty* nodes $\mathcal{F} \subset \mathcal{V}$. Moreover, $|\mathcal{F}| = \phi \leq f$, i.e. the set of faulty nodes is f -totally bounded (only the upper bound f is known and not the actual ϕ). We use the following classification from (LeBlanc et al., 2013).

Definition 3 *An agent $i \in \mathcal{F}$ is said to be Byzantine if it does not send the same value to all of its neighbors at some time-step, or if it does not follow the same update rule as the non-faulty agents. An agent $i \in \mathcal{F}$ is said to be malicious if it sends the same value to all of its neighbors, but it does not follow the same update rule as the non-faulty agents.*

Both malicious and Byzantine agents are allowed to update their states *arbitrarily* and can collude among them. Furthermore, the faulty agents are often assumed to be *omniscient*, i.e. they know the network topology, the dynamics of the other agents and their initial conditions. Note that all malicious agents are Byzantine but not vice versa. Therefore, Byzantine agents are more general and in this paper we use the terminology *faulty agent* to denote both Byzantine and malicious agents.

Each faulty agent $i \in \mathcal{F}$ can send a different, arbitrary, attack signal $\hat{x}_i^j(k)$ to each of its out-neighbors, $j \in \mathcal{N}_i^{\text{out}}$. The attack signals \hat{x}_i^j are injected by the faulty agents into the network to steer the multi-agent system towards some malicious consensus state $\hat{\theta}_\infty$. In applications, this malicious state $\hat{\theta}_\infty$ may either be dangerous for the integrity of the multi-agent system or it may be defined to corrupt the computation carried out by the network. Given that faulty agents are omniscient (worst case situation), the signal $\hat{x}_i^j(k)$ depends on the malicious consensus state $\hat{\theta}_\infty$ and on the present and past internal state history of non-faulty nodes $\mathcal{V} \setminus \mathcal{F}$. That is,

$$\hat{x}_i^j(k) = \hat{h}_i^j(\hat{\theta}_\infty; \theta(0), \dots, \theta(k)), \quad k \in \mathbb{Z}_{\geq 0}, \quad (4)$$

where $\theta(k) = [\theta_1(k), \dots, \theta_{n-\phi}(k)]^T$ is the stack vector of internal state of non-faulty nodes at time k . In the context of this paper, the non-faulty agents neither are aware of the attack law in (4) nor are interested to detect and isolate the faulty agents. In what follows, we denote with the *hat* all the quantities related to faulty agents.

We are now ready to give the following definition, adapted from (LeBlanc et al., 2013; Vaidya et al., 2012; Huang and Manton, 2009).

Definition 4 *The non-faulty agents are said to achieve resilient asymptotic consensus if for any choice of their initial values the following two conditions are verified:*

- **Validity:** $\min_{j \in \mathcal{V} \setminus \mathcal{F}} \mathbb{E}[\theta_j(k)] \leq \mathbb{E}[\theta_i(k+1)] \leq \max_{j \in \mathcal{V} \setminus \mathcal{F}} \mathbb{E}[\theta_j(k)]$, for all $k \geq 0$ and $i \in \mathcal{V} \setminus \mathcal{F}$;
- **Convergence:** $\lim_{k \rightarrow \infty} \mathbb{E}[|\theta_i(k) - \theta_j(k)|^2] = 0$, for all $i, j \in \mathcal{V} \setminus \mathcal{F}$, $i \neq j$.

Definition 4 implies that non-faulty agents achieve mean square consensus to a value that is between the smallest and largest values of their initial conditions. A consensus algorithm that guarantees both the validity and the convergence conditions is said to be *correct*, see (LeBlanc et al., 2013; Vaidya et al., 2012; Huang and Manton, 2009).

3.2 Differential privacy

We introduce differential privacy by closely following (Nozari et al., 2017). Here, we let $\theta(k)$, $x(k)$ and $\eta(k)$ be the stack vectors of all states/attack signals, messages sent and generated noises at time k in the whole network, respectively. These vectors are in \mathbb{R}^m , with $m = (n - \phi) + \sum_{i \in \mathcal{F}} d_i^{\text{out}}$. We denote by $\boldsymbol{\eta} = \{\eta(k)\}_{k=0}^\infty$ the sequence of vector noise on the total sample space $\Omega = (\mathbb{R}^m)^\mathbb{N}$. Given an initial state θ_0 , the sequences of random vector variables $\mathbf{x} = \{x(k)\}_{k=0}^\infty$ and $\boldsymbol{\theta} = \{\theta(k)\}_{k=0}^\infty$ are uniquely determined by $\boldsymbol{\eta}$. Therefore, the function $X_{\theta_0} : (\mathbb{R}^m)^\mathbb{N} \rightarrow (\mathbb{R}^m)^\mathbb{N}$ such that $X_{\theta_0}(\boldsymbol{\eta}) = \mathbf{x}$ is well defined and represents the *observation* corresponding to the *execution* $(\theta_0, \boldsymbol{\eta})$.

Definition 5 *Given $\delta \in \mathbb{R}_{\geq 0}$, the initial states of the non-faulty agents of the network $\theta_0^{(1)}$ and $\theta_0^{(2)}$ are δ -adjacent if, for some $i_0 \in \mathcal{V} \setminus \mathcal{F}$, $|\theta_{0,i}^{(2)} - \theta_{0,i}^{(1)}| \leq \delta$ if $i = i_0$ and $|\theta_{0,i}^{(2)} - \theta_{0,i}^{(1)}| = 0$ if $i \neq i_0$, for $i \in \mathcal{V} \setminus \mathcal{F}$. Given $\delta, \varepsilon \in \mathbb{R}_{\geq 0}$, the network dynamics is ε -differentially private if, for any pair $\theta_0^{(1)}$ and $\theta_0^{(2)}$ of δ -adjacent initial states and any set $\mathcal{O} \in \mathcal{B}((\mathbb{R}^{n-\phi})^\mathbb{N})$,*

$$\mathbb{P}\{\boldsymbol{\eta} \in \Omega | X_{\theta_0^{(1)}}(\boldsymbol{\eta}) \in \mathcal{O}\} \leq e^\varepsilon \mathbb{P}\{\boldsymbol{\eta} \in \Omega | X_{\theta_0^{(2)}}(\boldsymbol{\eta}) \in \mathcal{O}\}. \quad (5)$$

We also give a definition of accuracy, which is adapted from (Nozari et al., 2017).

Definition 6 *For $p \in [0, 1]$ and $r \in \mathbb{R}_{\geq 0}$, the network dynamics is (p, r) -accurate if, for any initial condition, the state of each non-faulty agent converges, in any sense, as $k \rightarrow \infty$ to a random variable θ_∞ , with $\mathbb{E}[\theta_\infty] < \infty$, and $\mathbb{P}\{|\theta_\infty - \mathbb{E}[\theta_\infty]| \leq r\} \geq 1 - p$.*

3.3 Problem statement

Given the set-up described above, consider a multi-agent system and assume that a f -totally bounded set of faulty agents exists, $\mathcal{F} \subset \mathcal{V}$. Our goal in this paper is to design a distributed consensus algorithm ensuring that: (i) resilient asymptotic consensus is achieved by the non-faulty agents with (p, r) -accuracy. That is, the algorithm allows the non-faulty agents to achieve consensus despite the presence of faulty agents trying to steer the multi-agent system toward some malicious consensus state $\hat{\theta}_\infty$; (ii) ϵ -differential privacy of the initial conditions of non-faulty agents is guaranteed when faulty agents behave arbitrarily within the communication policy adopted in the system. The algorithm presented in this paper consists of: (i) an update law of the form (2); (ii) an inter-agent message generator of the form (3); (iii) the distribution of the noise processes $\eta_i(k)$.

