2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

# A System-theoretic Privacy-informed Framework in Multi-agent Systems

Mahdieh S. Sadabadi

*Abstract*— **The problem of privacy-preserving in multi-agent systems corresponds to prohibiting the disclosure of agents' initial states while ensuring desired performance such as distributed average consensus. In this paper, a system-theoretic dynamic privacy-informed framework is developed. The proposed privacy framework relies on an obfuscation phase where a dynamic mask is inserted on the agents' state trajectories and masked outputs are exchanged amongst agents, rendering the physical states of an agent indiscernible by the other agents or external curious attackers (eavesdropper adversaries). Application of the proposed dynamic privacy scheme to well-known problems in multi-agent systems including average consensus and social opinion dynamics are presented.**

## I. Introduction

The cornerstone of most multi-agent systems is interaction and collaboration among multiple agents to achieve a certain common goal. The collaboration is achieved by information and data exchange through communication networks. The use of information and communication technologies (ICTs) and internet of things (IoT) schemes in multi-agent systems however comes at a price. One of the most notable issues is the loss of data privacy. This privacy issue is important particularly when agents carry sensitive and private information. The privacy-preserving of exchanged data is important in several applications such as smart energy systems where various private information such as power outputs, incremental costs, local sensitive grid operational information, and consumers' power usage profiles are usually exchanged amongst distributed generation units, network operators, and dispatch units [1]. The main question is how to preserve the privacy of the agents' information in a multi-agent system framework while ensuring a desired level of performance in reaching the common goal?

In order to address the privacy challenge and enhance the privacy-preserving feature in multi-agent systems, several approaches have been proposed in the literature. The common scheme is based on differential privacy mechanisms [2], which are the basis of most of existing privacy-related approaches reported in the literature., e.g., [3]–[6]. The differential privacy scheme usually relies on adding noises of appropriate statistical properties to exchanged data to attain the privacy guarantee. Nevertheless, corrupting the transmitted information by a large amount of injected noise in multi-agent systems might lead to the degradation of control performance and the accuracy of reaching an agreement amongst agents [6].

M. S. Sadabadi is with the Department of Electrical and Electronic Engineering, University of Manchester, Manchester, United Kingdom. mahdieh.sadabadi@manchester.ac.uk.

Another well-known approach for enhancing privacy preservation is based on data encryption approaches such as homomorphic encryption [7], [8], multi party or third party computation schemes [9], and distributed cryptography techniques [10]. In these approaches, the encrypted data are exchanged amongst agents by a communication graph and the cryptographic method is kept private by each agent. The main disadvantage of these methods is heavy computation/communication overhead [11].

Several studies have focused on the development of stochastic and deterministic privacy-preserving algorithms for the problem of average consensus in multi-agent systems where agents' initial states contain sensitive and private information. Examples of these approaches include the proposed differential privacy-preserving methods in [3], [4], gossip-based algorithms in [12], the deterministic privacy-preserving method in [13], the finite transmission event-triggered quantized average consensus algorithm in [14], the mask-based dynamic privacy-preserving strategy in [15], plausible deniability in [11], state decomposition-based approach in [16], and the privacy-preserving technique in [17] based on adding and subtracting random noises. These approaches either suffer from heavy computational burden or performance degradation in achieving a consensus on agents' initial values or are based on some restrictive assumptions on agents' knowledge about other agents' information or are only limited to multi-agent systems with discrete-time dynamics.

While the importance of the privacy-preserving of initial conditions is well-known in the context of average consensus, similar privacy concerns might exist for some other problems in a multi-agent framework such as continuous-time Friedkin–Johnsen model (FJ model) of opinion dynamics [15]. Such models describe how an opinion forms in a social network and the initial conditions of each agent describes each individual's initial opinion. Although the privacy of agents' initial opinion might be of importance, there are only a few methods that have considered the privacy-enhancing in FJ models. For instance, in [15], a system-theoretic algorithm for the privacy protection initial opinions in a continuous-time Friedkin–Johnsen model was developed.

This paper will address the privacy-preserving challenge in multi-agent systems by developing a system-theoretic privacy-informed framework. In particular, we focus on two well-known problems in the context of multi-agent systems: (i) average consensus and (ii) continuous-time Friedkin–Johnsen model. The proposed dynamic privacy framework relies on an obfuscation phase where a dynamic mask is

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

inserted on the agents' state trajectories and masked outputs are exchanged amongst agents, rendering the physical states of an agent indiscernible by the other agents or external curious attackers (eavesdropper adversaries). In the proposed context, the parameters of the dynamic mask applied to each agent is local, i.e., chosen by an individual agent; hence, improving the privacy of agents' states. The proposed privacy-informed framework in this paper is exact, i.e., the convergence to the exact value (the agreement among agents) is guaranteed.

The rest of the paper is organized as follows: Section II briefly explains the proposed privacy-informed framework in multi-agent systems. The applications of the proposed framework in average consensus and continuous-time Friedkin–Johnsen models are investigated in detail in Section III and Section IV, respectively. Section V presents numerical examples. Finally, the concluding remarks are given in Section VI.

*Notations.* Throughout this paper, $\mathbf{1}_n$ is an $n\times 1$ vector of ones, $\mathbf{0}_n$ is an $n\times 1$ zero vector, $\mathbf{I}_n$ is an $n\times n$ Identity matrix, and $\mathbf{0}_{n\times m}$ is a zero matrix of dimension $n\times m$. The symbols $X^T$, $X=[x_{i,j}]$, and $[x]$ respectively denote the transpose of matrix $X$, a matrix with entries $x_{i,j}$, and $[x]=\text{diag}(x_1,x_2,\ldots,x_n)$. For a symmetric matrix $X$, positive definite and positive semi-definite operators are respectively shown by $X\succ 0$ and $X\succeq 0$. We define $\mathbb{R}_+:=\{x\in\mathbb{R}\ |x>0\}$ and $\mathbb{R}_{\geq 0}:=\{x\in\mathbb{R}\ |x\geq 0\}$.

## II. Privacy-informed Framework

In a multi-agent system, agents usually aim to reach a collective agreement that requires communication and cooperation amongst agents. However, such communication and information exchanges might impose privacy concerns and disclose agents' sensitive and private information.

In order to enhance privacy in multi-agent systems, this paper will propose a system-theoretic privacy-informed control framework. The proposed dynamic control framework relies on an obfuscation phase where a dynamic mask is inserted on the agents' state trajectories $\mathbf{x}(t)\in\mathbb{R}^n$ and masked outputs are exchanged amongst neighbouring agents, rendering the physical states of an agent indiscernible by the other agents or external curious attackers (eavesdropper adversaries). By the term "dynamic", we mean that the mask is not static and has dynamics. In general, the masked outputs $\mathbf{y}(t)\in\mathbb{R}^n$ are defined as follows:

$$\mathbf{y}(t)=\mathbf{x}(t)+p(t,\mathbf{x}(t),\pi_1), \tag{1}$$

where $p(t,\mathbf{x}(t),\pi_1)\in\mathbb{R}^n$ is a continuous dynamic perturbation signal that has specific dynamics and $\pi_1$ is a set of the local parameters of agents.

This paper focuses on two specific problems in multi-agent systems including average consensus and the model of social opinion dynamics. In both problems, the main objective is to enhance the privacy of agents' initial states $\mathbf{x}(0)$ while reaching a collective agreement amongst agents.

## III. Average Consensus Problem

Distributed average consensus is one of the common problems in multi-agent systems, in which agents aim to reach an agreement on the average value of their initial states based on collaboration amongst agents [18]. The collaboration relies on communication and information exchanges amongst agents. The communication and information exchange however might disclose the information of agents' states and their initial values.

### A. Problem Statement

We consider a continuous-time multi-agent system consisting of $n$ ($n\geq 2$) agents whose dynamics are mathematically described as follows:

$$\dot{x}_i(t)=u_i(t), \tag{2}$$

for $i=1,\ldots,n$, where $x_i(t)\in\mathbb{R}$ is the state and $u_i(t)\in\mathbb{R}$ is the control input of agent $i$.

The information flow amongst agents is modeled by an undirected graph $\mathscr{G}=(\mathscr{V},\mathscr{E})$. The node set $\mathscr{V}=\{1,\ldots,n\}$ and the edge set $\mathscr{E}\subseteq\mathscr{V}\times\mathscr{V}$ represent agents and information exchange links, respectively. An edge $(i,j)\in\mathscr{E}$ indicates that node $j$ can receive information from node $i$ and vice versa. The adjacency matrix of the underlying graph is denoted by a matrix $\mathbb{A}$ with entries $a_{ij}$, where $a_{ii}=0$ and $a_{ij}=1$ if $(j,i)\in\mathscr{E}$; otherwise, $a_{ij}=0$. The following assumption is made on the communication graph $\mathscr{G}$:

**Assumption 1.** *The undirected graph is assumed to be connected.*

The main objective is to design a distributed average consensus algorithm for (2) so that all agents reach an average consensus, i.e.,

$$\lim_{t\to\infty}x_i(t)=x_{ave},\ \forall i\in\mathscr{V}. \tag{3}$$

where $x_{ave}=\frac{1}{n}\sum_{j=1}^{n}x_j(0)$.