4 The DP-MSR algorithm

We propose a modification of the original MSR algorithm, which we term as Differentially Private MSR algorithm (DP-MSR). The key steps of the DP-MSR algorithm are summarized as pseudo-code in Algorithm 1. An agent that follows the DP-MSR: (i) transmits to its neighbors a message that is *corrupted* with noise (rather than its internal state as in the MSR). Note that the *Transmit phase* in Algorithm 1 defines the communication policy of the network; (ii) discards, as in the MSR, some of the (noisy) messages from the neighbors before updating its internal state. In Section 5 we characterize correctness, accuracy and differential privacy of the DP-MSR algorithm. In particular, we show that these properties are related to the topology of the underlying graph \mathcal{G} .

Following Algorithm 1, each agent i executes at every time step k the following three actions:

- (1) *Transmit phase*: agent i transmits to all its out-neighbors the message x_i , generated as

$$x_i(k) = \theta_i(k) + \eta_i(k), \quad (6)$$

where $\eta_i(k)$ is a zero-mean noise with Laplacian distribution $\eta_i(k) \sim \text{Lap}(b(k))$ and $b(k) = cq^k$, $c > 0$, $\frac{1}{2} < q < 1$;

- (2) *Receive phase*: agent i receives the messages $\{x_j(k), j \in \mathcal{N}_i\}$ from all its in-neighbors. These values are used to create the vector $r_i(k)$ with dimension $|\mathcal{N}_i(k)|$;
- (3) *Update phase*: agent i stores the values contained in $r_i(k)$ and removes from such a vector the smallest f values and the largest f values (breaking ties arbitrarily). We denote by $\mathcal{N}_i^*(k)$ the set of in-neighbors of the i -th agent whose values have not

Algorithm 1 DP-MSR

Variables:

$\theta_i(k)$ internal state of agent i at time k
 $\eta_i(k)$ independent zero-mean noise with Laplace distribution, i.e. $\eta_i(k) \sim \text{Lap}(b(k))$, with: $b(k) = cq^k$, and $c > 0$, $\frac{1}{2} < q < 1$ being two design parameters
 f design parameter

Initialization:

Set $\theta_i(0) \leftarrow \theta_{0,i}$ and $a_i \leftarrow 1/(|\mathcal{N}_i| - 2f + 1)$

while True do

Transmit phase

Send to each out-neighbor $x_i(k) \leftarrow \theta_i(k) + \eta_i(k)$

Receive phase

Get $x_j(k)$ from each in-neighbor

Store the values in the vector $r_i(k)$ of size $|\mathcal{N}_i|$

Update phase

Remove from $r_i(k)$ the largest and smallest f values

Compute:

$$\theta_i(k+1) \leftarrow a_i \theta_i(k) + \sum_{j \in \mathcal{N}_i^*(k)} a_i x_j(k),$$

where $\mathcal{N}_i^*(k)$ is the set of in-neighbors that have not been removed from $r_i(k)$

end while

been removed. Clearly, $|\mathcal{N}_i^*(k)| = |\mathcal{N}_i(k)| - 2f$. Then, agent i updates its internal state as

$$\theta_i(k+1) = a_i \theta_i(k) + \sum_{j \in \mathcal{N}_i^*(k)} a_i x_j(k), \quad (7)$$

with $a_i = \frac{1}{|\mathcal{N}_i^*(k)|+1} = \frac{1}{|\mathcal{N}_i|-2f+1}$ (that is, the agent takes the average of its internal state $\theta_i(k)$ and the remaining messages in $r_i(k)$). Note that the a_i 's only depend on f and on the topology of the graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$.

5 Properties of the DP-MSR algorithm

We now investigate correctness, accuracy and differential privacy of Algorithm 1.

5.1 Correctness

We show that if the network topology is sufficiently robust, then Algorithm 1 is correct according to Definition 4. This is done by proving that: (i) each non-faulty agent $i \in \mathcal{V} \setminus \mathcal{F}$ (described by the stochastic process $\theta_i(k)$) converges in mean square to some common random variable θ_∞ , despite the presence of up to f faulty agents in the network (convergence condition); (ii) for each non-faulty agent $i \in \mathcal{V} \setminus \mathcal{F}$, $\mathbb{E}[\theta_i(k+1)]$ is the convex combination of $\{\mathbb{E}[\theta_j(k)]\}$, $j \in \mathcal{V} \setminus \mathcal{F}$, for $k \in \mathbb{Z}_{\geq 0}$ (validity condition).

Lemma 1 Consider the multi-agent system (6)-(7) and without loss of generality suppose that the

first $n - \phi$ agents are non-faulty and let $\theta(k) = [\theta_1(k), \dots, \theta_{n-\phi}(k)]^T$. Assume that the network topology is at least $(2f + 1)$ -robust, then the dynamics of the non-faulty agents can be expressed as

$$\theta_i(k+1) = M_i(k)\theta(k) + \sum_{j \in \mathcal{N}_i^*(k)} a_i \eta_j(k), \quad (8)$$

with $i \in \{1, \dots, n - \phi\}$ and $M_i(k)$ being a row-stochastic vector of size $(n - \phi)$ having the following properties: (1) the i -th element of $M_i(k)$, say $M_{ii}(k)$, is such that $M_{ii}(k) = a_i$; (2) the j -th element of $M_i(k)$, i.e. $M_{ij}(k)$ is non-zero only if $(j, i) \in \mathcal{E}$; (3) $\forall k$, at least $|\mathcal{N}_i \cap (\mathcal{V} \setminus \mathcal{F})| - f + 1$ elements of $M_i(k)$ are lower bounded by some positive constant.

Proof. This lemma can be proved by using arguments similar to those presented in (Vaidya, 2012; Su and Vaidya, 2017), where an analogous result was proved in the case where there is no noise in the network (i.e. $\eta_i = 0, \forall i = 1, \dots, n$). By combining (6) and (7) the closed loop dynamics of non-faulty agents can be written as

$$\begin{aligned} \theta_i(k+1) &= a_i \theta_i(k) + \sum_{j \in \mathcal{N}_i^*(k)} a_i (\theta_j(k) + \eta_j(k)) \\ &= \sum_{j \in \{i\} \cup \mathcal{N}_i^*(k)} a_i \theta_j(k) + \sum_{j \in \mathcal{N}_i^*(k)} a_i \eta_j(k). \end{aligned} \quad (9)$$

Note that the set $\mathcal{N}_i^*(k)$ in the above dynamics may contain values coming from faulty agents. Indeed, faulty agents might generate values between the f largest and the f smallest values received by a non-faulty agent and therefore those values would not be eliminated by the algorithm. In this case, as shown in details in (Vaidya, 2012), under the hypothesis that the network topology is $(2f + 1)$ -robust, it is always possible to express the value $\theta_j(k)$ from any faulty agent in $\mathcal{N}_i^*(k)$ as the convex combination of those of the non-faulty agents in $\mathcal{N}_i \setminus \mathcal{N}_i^*(k)$, that were removed “accidentally” by the algorithm. \square

Lemma 1 states that the time evolution of the non-faulty agent can be conveniently expressed as in (8) and, moreover, the noise-free term $M_i(k)\theta(k)$ depends only on the non-faulty agents.