For a connected undirected graph, the average consensus in (3) is achieved under the following updating rule [19]:

$$u_i(t)=-\sum_{j\in\mathscr{N}_i}a_{ij}\left(x_i(t)-x_j(t)\right),\forall t\geq 0, \tag{4}$$

for $i\in\mathscr{V}$, where $\mathscr{N}_i$ is the set of the neighboring agents of agent $i$ and $a_{ij}$ denotes the $(i,j)$-th element of the adjacency matrix $\mathbb{A}$. Note that $a_{ij}\neq 0$ if and only if $j\in\mathscr{N}_i$.

The conventional average consensus in (4) does not preserve the privacy of the initial value of agents' state, as it requires exchanging $x_j(t)$ to neighboring agents. This implies that an agent's privacy ($x_j(0)$) will be revealed to its neighbors and all agents' privacy will be revealed to external eavesdroppers.

The main objective of this section is to explore a privacy-informed average consensus algorithm that ensures the average consensus objective in (3) while simultaneously preserving the privacy of the individual agent's initial state.

### B. Privacy-informed Average Consensus Algorithm

The proposed privacy-informed average consensus algorithm in this paper relies on an obfuscation phase where a dynamic mask is used for agents' state trajectories $x_i(t)$. To this end, the proposed privacy-informed updating rule is given as follows:

$$u_i(t) = -\sum_{j\in\mathscr{N}_i} a_{ij}\left(y_i(t)-y_j(t)\right), \tag{5}$$

$$y_i(t) = x_i(t)+\beta_{1,i}h_i(t) \tag{6}$$

where $\beta_{1,i}\in\pi_1$ is a positive constant that is local to agent $i$ and $h_i(t)$ is an auxiliary state whose dynamics are defined as follows:

$$\dot{h}_i(t) = -\beta_{2,i}h_i(t)+\beta_{3,i}x_i(t);\ h_{i,0}=h_i(0)\neq 0, \tag{7}$$

for $i\in\mathscr{V}$, where $\beta_{2,i}\in\pi_1$ and $\beta_{3,i}\in\pi_1$ are positive scalars that are local to each agent and $h_{i,0}\in\pi_1$ is the initial condition of the auxiliary state $h_i(t)$ that is local to agent $i$. We define $\mathbf{x}(t)=[x_1(t),\ldots,x_n(t)]^T$ and $\mathbf{h}(t)=[h_1(t),\ldots,h_n(t)]^T$. Also, the value of $\beta_{2,i}$ is chosen to be equal to $\beta_{3,i}\beta_{1,i}$. The closed-loop dynamics under the proposed distributed average consensus control approach in (5)-(7) can be described in a vector format as follows:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= -\mathbf{L}\left(\underbrace{\mathbf{x}(t)+\beta_1\mathbf{h}(t)}_{\mathbf{y}(t)}\right),\ \mathbf{x}(0)\in\mathbb{R}^n\\ \dot{\mathbf{h}}(t) &= -\beta_2\mathbf{h}(t)+\beta_3\mathbf{x}(t),\quad \beta_2=\beta_3\beta_1\end{aligned} \tag{8}$$

where $\beta_1=\mathrm{diag}(\beta_{1,1},\ldots,\beta_{1,n})$, $\beta_2=\mathrm{diag}(\beta_{2,1},\ldots,\beta_{2,n})$, $\beta_3=\mathrm{diag}(\beta_{3,1},\ldots,\beta_{3,n})$, and $\mathbf{L}$ is the Laplacian matrix of the underlying communication graph. As one can observe from (5), each agent $i\in\mathscr{V}$ communicates $y_i(t)$ with its neighbors. As $h_i(t)$ are auxiliary variables, their initial condition can be set at any non-zero value. Moreover, as they are local to each agent (see the dynamics of $h_i(t)$ in (7)), their initial conditions are local to each agent. This helps in enhancing the privacy of the initial value of $x_i(0)$ from other agents.

Under Assumption 1, it can be shown that the state trajectories $\mathbf{x}(t)$ in (8) satisfy the following condition:

$$\mathbf{1}_n^T\mathbf{x}(t)=\mathbf{1}_n^T\mathbf{x}(0),\ \forall t>0. \tag{9}$$

The above equation is a conservation law for the dynamics of (8).

The following assumptions are made on an adversarial internal agent's knowledge and the topology of the communication graph $\mathscr{G}$.