Proposition 1 Consider the multi-agent system (6)-(7). Assume that the network topology is at least $(2f + 1)$ -robust. Then, the DP-MSR algorithm is correct and the closed loop dynamics (9) reaches mean square consensus, i.e. there exists a random variable θ_∞ such that $\lim_{k \rightarrow \infty} \mathbb{E} [|\theta_i(k) - \theta_\infty|^2] = 0$ for all $i \in \mathcal{V} \setminus \mathcal{F}$.

Proof. Without loss of generality, the attack signal (4) can be modeled as the second order random process

$$\hat{x}_i^j(k) = \hat{\theta}_i^j(k) + \hat{\eta}_i^j(k),$$

where $\hat{\theta}_i^j(k) : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}$ is some arbitrary deterministic signal, and $\hat{\eta}_i^j(k)$ is a zero-mean white noise with bounded variance $\hat{\sigma}_{i,j}^2(k)$, for all $k \geq 0$, and such that $\sum_{k=0}^{\infty} \hat{\sigma}_{i,j}^2(k) < \infty$.

Now, from Lemma 1 the evolution of a non-faulty agent can be expressed as in (8). Therefore, by taking the expected value we get $\mathbb{E} [\theta_i(k+1)] = M_i(k) \mathbb{E} [\theta(k)]$, $k \geq 0$. Since M_i is a stochastic vector, then the validity condition as in Definition 4 is satisfied for all $k \geq 0$.

We can prove convergence by means of Theorem 1 after noticing that (8) can be written in compact form as

$$\theta(k+1) = M(k)\theta(k) + \check{\eta}(k), \quad (10)$$

where $M(k)$ is a $(n - \phi) \times (n - \phi)$ row-stochastic matrix and

$$\check{\eta}(k) = \sum_{j \in \mathcal{N}_i^*(k) \setminus \mathcal{Q}_i(k)} a_i \eta_j(k) + \sum_{j \in \mathcal{Q}_i(k)} a_i \hat{\eta}_j^i(k), \quad (11)$$

with $\mathcal{Q}_i(k) = \mathcal{N}_i^*(k) \cap \mathcal{F}$ being the set of faulty agents in $\mathcal{N}_i^*(k)$ at time k . Indeed, observe that the dynamics (10) has the same form as (1), with $A(k) = M(k)$, $H(k) = I_{n-\phi}$, $v(k) = \check{\eta}(k)$ and that hypothesis (i) of such theorem is satisfied by construction (recall indeed that $\check{\eta}(k)$ is a sequence of independent random vectors with zero-mean and bounded covariance). We now show that hypotheses (ii) and (iii) of Theorem 1 are also satisfied by (10). To show that (ii) holds, we note that, in this case, $\sum_{k=0}^{\infty} \mathbb{E} [\text{tr}(H(k)^T H(k))] \mathbb{E} [v(k)^T v(k)] = (n - \phi) \sum_{k=0}^{\infty} \mathbb{E} [\check{\eta}^T(k) \check{\eta}(k)]$ and hence

$$\begin{aligned} &\sum_{k=0}^{\infty} \mathbb{E} [\check{\eta}^T(k) \check{\eta}(k)] \\ &\leq \sum_{k=0}^{\infty} \sum_{i=1}^{n-\phi} a_i^2 (|\mathcal{N}_i^*(k) \setminus \mathcal{Q}_i(k)| 2b^2(k) + |\mathcal{Q}_i(k)| \hat{\sigma}_{j,i}^2(k)) \\ &\leq \sum_{k=0}^{\infty} \sum_{i=1}^{n-\phi} a_i^2 |\mathcal{N}_i^*(k)| \max\{2b^2(k), \hat{\sigma}_{j,i}^2(k)\} \\ &\leq (n - f)^2 \sum_{k=0}^{\infty} \max\{2b^2(k), \hat{\sigma}_{j,i}^2(k)\} < \infty, \end{aligned}$$

where we used the fact that $0 < a_i \leq 1$, that the variance of the noise signals goes to zero, and that, for each agent i , the quantity $|\mathcal{N}_i^*(k)| = |\mathcal{N}_i| - 2f$ is constant for all $k \geq 0$ and such that $|\mathcal{N}_i^*(k)| \leq (n - 1) - 2f \leq n - f$. Now, hypothesis (iii) of Theorem 1 can be shown by noting that, since the network is $(2f + 1)$ -robust, then from (Vaidya, 2012, Theorem 1) we can immediately conclude

that the sequence $\{M(k), k \geq 0\}$ has ergodic backward product. Therefore, Theorem 1 holds and hence we can conclude that there exists a random variable θ_∞ such that $\lim_{k \rightarrow \infty} \mathbb{E} [\|\theta_i(k) - \theta_\infty\|^2] = 0$, for all $i \in \mathcal{V} \setminus \mathcal{F}$. This proves the result. \square

5.2 Accuracy

Next we analyze the statistical properties of θ_∞ , together with the accuracy of the algorithm.

Proposition 2 *Consider the multi-agent system (6)-(7). Assume that the network topology is at least $(3f+1)$ -robust, then the random variable θ_∞ towards which the state of non-faulty agents, $\theta_i(k)$, converges as $k \rightarrow \infty$ is such that $\mathbb{E}[\theta_\infty] = w^T \theta_0$, for some vector $w \in \mathbb{R}^n$ having nonnegative elements with $\mathbf{1}^T w = 1$, and*

$$\frac{2c^2 \min_i a_i^2}{n(1-q^2)} \leq \text{var}(\theta_\infty) \leq \frac{c^2(n-f)}{2(1-q^2)}. \quad (12)$$

Proof. Proposition 1 implies that $\lim_{k \rightarrow \infty} \mathbb{E}[\|\theta(k) - \theta_\infty \mathbf{1}\|^2] = 0$, which in turns implies that $\lim_{k \rightarrow \infty} \mathbb{E}[\theta(k)] = \mathbb{E}[\theta_\infty \mathbf{1}]$. Denoting with $\Psi_{k,s}$ the backward product of the sequence $\{M(k), k \geq 0\}$, from equation (10) we get

$$\theta(k) = \Psi_{k,0} \theta_0 + \sum_{h=0}^{k-1} \Psi_{k,h+1} \check{\eta}(h), \quad k \geq 0.$$

Now, from the fact that $\mathbb{E}[\check{\eta}(k)] = 0, \forall k \geq 0$, and since the sequence $\{M(k), k \geq 0\}$ has ergodic backward product, it follows that $\lim_{k \rightarrow \infty} \mathbb{E}[\theta(k)] = \lim_{k \rightarrow \infty} \Psi_{k,0} \theta_0 = \mathbf{1} w^T \theta_0$, for some $w \in \mathbb{R}^n$ such that $\mathbf{1}^T w = 1$ (Ren and Beard, 2005, Lemma 3.7). Hence, we have that $\mathbb{E}[\theta_\infty] = w^T \theta_0$. Now, since

$$\begin{aligned} \theta(k)^T \theta(k) &= \theta_0^T \Psi_{k,0}^T \Psi_{k,0} \theta_0 + 2 \sum_{h=0}^{k-1} \theta_0^T \Psi_{k,0}^T \Psi_{k,h+1} \check{\eta}(h) \\ &\quad + \sum_{h=0}^{k-1} \check{\eta}^T(h) \Psi_{k,h+1}^T \Psi_{k,h+1} \check{\eta}(h) \end{aligned}$$