**Assumption 2.** *(Knowledge set of the adversary): It is assumed that an adversarial internal agent $i\in\mathscr{V}$ has knowledge about (i) signals that it receives from its neighbors, i.e., $y_j(t)=x_j(t)+\beta_{1,j}h_j(t)$, $j\in\mathscr{N}_i$, $\forall t\geq 0$ and (ii) $(x_i(t),y_i(t))$, $\forall t\geq 0$.*

**Assumption 3.** *(Incapacity of the adversary's knowledge): It is assumed that an adversarial internal agent $i\in\mathscr{V}$ does not have knowledge about (i) the set $\pi_1$ (the value of $\beta_{1,j}$, $\beta_{3,j}$, and the initial value of $h_{j,0}$ for $j\neq i$), (ii) output trajectories $y_j(t)$, $j\notin\mathscr{N}_i$, and (ii) the adjacency matrix $\mathbb{A}$ of the network (network topology).*

**Assumption 4.** *(No overlapping neighborhood in $\mathscr{G}$): $\{\mathscr{N}_i\cup\{i\}\}\nsubseteq\{\mathscr{N}_j\cup\{j\}\}$, $\forall i,j\in\mathscr{V}$, $i\neq j$.*

Assumption 4 enforces a condition on the design of the topology of the communication graph $\mathscr{G}$, ensuring that node $i\in\mathscr{V}$ does not have complete information about other nodes in the graph. It is worth mentioning that Assumption 4 has been used in other privacy-related studies, e.g., [15] and [17].

The main purpose of the output masks $y_j(t)$ is to offset the agent's initial condition $(x_j(0))$ so that an eavesdropping attacker, either external or internal, cannot reconstruct it. Due to Assumption 3 (i), the problem of estimating $x_j(0)$ from the exchanged output $y_j(t)$ cannot be cast as a state observability problem. However, it can be considered as a joint system identification and state observability problem. In fact, detecting agents' initial states requires the identification of the dynamics of the auxiliary state $h_j(t)$ and the parameters of output masks (e.g., $\beta_{1,j}$) as well as a state observer. We then use the definition of "indiscernibility" in [15] that refers to the infeasibility of this joint identification and observability problem.

### C. Stability and Indiscernibility Analysis

It can be shown that the closed-loop dynamics in (8) has a continuum of equilibria in the form of $\bar{\mathbf{x}}=\mathbf{1}_n\alpha$ and $\bar{\mathbf{h}}=\beta_1^{-1}\mathbf{1}_n\alpha$, where $\alpha\in\mathbb{R}$ in which there is a unique equilibrium point $\bar{\mathbf{x}}=\mathbf{1}_n x_{ave}$ and $\bar{\mathbf{h}}=\beta_1^{-1}\mathbf{1}_n x_{ave}$ corresponding to every initial state value $\mathbf{x}(0)$, i.e., $\alpha=x_{ave}$.

**Lemma 1.** *All agents under the closed-loop dynamics in* (8) *globally asymptotically reach an average-consensus. Moreover, under Assumption 2-Assumption 4, the system in* (8) *guarantees the indiscernibility of the agents' initial condition* $\mathbf{x}(0)$.

*Proof:* Let $\bar{\mathbf{x}}=\mathbf{1}_n x_{ave}$ and $\bar{\mathbf{h}}=\beta_1^{-1}\mathbf{1}_n x_{ave}$ be an equilibrium of (8). We choose the following quadratic Lyapunov function:

$$\begin{aligned}V_1 &= \frac{1}{2}(\mathbf{x}(t)-\bar{\mathbf{x}})^T\mathbf{L}^+(\mathbf{x}(t)-\bar{\mathbf{x}})\\ &\quad+\frac{1}{2}(\mathbf{h}(t)-\bar{\mathbf{h}})^T\beta_1\beta_3^{-1}(\mathbf{h}(t)-\bar{\mathbf{h}}),\end{aligned} \tag{10}$$

where $\mathbf{L}^+$ is the generalized inverse of $\mathbf{L}$. Note that $V_1\geq 0$ and $(\beta_1\beta_3^{-1})^T=\beta_1\beta_3^{-1}$. Also, as $\mathbf{1}_n$ is in the kernel of $\mathbf{L}^+$, i.e., $\mathbf{L}^+\mathbf{1}_n=\mathbf{0}_n$, $V_1=0$ implies that $\mathbf{x}(t)-\bar{\mathbf{x}}=\mathbf{1}_n x^*$, where $x^*$ is a scalar. However, according to the conservation law in (9), it can be shown that $x^*=0$. As a result, $V_1=0$ if and only if $\mathbf{x}(t)=\bar{\mathbf{x}}$ and $\mathbf{h}(t)=\bar{\mathbf{h}}$.