and

$$\mathbb{E} \left[\sum_{h=0}^{k-1} \theta_0^T \Psi_{k,0}^T \Psi_{k,h+1} \check{\eta}(h) \right] = \mathbb{E} \left[\sum_{h=0}^{k-1} \zeta_h^T \check{\eta}(h) \right] = 0,$$

for some vectors $\zeta_h \in \mathbb{R}^n$, we have that

$$\begin{aligned} \text{var}(\theta(k)) &= \mathbb{E}[\theta^T(k) \theta(k)] - \mathbb{E}[\theta(k)]^T \mathbb{E}[\theta(k)] \\ &= \sum_{h=0}^{k-1} \mathbb{E}[\check{\eta}^T(h) \Psi_{k,h+1}^T \Psi_{k,h+1} \check{\eta}(h)] \\ &= \sum_{h=0}^{k-1} \sum_{i=1}^{n-\phi} \text{var}(\check{\eta}_i(h)) (\Psi_{k,h+1}^T \Psi_{k,h+1})_{ii}, \end{aligned}$$

having taken into account that $\check{\eta}_i$ are independent random variables. Now, from Proposition 1 it also follows that $\lim_{k \rightarrow \infty} \text{var}(\theta_i(k)) = \text{var}(\theta_\infty)$ and hence we have

$$\begin{aligned} \text{var}(\theta_\infty) &= \frac{1}{n-\phi} \lim_{k \rightarrow \infty} \text{var}(\theta(k)) \\ &= \frac{1}{n-\phi} \lim_{k \rightarrow \infty} \sum_{h=0}^{k-1} \sum_{i=1}^{n-\phi} \text{var}(\check{\eta}_i(h)) (\Psi_{k,h+1}^T \Psi_{k,h+1})_{ii}. \end{aligned} \quad (13)$$

Based on this, we now give upper and lower bounds on $\text{var}(\theta_\infty)$.

Upper bound: Consider the worst-case scenario where every faulty agent injects noise $\hat{\eta}_i^j(k)$. Without loss of generality, we can assume that

$$\hat{\sigma}_{i,j}^2(k) \leq 2b^2(k) = 2c^2 q^{2k}, \quad k \in \mathbb{Z}_{\geq 0},$$

i.e. $\hat{\sigma}_{i,j}^2(k)$ decays at least as the variance of the non-faulty agents in (6) and any difference from this is taken into account in $\hat{\theta}_j^i(k)$. From the hypothesis on the robustness of the network, we have that $|\mathcal{N}_i| \geq 3f+1$, and so $a_i = \frac{1}{|\mathcal{N}_i| - 2f + 1} \leq \frac{1}{f+2} < \frac{1}{2}$, for every node $i \in \mathcal{V} \setminus \mathcal{F}$. Hence, from (11) and recalling that $|\mathcal{N}_i^*(k)| \leq n-f$, we have that $\text{var}(\check{\eta}_i(h)) \leq (n-f) a_i^2 \hat{\sigma}_{j,i}^2(h) \leq \frac{1}{2} (n-f) c^2 q^{2h}$. Therefore, from (13) we get

$$\begin{aligned} \text{var}(\theta_\infty) &\leq \frac{n-f}{2(n-\phi)} \lim_{k \rightarrow \infty} \sum_{h=0}^{k-1} \sum_{i=1}^{n-\phi} c^2 q^{2h} (\Psi_{k,h+1}^T \Psi_{k,h+1})_{ii} \\ &\leq \frac{n-f}{2(n-\phi)} \lim_{k \rightarrow \infty} \sum_{h=0}^{k-1} c^2 q^{2h} \text{tr}(\Psi_{k,h+1}^T \Psi_{k,h+1}) \\ &\leq \frac{n-f}{2} \sum_{h=0}^{\infty} c^2 q^{2h} = \frac{c^2(n-f)}{2(1-q^2)}, \end{aligned}$$

where we used the fact that, for any $m \times m$ row-stochastic matrix Ψ , it holds that $\text{tr}(\Psi^T \Psi) = \text{tr}(\Psi \Psi^T) \leq m$.

Lower bound: In the case that the faulty agents do not add any noise to their signals, i.e. $\hat{\eta}_j^i(k) = 0$, from (11) we have that $\text{var}(\check{\eta}_i(h)) = |\mathcal{N}_i^*(h) \setminus \mathcal{Q}_i(h)| a_i^2 2b^2(h)$. Again, from the hypothesis on the robustness of the network we have that $|\mathcal{N}_i^*(h) \setminus \mathcal{Q}_i(h)| = |\mathcal{N}_i| - 2f - |\mathcal{Q}_i(h)| \geq$

$3f + 1 - 2f - \phi \geq 1$ for all i and $h \geq 0$. Hence, from (13) we get

$$\begin{aligned} \text{var}(\theta_\infty) &\geq \frac{1}{n - \phi} \lim_{k \rightarrow \infty} \sum_{h=0}^{k-1} \sum_{i=1}^{n-\phi} a_i^2 2b^2(h) (\Psi_{k,h+1}^T \Psi_{k,h+1})_{ii} \\ &\geq \frac{1}{n} \lim_{k \rightarrow \infty} \sum_{h=0}^{k-1} \left[\min_i (a_i^2) 2c^2 q^{2h} \right] \text{tr}(\Psi_{k,h+1}^T \Psi_{k,h+1}) \\ &\geq \frac{2c^2 \min_i a_i^2}{n} \sum_{h=0}^{\infty} q^{2h} = \frac{2c^2 \min_i a_i^2}{n(1-q^2)}, \end{aligned}$$

where we used the fact that, for any row-stochastic matrix Ψ , it holds that $\text{tr}(\Psi^T \Psi) = \|\Psi\|_F^2 \geq 1$. \square

Remark 1 Note that $(3f + 1)$ -robustness is required to obtain a nonzero lower bound in (12) independently on the network topology (the upper bound, instead, only requires $(2f + 1)$ -robustness). The fact that the network is $(3f + 1)$ -robust guarantees that, after Algorithm 1 removes $2f$ messages coming from in-neighbors, at least one of the remaining $f + 1$ messages comes from a noisy non-faulty agent in \mathcal{N}_i^* . On the other hand, even a slight relaxation of this condition, say the network is $3f$ -robust, would not allow to get a nonzero lower bound on $\text{var}(\theta_\infty)$ for every network topology. Indeed, for any $3f$ -robust network we know that $|\mathcal{N}_i| \geq 3f$, for all i , hence $|\mathcal{N}_i^*(h) \setminus \mathcal{Q}_i(h)| = |\mathcal{N}_i| - 2f - |\mathcal{Q}_i(h)| \geq 3f - 2f - \phi \geq f - f = 0$. Therefore, if the messages received by some non-faulty agent do not contain any noise, i.e. $\hat{\eta}_j^i(h) = 0$, $\forall i, j$, $h \geq 0$, then from (11) it follows that for some i and h it may occur that $\text{var}(\hat{\eta}_i(h)) = 0$. In other words, there exists a $3f$ -robust network such that the lower bound on $\text{var}(\theta_\infty)$ is zero.

Proposition 3 Under same assumptions of Proposition 2, Algorithm 1 is $\left(p, c\sqrt{\frac{(n-f)}{2p(1-q^2)}}\right)$ -accurate for any $p \in (0, 1)$.