The time derivative of $V_1$ in (10) along with the closed-loop trajectories in (8) is obtained as follows:

$$\dot{V}_1=-(\mathbf{x}(t)-\bar{\mathbf{x}})^T(\mathbf{x}(t)-\bar{\mathbf{x}})-(\mathbf{h}(t)-\bar{\mathbf{h}})^T\beta_1^2(\mathbf{h}(t)-\bar{\mathbf{h}}). \tag{11}$$

Note that in the derivation of $\dot{V}_1$ in (11), we have used the fact that $\mathbf{L}^+\mathbf{L}=\mathbf{I}_n-\frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ [20] and $\mathbf{1}_n^T(\mathbf{x}(t)-\bar{\mathbf{x}})=0$;

$\forall t \geq 0$. As $\dot{V}_1 < 0$, $\mathbf{x}(t)$ globally asymptotically converges to $\bar{\mathbf{x}} = \mathbf{1}_n x_{ave}$.

For the indiscernibility analysis, we consider the case where an eavesdropping agent $j$ aims to estimate $x_i(0)$, $i \neq j$. One approach is to identify the masked output $y_i(t)$ and the local parameter of the agent $i$. Nevertheless, based on Assumption 2, agent $j$ has only access to the set of $\mathscr{M}_{ij} = \{\mathscr{N}_i \cup \{i\}\} \cap \{\mathscr{N}_j \cup \{j\}\}$ of all output trajectories entering into the right hand side of (5). As a result, the system identification for (5) cannot be correctly carried out. Another approach for agent $j$ is to find $x_i(0)$ as follows:

$$x_i(0) = \bar{x}_i - \int_0^\infty (-\sum_{j \in \mathscr{N}_i} a_{ij}(y_i(t) - y_j(t))), \tag{12}$$

where $\bar{x}_i = x_{ave}$ is available to agent $j$. However, due to Assumption 4, agent $j$ cannot correctly estimate the integral in the right hand side of (12) (see Corollary 1 in [17]). As Assumption 4 holds for all agents, $x_i(0)$ is indiscernible.

## IV. Continuous-time Friedkin–Johnsen Model

Consider a continuous-time Friedkin–Johnsen model described as follows:

$$\dot{\mathbf{x}}(t) = -(\mathbf{L} + \Theta)\mathbf{x}(t) + \Theta\mathbf{x}(0), \tag{13}$$

where $\mathbf{L}$ is a Laplacian matrix and $\Theta = \text{diag}(\theta_1, ..., \theta_n)$, $\theta_i \in [0,1]$. As one can observer from (13), the conventional Friedkin–Johnsen model does not preserve the privacy of the initial opinions, i.e., $\mathbf{x}(0)$. Note that if $\Theta = \mathbf{0}_{n\times n}$, the FJ model reduces to the classical average consensus.

In the following, the conventional Friedkin–Johnsen model will be modified by adding a privacy-informed feature.

### A. Privacy-informed Friedkin–Johnsen Model

The proposed privacy-informed Friedkin–Johnsen model is based on the following updating rule:

$$\dot{\mathbf{x}}(t) = -(\mathbf{L} + \Theta)\mathbf{y}(t) + \Theta\mathbf{y}(\mathbf{x}(0)), \tag{14}$$

where

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{x}(t) + \beta_1 \mathbf{h}(t), \\ \dot{\mathbf{h}}(t) &= -\beta_2 \mathbf{h}(t) + \beta_3 \mathbf{x}(t), \end{aligned} \tag{15}$$

and $\beta_1 \in \mathbb{R}^{n\times n}$ and $\beta_3 \in \mathbb{R}^{n\times n}$ are diagonal matrices with positive diagonal elements, $\beta_2 = \beta_3\beta_1$, and $\mathbf{y}(0) = \mathbf{x}(0) + \beta_1\mathbf{h}(0)$, where $\mathbf{h}(0)$ is the initial values of $\mathbf{h}(t)$. In (15), the diagonal entities of $\beta_1$ and $\beta_3$, i.e., $\beta_{1,i} \in \pi_1$, and $\beta_{3,i} \in \pi_1$ are positive scalars local to each individual. Moreover, $h_i(0) \neq 0$ for $i = 1, \ldots, n$