Proof. From Chebyshev's inequality it follows that

$$\mathbb{P}\{|\theta_\infty - w^T \theta_0| \leq r\} \geq 1 - \frac{\text{var}(\theta_\infty)}{r^2} \geq 1 - \frac{\overline{\text{var}}(\theta_\infty)}{r^2},$$

where we denoted by $\overline{\text{var}}(\theta_\infty)$ the upper bound in (12). Now, by picking $p = \overline{\text{var}}(\theta_\infty)/r^2$ it follows that $r = \sqrt{\frac{\overline{\text{var}}(\theta_\infty)}{p}} = c\sqrt{\frac{(n-f)}{2p(1-q^2)}}$. \square

As in (Nozari et al., 2017), it follows that the ideal case of $(0, 0)$ -accuracy cannot be obtained because the radius r is a decreasing function of p .

5.3 Differential privacy

We now investigate the ε -differential privacy properties of the algorithm. To this aim, we consider two important special cases: (i) absence of faulty agents, i.e. $\phi = 0$; (ii) presence of $\phi \neq 0$ faulty agents that, similarly to (Sundaram and Hadjicostis, 2011), behave arbitrarily within the communication policy of the network, i.e. $\hat{\eta}_i^j(k) \sim \text{Lap}(\hat{b}(k))$, with $\hat{b}(k) = cq^k$. A key difference between the model considered here and the one in (Sundaram and Hadjicostis, 2011) is that in the latter it is assumed that faulty agents send the same value to their neighbors. In our case, omniscient faulty agents have the capability to adapt their attack signals to the initial conditions of non-faulty agents to drive the multi-agent system towards some malicious consensus value $\hat{\theta}_\infty$. Specifically, let $\theta_0^{(1)}$ and $\theta_0^{(2)}$ be two δ -adjacent initial conditions of non-faulty agents and let $\hat{\theta}_i^{j,(1)}(k)$ and $\hat{\theta}_i^{j,(2)}(k)$ be the corresponding input attack signals injected by the i -th faulty agent to steer the multi-agent system towards the same $\hat{\theta}_\infty$. Then, without loss of generality, the two input signals are such that

$$\hat{\theta}_i^{j,(2)}(k) = \hat{\theta}_i^{j,(1)}(k) + \hat{\delta}_i^j(k), \quad (14)$$

where $\hat{\delta}_i^j(k)$ converges to zero faster than the noise injected by the agents, that is $|\hat{\delta}_i^j(k)| \leq \bar{\delta} \lambda^k$, with $\bar{\delta} \geq 1$, $0 < \lambda < q$.

Proposition 4 Consider the multi-agent system (6)-(7) and let δ be the adjacency bound of the initial conditions of non-faulty agents. The following statements hold:

- (i) if $\phi = 0$, then $\bar{\varepsilon}$ -differential privacy of the agents is guaranteed, where $\bar{\varepsilon} = \delta \frac{2q}{c(2q-1)}$, $\frac{1}{2} < q < 1$;
- (ii) if $\phi \neq 0$, $\phi < f$, with $\hat{\eta}_i^j(k) \sim \text{Lap}(cq^k)$, for all i, j and $k \geq 0$, then ε -differential privacy is guaranteed for non-faulty agents, with $\varepsilon = \bar{\varepsilon} + \bar{\delta} \frac{f d_{\max}^{\text{out}} q}{c(q-\lambda)}$.

Proof. The proof follows similar steps and notation to those presented in (Nozari et al., 2017; Huang et al., 2012). Essentially, equation (5) is proved by establishing a bijective correspondence between any two executions $(\theta_0^{(1)}, \boldsymbol{\eta}^{(1)})$ and $(\theta_0^{(2)}, \boldsymbol{\eta}^{(2)})$ that generate the same observations (that is, such that $X_{\theta_0^{(1)}}(\boldsymbol{\eta}^{(1)}) = X_{\theta_0^{(2)}}(\boldsymbol{\eta}^{(2)})$) and then by evaluating the probability by integrating over all the possible executions starting from the same δ -adjacent initial conditions. Without loss of generality, assume that the agents from 1 to $n - \phi$ are non-faulty, and that the agents from $(n - \phi + 1)$ through n are faulty. At a generic time instant h the vector noise generated by all the agents is given as $\boldsymbol{\eta} = [\eta_1 \ \dots \ \eta_{n-\phi} \ \hat{\eta}_{n-\phi+1}^T \ \dots \ \hat{\eta}_n^T]^T$. That is, the first

$n - \phi$ components of η are the $\eta_i \in \mathbb{R}$ generated by non-faulty agents, while $\hat{\eta}_i = [\hat{\eta}_i^1 \dots \hat{\eta}_i^{d_i^{\text{out}}}]^T$, $i \in \mathcal{F}$, is the noise generated by the i -th faulty agent and received by each of its d_i^{out} out-neighbors (recall that faulty agents can send a different noise signal to each of their out-neighbors). Moreover, note that the vector $\eta \in \mathbb{R}^m$, where $m = (n - \phi) + \sum_{i=1}^{\phi} d_i^{\text{out}}$. Consider now any pair of δ -adjacent initial conditions $\theta_0^{(1)}$ and $\theta_0^{(2)}$ of the non-faulty agents and an arbitrary set $\mathcal{O} \subset (\mathbb{R}^m)^{\mathbb{N}}$ of observations. For any $k \geq 0$, let $R_k^{(\ell)} = \{\eta_k \in \Omega_k | X_{k, \theta_0^{(\ell)}}(\eta_k) \in \mathcal{O}_k\}$, $\ell = 1, 2$, where $\Omega_k = \mathbb{R}^{m(k+1)}$ being the sample space up to time k , and $\mathcal{O}_k \subseteq \mathbb{R}^{m(k+1)}$ being the set of observations obtained by truncating the elements of \mathcal{O} to finite subsequences of length $k + 1$. Hence, by the continuity of probability (Durrett, 2010, Theorem 1.1.1.iv)

$$\mathbb{P}\{\eta \in \Omega | X_{\theta_0^{(\ell)}}(\eta_k) \in \mathcal{O}\} = \lim_{k \rightarrow \infty} \int_{R_k^{(\ell)}} f_{m(k+1)}(\eta_k^{(\ell)}) d\eta_k^{(\ell)},$$

for $\ell = 1, 2$, where $f_{m(k+1)}$ is the $m(k+1)$ -dimensional joint probability distribution given as

$$\begin{aligned} f_{m(k+1)}(\eta_k) &= \prod_{h=0}^k \left[\prod_{i=1}^{n-\phi} \mathcal{L}(\eta_i(h); b(h)) \cdot \prod_{i=n-\phi+1}^n \prod_{j=1}^{d_i^{\text{out}}} \mathcal{L}(\hat{\eta}_i^j(h); \hat{b}(h)) \right] \\ &= \prod_{h=0}^k \prod_{i=1}^{n-\phi} \mathcal{L}(\eta_i(h); b(h)) \cdot \prod_{h=0}^k \prod_{i=n-\phi+1}^n \prod_{j=1}^{d_i^{\text{out}}} \mathcal{L}(\hat{\eta}_i^j(h); \hat{b}(h)) \\ &= f_{(n-\phi)(k+1)}(\eta_k) \cdot \hat{f}_{(m-n+\phi)(k+1)}(\eta_k). \end{aligned} \quad (15)$$