### B. Stability and indiscernibility Analysis

If $\mathbf{L}$ is irreducible and there exists at-least one non-zero $\theta_i$, the privacy-informed Friedkin–Johnsen dynamic model in (14) and (15) has a unique equilibrium of $\bar{\mathbf{x}} = (\mathbf{L} + \Theta)^{-1}\Theta\mathbf{x}(0)$ and $\bar{\mathbf{h}} = \beta_3\beta_2^{-1}(\mathbf{L} + \Theta)^{-1}\Theta\mathbf{x}(0)$. Denoting $\hat{\mathbf{x}}(t) = \mathbf{x}(t) - \bar{\mathbf{x}}$ and $\hat{\mathbf{h}}(t) = \mathbf{h}(t) - \bar{\mathbf{h}}$, then (14) and (15) are expressed in the new coordinates as follows:

$$\begin{aligned} \dot{\hat{\mathbf{x}}}(\mathbf{t}) &= -(\mathbf{L} + \Theta)\left(\hat{\mathbf{x}}(t) + \beta_1\hat{\mathbf{h}}(t)\right), \\ \dot{\hat{\mathbf{h}}}(\mathbf{t}) &= -\beta_2\hat{\mathbf{h}}(t) + \beta_3\hat{\mathbf{x}}(t), \quad \beta_2 = \beta_3\beta_1 \end{aligned} \tag{16}$$

**Lemma 2.** *The origin of the closed-loop system in* (16) *is globally asymptotically stable.*

*Proof:* We consider the following quadratic Lyapunov function:

$$V_2 = \frac{1}{2}\hat{\mathbf{x}}^T(t)(\mathbf{L} + \Theta)^{-1}\hat{\mathbf{x}}(t) + \frac{1}{2}\hat{\mathbf{h}}^T(t)\beta_1\beta_3^{-1}\hat{\mathbf{h}}(t). \tag{17}$$

The time derivative of the proposed Lyapunov function $V_2$ is obtained as follows:

$$\dot{V}_2 = -\hat{\mathbf{x}}^T(t)\hat{\mathbf{x}}(t) - \hat{\mathbf{h}}^T(t)\beta_1^2\hat{\mathbf{h}}(t) \tag{18}$$

Since $\dot{V}_2 < 0$; $\forall\ \hat{\mathbf{x}}(t) \neq \mathbf{0}_n$ and $\hat{\mathbf{h}}(t) \neq \mathbf{0}_n$, the origin of the closed-loop system in (16) is globally asymptotically stable.

**Remark 1.** *Similar to the case of the average consensus in Section III-C, it can be shown that the proposed dynamics in* (14) *and* (15) *ensure the indiscernibility of agents' initial condition* $\mathbf{x}(0)$ *under Assumption 2-Assumption 4.*

## V. Numerical Examples

In this section, two numerical examples are presented to illustrate some key results of the proposed privacy-preserving techniques.

**Example 1.** *In this example, we consider the problem of average consensus in a multi-agent system composed of* $n = 5$ *agents. The communication topology amongst agents is described by an undirected graph with the following Laplacian matrix:*

$$\mathbf{L} = \begin{bmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 2 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 & -1 \\ 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & -1 & -1 & 2 \end{bmatrix}. \tag{19}$$

*The initial value of the states of agents are* $\mathbf{x}(0) = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix}^T$*; hence,* $x_{ave} = 3$*. The initial values of* $\mathbf{h}(t)$ *are randomly selected in the interval* $(0, 10]$*. The control parameters are chosen as follows:*

$$\beta_1 = \text{diag}(10,8,9,5,4),\ \beta_3 = \text{diag}(5,8,6,5,4),\ \beta_2 = \beta_3\beta_1.$$

*The graph topology based on the Laplacian matrix in* (19) *satisfies Assumption 4. Fig. 1 depicts the agents' physical state errors* $\mathbf{x}(t) - \frac{1}{n}\sum_{i=1}^n x_i(0)$*, masked outputs* $\mathbf{y}(t)$*, and auxiliary states* $\mathbf{h}(t)$*, with the proposed control strategy in (5). As one can observe from Fig. 1(b) and (c), the introduction of* $\mathbf{h}(t)$ *scrambles the initial conditions of agents' states; hence, it enhances the privacy-preserving feature in the average consensus problem.*

**Example 2.** *In this example, we consider a continuous-time Friedkin–Johnsen model of* $n = 34$ *agents. The initial values of the agents' state* $\mathbf{x}(0)$ *are chosen randomly in the interval* $[-3, 7]$*. The results shown in Fig. 2 indicate that while the masked outputs* $\mathbf{y}(t)$ *are exchanged among agents (see Fig. 2 (c))* $\mathbf{x}(t)$ *converges to* $\bar{\mathbf{x}} = (\mathbf{L} + \Theta)^{-1}\Theta\mathbf{x}(0)$ *(see Fig. 2 (a)).*