Now, without loss of generality, assume that $\theta_{0, i_0}^{(2)} = \theta_{0, i_0}^{(1)} + \delta_1$ for some $i_0 \in \{1, \dots, n - \phi\}$, where $0 \leq \delta_1 \leq \delta$, and $\theta_{0, i}^{(2)} = \theta_{0, i}^{(1)}$ for all $i \neq i_0$. Then, for any $\eta_k^{(1)} \in R_k^{(1)}$, define $\eta_k^{(2)}$ as

$$\eta_i^{(2)}(h) = \begin{cases} \eta_i^{(1)}(h) - a_i^h \delta_1, & \text{if } i = i_0, \\ \eta_i^{(1)}(h), & \text{if } i \neq i_0, \end{cases} \quad (16)$$

for $h \in \{0, \dots, k\}$. Then, from (6), for $i = i_0$ we have $x_{i_0}^{(2)}(0) = \theta_{0, i_0}^{(2)} + \eta_{i_0}^{(2)}(0) = \theta_{0, i_0}^{(1)} + \delta_1 + \eta_{i_0}^{(1)}(0) - \delta_1 = x_{i_0}^{(1)}(0)$, while for $i \neq i_0$ we have $x_i^{(2)}(0) = \theta_{0, i}^{(2)} + \eta_i^{(2)}(0) = \theta_{0, i}^{(1)} + \eta_i^{(1)}(0) = x_i^{(1)}(0)$. That is, (16) implies $x^{(2)}(0) = x^{(1)}(0)$. Since Algorithm 1 performs only deterministic operations on the data that receives as an input (in fact, it first sorts the values and then cuts the smallest f and largest f values), then $x^{(2)}(0) = x^{(1)}(0)$ implies $\mathcal{N}_i^{*,(2)}(0) = \mathcal{N}_i^{*,(1)}(0)$, for $i \in \{1, \dots, n - \phi\}$ (i.e. the

sets \mathcal{N}_i^* generated from the two δ -adjacent conditions are the same, because in both cases the same values are removed by the algorithm). Therefore, by induction, it can be easily shown that from (16) we get $\theta_i^{(2)}(h) = \theta_i^{(1)}(h) + a_i^h \delta_1$, if $i = i_0$, and $\theta_i^{(2)}(h) = \theta_i^{(1)}(h)$, if $i \neq i_0$, and also $x^{(2)}(h) = x^{(1)}(h)$, which implies $\mathcal{N}_i^{*,(2)}(h) = \mathcal{N}_i^{*,(1)}(h)$, for $i \in \{1, \dots, n - \phi\}$ and $h \in \{0, \dots, k\}$. That is, $X_{k, \theta_0^{(1)}}(\eta_k^{(1)}) = X_{k, \theta_0^{(2)}}(\eta_k^{(2)})$, so $\eta^{(2)} \in R_k^{(2)}$, and this defines a bijective correspondence between the two executions. Hence, for any $\eta_k^{(2)} \in R_k^{(2)}$ there exists a unique $(\eta_k^{(1)}, \Delta \eta_k) \in R_k^{(1)} \times \mathbb{R}^{(n-\phi)(k+1)}$ such that $\eta_k^{(2)} = \eta_k^{(1)} + \Delta \eta_k$, where $\Delta \eta_k$ is fixed. Therefore, using a change of variable as in (Nozari et al., 2017), we get $\mathbb{P}\{\eta \in \Omega | X_{\theta_0^{(2)}}(\eta) \in \mathcal{O}\} = \lim_{k \rightarrow \infty} \int_{R_k^{(1)}} f_{m(k+1)}(\eta_k^{(1)} + \Delta \eta_k) d\eta_k^{(1)}$.

Now, we first prove statement (i) of the proposition (i.e. $\phi = 0$) and then we prove statement (ii).

Proof of statement (i): In this case $\phi = 0$. This implies that $f_{m(k+1)}(\eta_k)$ defined in (15) is simply equal to $f_{(n-\phi)(k+1)}(\eta_k)$. Hence, the relation between the probability of any subsequence of an individual execution $\eta_k^{(1)}$ and its corresponding execution $\eta_k^{(2)}$ for a particular observation is given by

$$\begin{aligned} \frac{f_{(n-\phi)(k+1)}(\eta_k^{(1)})}{f_{(n-\phi)(k+1)}(\eta_k^{(1)} + \Delta \eta_k)} &= \frac{\prod_{h=0}^k \prod_{i=1}^{n-\phi} \mathcal{L}(\eta_i^{(1)}(h); b(h))}{\prod_{h=0}^k \prod_{i=1}^{n-\phi} \mathcal{L}(\eta_i^{(1)}(h) + \Delta \eta_i(h); b(h))} \\ &= \frac{\prod_{h=0}^k \mathcal{L}(\eta_{i_0}^{(1)}(h); b(h))}{\prod_{h=0}^k \mathcal{L}(\eta_{i_0}^{(1)}(h) + \Delta \eta_{i_0}(h); b(h))} \leq \prod_{h=0}^k e^{\frac{|\Delta \eta_{i_0}(h)|}{b(h)}} \\ &\leq e^{\sum_{h=0}^k \frac{a_{i_0}^h \delta}{c q^h}} \leq e^{\frac{\delta}{c} \sum_{h=0}^k (\frac{1}{2q})^h} \end{aligned} \quad (17)$$

where the last step follows from $a_i < 1/2$. Thus, we have

$$f_{(n-\phi)(k+1)}(\eta_k^{(1)}) \leq e^{\frac{\delta}{c} \sum_{h=0}^k (\frac{1}{2q})^h} f_{(n-\phi)(k+1)}(\eta_k^{(1)} + \Delta \eta_k). \quad (18)$$

Therefore, integrating both sides of (18) over all the executions $R_k^{(1)}$, taking the limit for $k \rightarrow \infty$ and considering that the geometric series in the exponent is convergent since from the hypotheses $\frac{1}{2} < q < 1$, we have

$$\mathbb{P}\{\eta \in \Omega | X_{\theta_0^{(1)}}(\eta) \in \mathcal{O}\} \leq e^{\frac{\delta}{c} \frac{2q}{(2q-1)}} \mathbb{P}\{\eta \in \Omega | X_{\theta_0^{(2)}}(\eta) \in \mathcal{O}\}.$$

The result is then proved by definition of differential privacy.

Proof of statement (ii): Now, in this case $\phi > 0$. For any $\eta_k^{(1)}$ define

$$\hat{\eta}_i^{j,(2)}(h) = \hat{\eta}_i^{j,(1)}(h) - \hat{\delta}_i^j(h), \quad (19)$$

for $h = \{0, \dots, k\}$. Similarly to what has been previously done, it can be shown that $x_i^{j,(2)}(0) = \hat{\theta}_i^{j,(2)}(0) + \hat{\eta}_i^{j,(2)}(0) = \hat{\theta}_i^{j,(1)}(0) + \hat{\delta}_i^j(0) + \hat{\eta}_i^{j,(1)}(0) - \hat{\delta}_i^j(0) = x_i^{j,(1)}(0)$. Therefore, also in this case, due to the deterministic operations carried out by Algorithm 1 on its input data, we have that $\mathcal{N}_i^{*,(2)}(0) = \mathcal{N}_i^{*,(1)}(0)$, for $i = \{1, \dots, n - \phi\}$, and hence, from (16) and (19), by induction, we have again $X_{k,\theta_0^{(1)}}(\eta_k^{(1)}) = X_{k,\theta_0^{(2)}}(\eta_k^{(2)})$. Thus, from (15), the relation between probabilities of any of these two corresponding executions is given by

$$\frac{f_{m(k+1)}(\eta_k^{(1)})}{f_{m(k+1)}(\eta_k^{(2)})} = \frac{f_{(n-\phi)(k+1)}(\eta_k^{(1)})}{f_{(n-\phi)(k+1)}(\eta_k^{(2)})} \cdot \frac{\hat{f}_{(m-n+\phi)(k+1)}(\eta_k^{(1)})}{\hat{f}_{(m-n+\phi)(k+1)}(\eta_k^{(2)})}, \quad (20)$$