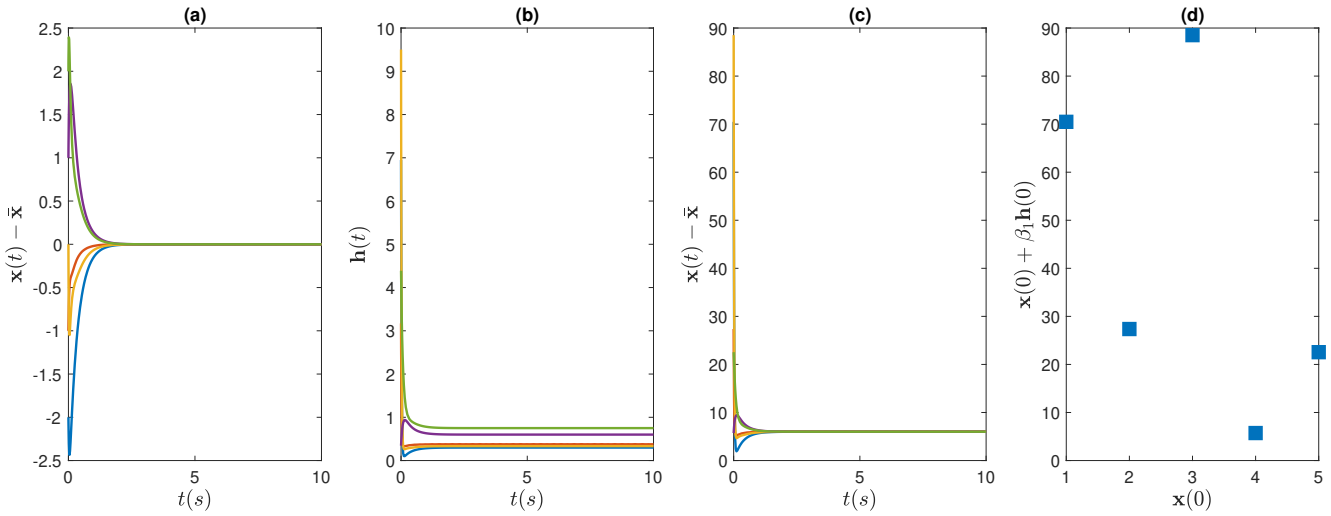


Fig. 1. Privacy-preserving average consensus in Example 1. (a) State error $\mathbf{x}(t)-\bar{\mathbf{x}}$; (b) auxiliary states $\mathbf{h}(t)$; (c) masked outputs $\mathbf{y}(t)$; (d): initial conditions $\mathbf{x}(0)$ vs. $\mathbf{x}(0)+\beta_1\mathbf{h}(0)$.

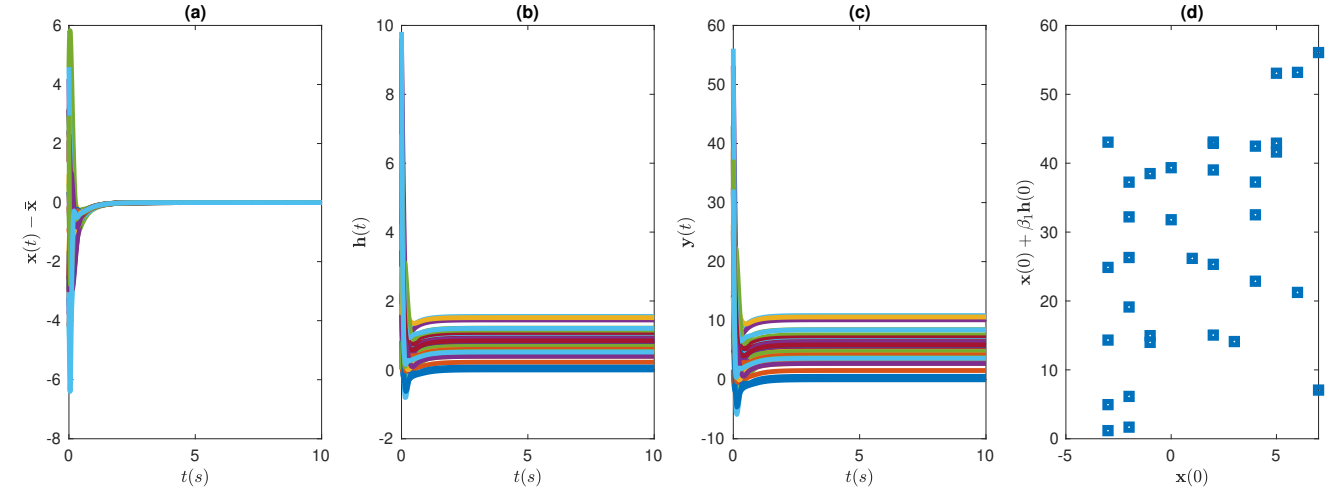


Fig. 2. Privacy-preserving continuous-time Friedkin–Johnsen model in Example 2. (a):State error $\mathbf{x}(t)-\bar{\mathbf{x}}$; (b) auxiliary states $\mathbf{h}(t)$; (c) masked outputs $\mathbf{y}(t)$; (d): initial conditions $\mathbf{x}(0)$ vs. $\mathbf{x}(0)+\beta_1\mathbf{h}(0)$.

## VI. Conclusion

This paper focuses on the enhancing of privacy in multi-agent systems with a particular emphasis on the problems of average consensus and continuous-time Friedkin–Johnsen model. A system-theoretic scheme is introduced which is based on introducing a dynamic mask whose dynamics and initial conditions are local to each agent. The proposed masks are inserted on the agents' states and masked data are exchanged among agents, boosting the privacy of the physical states of an agent from the other agents or eavesdropper adversaries. Simulation results are given to evaluate the performance of the proposed results in this paper.

## References


[1] M. S. Sadabadi, "Privacy-informed consensus-based secondary control in cyber-physical DC microgrids," *IEEE Control Systems Letters*, vol. 7, pp. 2089–2094, 2023.

[2] C. Dwork, "Differential privacy: A survey of results," in *International Conference on Theory and Applications of Models of Computation*, 2008, pp. 1–19.

[3] N. Gupta, J. Katz, and N. Chopra, "Privacy in distributed average consensus," in *20th IFAC World Congress*, vol. 50, no. 1, 2017, p. 9515–9520.

[4] E. Nozari, P. Tallapragada, and J. Cortés, "Differentially private average consensus: Obstructions, trade-offs, and optimal algorithm design," *Automatica*, vol. 81, p. 221–231, 2017.

[5] Y. Kawano and M. Cao, "Design of privacy-preserving dynamic controllers," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3863–3878, 2020.

[6] W. Zhang, Z. Zuo, Y. Wang, and G. Hu, "How much noise suffices for privacy of multi-agent systems?" *IEEE Transactions on Automatic Control*, pp. 1–16, 2022.

[7] Y. Shoukry, K. Gatsis, A. Alanwar, G. J. Pappas, S. A. Seshia, M. Srivastava, and P. Tabuada, "Privacy-aware quadratic optimization using partially homomorphic encryption," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, 2016, pp. 5053–5058.

[8] C. N. Hadjicostis and A. D. Domínguez-García, "Privacy-preserving distributed averaging via homomorphically encrypted ratio consensus," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3887–

3894, 2020.
[9] R. Lazzeretti, S. Horn, P. Braca, and P. Willett, "Secure multi-party consensus gossip algorithms," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 7406–7410.
[10] M. Ruan, H. Gao, and Y. Wang, "Secure and privacy-preserving consensus," *IEEE Transactions on Automatic Control*, vol. 64, no. 10, pp. 4035–4049, 2019.
[11] N. Monshizadeh and P. Tabuada, "Plausible deniability as a notion of privacy," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, Nice, France, Dec. 2019, pp. 1710–1715.
[12] N. E. Manitara, A. I. Rikos, and C. N. Hadjicostis, "Privacy-preserving distributed average consensus in finite time using random gossip," in *2022 European Control Conference (ECC)*, 2022, pp. 1282–1287.
[13] A.-S. Esteki and S. S. Kia, "Deterministic privacy preservation in static average consensus problem," *IEEE Control Systems Letters*, vol. 5, no. 6, pp. 2036–2041, 2021.
[14] A. I. Rikos, C. N. Hadjicostis, and K. H. Johansson, "Finite-time privacy-preserving quantized average consensus with transmission stopping," in *2022 IEEE 61st Conference on Decision and Control (CDC)*, 2022, pp. 6762–6768.
[15] C. Altafini, "A system-theoretic framework for privacy preservation in continuous-time multiagent dynamics," *Automatica*, vol. 122, p. 109253, Dec. 2020.
[16] Y. Wang, "Privacy-preserving average consensus via state decomposition," *IEEE Transactions on Automatic Control*, vol. 64, no. 11, pp. 4711–4716, 2019.
[17] Y. Mo and R. M. Murray, "Privacy preserving average consensus," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 753–765, 2017.
[18] M. S. Sadabadi and A. Gusrialdi, "Resilient average consensus on general directed graphs in presence of cyber-attacks," *European Journal of Control*, vol. 68, no. 100669, 2022.
[19] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
[20] P. Van Mieghem, K. Devriendt, and H. Cetinay, "Pseudoinverse of the Laplacian and best spreader node in a network," *Physical Review E*, vol. 96, no. 3, p. 032311, 2017.