where the first term is given in (17) and the second one, omitting $(m - n + \phi)(k + 1)$ in the subscripts, is

$$\begin{aligned} \frac{\hat{f}(\eta_k^{(1)})}{\hat{f}(\eta_k^{(1)} + \Delta \eta_k)} &= \frac{\prod_{h=0}^k \prod_{i=n-\phi+1}^n \prod_{j=1}^{d_i^{\text{out}}} \mathcal{L}(\hat{\eta}_i^{j,(1)}(h); \hat{b}(h))}{\prod_{h=0}^k \prod_{i=n-\phi+1}^n \prod_{j=1}^{d_i^{\text{out}}} \mathcal{L}(\hat{\eta}_i^{j,(1)}(h) + \Delta \hat{\eta}_i^j(h); \hat{b}(h))} \\ &\leq \prod_{h=0}^k \prod_{i=n-\phi+1}^n \prod_{j=1}^{d_i^{\text{out}}} e^{\frac{|\Delta \hat{\eta}_i^j(h)|}{\hat{b}(h)}} \leq \prod_{h=0}^k e^{\frac{\phi d_{\max}^{\text{out}} \bar{\delta} \lambda^h}{c q^h}} \leq e^{\frac{\bar{\delta} f d_{\max}^{\text{out}}}{c}} \sum_{h=0}^k \left(\frac{\lambda}{q}\right)^h \end{aligned}$$

where we used the bound in (14). Combining the above two relations and (17) we finally get

$$f_{m(k+1)}(\eta_k^{(1)}) \leq e^{\frac{\bar{\delta}}{c} \sum_{h=0}^k \left(\frac{\lambda}{2q}\right)^h + \frac{\bar{\delta} f d_{\max}^{\text{out}}}{c}} \sum_{h=0}^k \left(\frac{\lambda}{q}\right)^h f_{m(k+1)}(\eta_k^{(1)} + \Delta \eta_k).$$

In the same way as before, integrating over all the possible executions, using a change of variables and taking the limit, we finally get

$$\mathbb{P}\{\eta \in \Omega | X_{\theta_0^{(1)}}(\eta) \in \mathcal{O}\} \leq e^{\bar{\varepsilon} + \bar{\delta} \frac{f d_{\max}^{\text{out}}}{c(q-\lambda)}} \mathbb{P}\{\eta \in \Omega | X_{\theta_0^{(2)}}(\eta) \in \mathcal{O}\},$$

that establishes the desired result by definition of differential privacy. \square

Note that, if the convergence to zero of $\hat{\delta}_i^j(k)$ in (14) is in one-step, i.e. $\lambda = 0$, then $\varepsilon = \bar{\varepsilon} + \bar{\delta} \frac{f d_{\max}^{\text{out}}}{c}$. Moreover, if $\bar{\delta} = 0$, i.e. the faulty agents are not completely omniscient, then Proposition 4 implies that faulty agents do not perturb the differential privacy of the multi-agent system.

Remark 2 The major difference of Proposition 4 in the case $\phi = 0$ with respect to the results in (Nozari et al., 2017) is that we consider directed graphs whose topology changes over time, due to the operations performed by Algorithm 1. With respect to this, we note that the only feature of Algorithm 1 used in the previous proof is that the algorithm performs deterministic operations on its input data. Hence, the proof is valid for any deterministic algorithm.

6 Simulation example

We now report simulation results for a multi-agent system of $n = 25$ agents having $f = 1$ faulty agent (say, agent 1). The agents interact over a circulant digraph where each agent communicates with 8 agents ahead. It can be shown that such a graph is 4-robust, see (Usevitch and Panagou, 2017, Theorem 1). Initial conditions of the non-faulty agents are taken from the standard normal distribution and are kept the same across the simulations. At each time-step k , non-faulty agents update their internal state and send messages to their out-neighbors in accordance to the DP-MSR, while the faulty agent sends the message $\hat{x}_1^j(k) = 0.5 \sin(k) + \eta_1^j(k)$, where $j \in \mathcal{N}_1^{\text{out}}$ and $\eta_1^j(k) \sim \text{Lap}(\hat{c} \hat{q}^k)$, with $\hat{q} = 0.9$ and $\hat{c} = 0.8$. We set the parameters in Algorithm 1 to $c = 1$ and $q = 0.75$. Since the graph is 4-robust, then the DP-MSR algorithm is correct (Proposition 1) and the system reaches mean square consensus to some value of the random variable θ_∞ , whose expected value $\mathbb{E}[\theta_\infty]$ is contained between the smallest and the largest values of the initial conditions of the non-faulty agents. This prediction is confirmed in Figure 1. In Figure 2, the statistical distribution of θ_∞ is shown, which has been obtained by running 10^4 simulations. Further, we numerically found that $\text{var}(\theta_\infty) \approx 0.05$, in accordance with the theoretical bound defined in (12), that is $0.0037 \leq \text{var}(\theta_\infty) \leq 27.4286$.

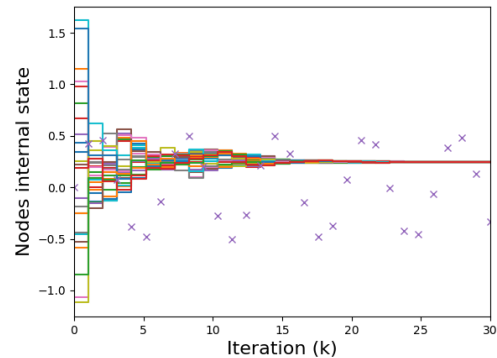


Fig. 1. Time evolution of the internal states of non-faulty agents for the multi-agent system in Section 6. The time evolution of the attack signal sent by the faulty agent is denoted by a cross.

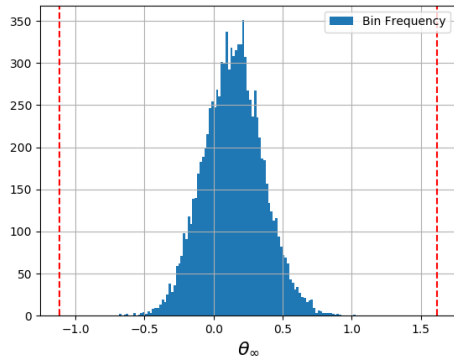


Fig. 2. Statistical distribution of θ_∞ . Dashed lines indicate the smallest and the largest initial conditions of non-faulty agents.

7 Conclusions

We considered multi-agent systems interacting over a directed network topology, where a subset of agents is faulty and where the non-faulty agents have the goal of achieving consensus and are subject to an additional differential privacy requirement. Our main contribution is the introduction of an algorithm, the DP-MSR algorithm, and the characterization of its correctness, accuracy and differential privacy properties. Simulations were used to illustrate the results. Future work will be aimed at extending the results presented in this paper to consider: (i) the analysis of differential privacy in more general scenarios, where the faulty agents do not respect the network communication policy; (ii) asynchronous systems; (iii) the design of algorithms to detect and isolate faulty agents.

Acknowledgements

The authors wish to thank Prof. Mario di Bernardo for reading this manuscript and for the comments he provided, Dr. Pietro De Lellis, Dr. Noise Holohan and Dr. Andrea Simonetto for the useful discussions on differential privacy and stochastic processes. The authors are also grateful to the anonymous reviewers and the AE for their constructive feedback.

References

Abbas, W., Laszka, A., Koutsoukos, X., 2018. Improving network connectivity and robustness using trusted nodes with application to resilient consensus. *IEEE Transactions on Control of Network Systems* 5, 2036–2048.

Albert, R., Jeong, H., Barabási, A.L., 2000. Error and attack tolerance of complex networks. *Nature* 406, 378.

di Bernardo, M., Fiore, D., Russo, G., Scafuti, F., 2016. Convergence, Consensus and Synchronization of Complex Networks via Contraction Theory, in: Lü, J., Yu, X., Chen, G., Yu, W. (Eds.), *Complex Systems and Networks. Understanding Complex Systems*. Springer, Berlin, Heidelberg, pp. 313–339.

Dibaji, S.M., Ishii, H., 2017. Resilient consensus of second-order agent networks: Asynchronous update rules with delays. *Automatica* 81, 123–132.

Dibaji, S.M., Ishii, H., Tempo, R., 2018. Resilient randomized quantized consensus. *IEEE Transactions on Automatic Control* 63, 2508–2522.

Dolev, D., 1982. The Byzantine generals strike again. *Journal of algorithms* 3, 14–30.

Dörfler, F., Bullo, F., 2014. Synchronization in complex networks of phase oscillators: A survey. *Automatica* 50, 1539–1564.

Durrett, R., 2010. *Probability: theory and examples*. 4th ed., Cambridge University Press.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., Naor, M., 2006a. Our data, ourselves: Privacy via distributed noise generation, in: *Eurocrypt*, Springer. pp. 486–503.

Dwork, C., McSherry, F., Nissim, K., Smith, A., 2006b. Calibrating noise to sensitivity in private data analysis, in: *Proc. of the 3rd Theory of Cryptography Conference*, pp. 265–284.

Fiore, D., Russo, G., di Bernardo, M., 2017. Exploiting nodes symmetries to control synchronization and consensus patterns in multiagent systems. *IEEE Control Systems Letters* 1, 364–369.

Griggs, W., Russo, G., Shorten, R., 2018. Leader and leaderless multi-layer consensus with state obfuscation: An application to distributed speed advisory systems. *IEEE Transactions on Intelligent Transportation Systems* 19, 711–721.

Han, S., Topcu, U., Pappas, G.J., 2014. Differentially private convex optimization with piecewise affine objectives, in: *Proc. of the 53rd IEEE Conference on Decision and Control*, pp. 2160–2166.

Huang, M., 2012. Stochastic approximation for consensus: a new approach via ergodic backward products. *IEEE Transactions on Automatic Control* 57, 2994–3008.

Huang, M., Dey, S., Nair, G.N., Manton, J.H., 2010. Stochastic consensus over noisy networks with Markovian and arbitrary switches. *Automatica* 46, 1571–1583.

Huang, M., Manton, J.H., 2009. Coordination and consensus of networked agents with noisy measurements: Stochastic algorithms and asymptotic behavior. *SIAM Journal on Control and Optimization* 48, 134–161.

Huang, M., Manton, J.H., 2010. Stochastic consensus seeking with noisy and directed inter-agent communication: Fixed and randomly varying topologies. *IEEE Transactions on Automatic Control* 55, 235–241.

Huang, Z., Mitra, S., Dullerud, G., 2012. Differentially private iterative synchronous consensus, in: *Proc. of*

- the 2012 ACM workshop on Privacy in the electronic society, pp. 81–90.
- Huang, Z., Mitra, S., Vaidya, N., 2015. Differentially private distributed optimization, in: Proc. of the 2015 International Conference on Distributed Computing and Networking, pp. 4:1–4:10.
- Kieckhafer, R., Azadmanesh, M., 1994. Reaching approximate agreement with mixed-mode faults. *IEEE Transactions on Parallel and Distributed Systems* 5, 53–63.
- Lamport, L., Shostak, R., Pease, M., 1982. The Byzantine generals problem. *ACM Transactions on Programming Languages and Systems* 4, 382–401.
- Le Ny, J., Pappas, G.J., 2014. Differentially private filtering. *IEEE Transactions on Automatic Control* 59, 341–354.
- LeBlanc, H.J., Koutsoukos, X.D., 2011. Consensus in networked multi-agent systems with adversaries, in: Proc. of the 14th ACM International Conference on Hybrid Systems: Computation and Control, pp. 281–290.
- LeBlanc, H.J., Koutsoukos, X.D., 2012. Low complexity resilient consensus in networked multi-agent systems with adversaries, in: Proc. of the 15th ACM International Conference on Hybrid Systems: Computation and Control, pp. 5–14.
- LeBlanc, H.J., Zhang, H., Koutsoukos, X., Sundaram, S., 2013. Resilient asymptotic consensus in robust networks. *IEEE Journal on Selected Areas in Communications* 31, 766–781.
- Mo, Y., Murray, R.M., 2017. Privacy preserving average consensus. *IEEE Transactions on Automatic Control* 62, 753–765.
- Monteil, J., Russo, G., 2017. On the design of nonlinear distributed control protocols for platooning systems. *IEEE Control Systems Letters* 1, 140–145.
- Nozari, E., Tallapragada, P., Cortés, J., 2017. Differentially private average consensus: obstructions, trade-offs, and optimal algorithm design. *Automatica* 81, 221–231.
- Nozari, E., Tallapragada, P., Cortes, J., 2018. Differentially private distributed convex optimization via functional perturbation. *IEEE Transactions on Control of Network Systems* 5, 395–408.
- Pasqualetti, F., Bicchi, A., Bullo, F., 2012. Consensus computation in unreliable networks: A system theoretic approach. *IEEE Transactions on Automatic Control* 57, 90–104.
- Ren, W., Beard, R.W., 2005. Consensus seeking in multi-agent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control* 50, 655–661.
- Senejohnny, D., Tesi, P., De Persis, C., 2018. A jamming-resilient algorithm for self-triggered network coordination. *IEEE Transactions on Control of Network Systems* 5, 981–990.
- Su, L., Vaidya, N.H., 2017. Reaching approximate byzantine consensus with multi-hop communication. *Information and Computation* 255, 352 – 368.
- Sundaram, S., Hadjicostis, C.N., 2011. Distributed function calculation via linear iterative strategies in the presence of malicious agents. *IEEE Transactions on Automatic Control* 56, 1495–1508.
- Tseng, L., Vaidya, N.H., 2016. A note on fault-tolerant consensus in directed networks. *ACM SIGACT News* 47, 70–91.
- Usevitch, J., Panagou, D., 2017. r-Robustness and (r,s)-Robustness of Circulant Graphs, in: Proc. of the 56th IEEE Conference on Decision and Control, pp. 4416–4421.
- Vaidya, N., 2012. Matrix representation of iterative approximate Byzantine consensus in directed graphs. *arXiv preprint arXiv:1203.1888*.
- Vaidya, N., Tseng, L., Liang, G., 2012. Iterative approximate Byzantine consensus in arbitrary directed graphs, in: Proceedings of the 2012 ACM Symposium on Principles of Distributed Computing, pp. 365–374.
- Zhang, H., Fata, E., Sundaram, S., 2015. A notion of robustness in complex networks. *IEEE Transactions on Control of Network Systems* 2, 310–320.
- Zhang, H., Sundaram, S., 2012. Robustness of information diffusion algorithms to locally bounded adversaries, in: Proc. of the American Control Conference, pp. 5855–5861.