Initial-Value Privacy of Linear Dynamical Systems

Lei Wang, Ian R. Manchester, Jochen Trumpf and Guodong Shi

Abstract— **This paper studies initial-value privacy problems of linear dynamical systems. We consider a standard linear time-invariant system with random process and measurement noises. For such a system, eavesdroppers having access to system output trajectories may infer the system initial states, leading to initial-value privacy risks. When a finite number of output trajectories are eavesdropped, we consider a requirement that any guess about the initial values can be plausibly denied. When an infinite number of output trajectories are eavesdropped, we consider a requirement that the initial values should not be uniquely recoverable. In view of these two privacy requirements, we define differential initial-value privacy and intrinsic initial-value privacy, respectively, for the system as metrics of privacy risks. First of all, we prove that the intrinsic initial-value privacy is equivalent to unobservability, while the differential initial-value privacy can be achieved for a privacy budget depending on an extended observability matrix of the system and the covariance of the noises. Next, the inherent network nature of the considered linear system is explored, where each individual state corresponds to a node and the state and output matrices induce interaction and sensing graphs, leading to a network system. Under this network system perspective, we allow the initial states at some nodes to be public, and investigate the resulting intrinsic initial-value privacy of each individual node. We establish necessary and sufficient conditions for such individual node initial-value privacy, and also prove that the intrinsic initial-value privacy of individual nodes is generically determined by the network structure.**

Index Terms— **Initial-Value Privacy; Differential Privacy; Identifiability; Observability**

I. Introduction

The rapid developments in networked control systems [1], internet of things [2], [3], smart grids [4], intelligent transportation [5], [6] during the past decade shed lights on how future infrastructures of our society can be made *smart* via interconnected sensing, dynamics, and control over cyber-physical systems. The operation of such systems inherently relies on users and subsystems sharing signals such as measurements, dynamical states, control inputs in their local views, so that collective decisions become possible. The shared signals might directly contain sensitive information of a private nature, e.g., loads and currents in a grid reflect directly activities in a residence or productions in a company [7]; or they might indirectly encode physical parameters, user preferences, economic inclination, e.g., control inputs in economic model predictive control implicitly carry information about the system's economic objective as it is used as the objective function [8].

Several privacy metrics have been developed to address privacy expectations of dynamical systems. A notable metric is differential privacy, which originated in computer science [9]–[11]. When a mechanism is applied taking the sensitive information as input and producing an output as the learning outcome, differential privacy guarantees plausible deniability of any inference about the private information by eavesdroppers having access to the output. Differential privacy has become the canonical solutions for privacy risk characterization in dataset processing, due to its quantitative nature and robustness to post-processing and side information [10], [11]. The differential privacy framework has also been generalized to dynamical systems for problems ranging from average consensus seeking [12] and distributed optimization [13], [14] to estimation and filtering [7], [15] and feedback control [16], [17]. Consistent with its root, differential privacy for a dynamical system provides the system with the ability to have plausible deniability facing eavesdroppers, e.g., recent surveys in [18], [19].

Besides differential privacy, another related but different privacy risk lies in the possibility that an eavesdropper makes an accurate enough estimation of the sensitive parameters or signals, perhaps from a number of repeated observations. In [20], the variance matrix of the maximum likelihood estimation was utilized to measure how accurate the initial node states in a consensus network maybe estimated from the trajectories of one or more malicious nodes. In [21], a measure of privacy was developed using the inverse of the trace of the Fisher information matrix, which is a lower bound of the variance of estimation error of unbiased estimators.

In particular, initial values of a dynamical system may carry sensitive private information, leading to privacy risks related to the initial values. For instance, when distributed load shedding in micro-grid systems is performed by employing an average consensus dynamics, initial values represent load demands of individual users [22]. In [12], [20], the initial-value privacy of the average consensus algorithm over dynamical networks was studied, and injecting exponentially decaying noises was used as a privacy-protection approach. The privacy of the initial value for a dynamical system is also of significant theoretical interest as the system trajectories or distributions of the system trajectories are fully parametrized by the initial value, in the presence of the plant knowledge. In [21], the initial-value privacy of a linear system was studied,

*This research is supported by the Australian Research Council Discovery Project DP190103615.

L. Wang, I. R. Manchester and G. Shi are with Australian Center for Field Robotics, The University of Sydney, Australia. (E-mail: lei.wang2; ian.manchester; guodong.shi@sydney.edu.au)

J. Trumpf is with College of Engineering and Computer Science, The Australian National University, Australia. (E-mail: Jochen.Trumpf@anu.edu.au)

and an optimal privacy-preserving policy was established for the probability density function of the additive noise such that the balance between the Fisher information-based privacy level and output performance is achieved.

In this paper, we study initial-value privacy problems of linear dynamical systems in the presence of random process and sensor noises. For such a system, eavesdroppers having access to system output trajectories may infer the system initial states. When a finite number of output trajectories are eavesdropped, we consider a requirement that any guess about the initial values can be plausibly denied. When an infinite number of output trajectories are eavesdropped, we consider a requirement that the initial values should not be uniquely recoverable. These requirements inspire us to define and investigate two initial-value privacy metrics for the considered linear system: differential initial-value privacy on the plausible deniability, and intrinsic initial-value privacy on the fundamental non-identifiability. Next, we turn to the inherent network nature of linear systems, where each dimension of the system state corresponds to a node, and the state and output matrices induce interaction and sensing graphs. In the presence of malicious users or additional observations, the initial states at a subset of the nodes may be known to the eavesdroppers as well. With such a public disclosure set, the intrinsic initial-value privacy of each individual node, and the structural privacy metric of the entire network, become interesting and challenging questions.

The main results of this paper are summarized in the following:

- For general linear systems, we prove that intrinsic initial-value privacy is equivalent to unobservability; and that differential initial-value privacy can be achieved for a privacy budget depending on an extended observability matrix of the system and the covariance of the noises.
- For networked linear systems, we establish necessary and sufficient conditions for intrinsic initial-value privacy of individual nodes, in the presence of a public disclosure set consisting of nodes with known initial states. We also show that the network structure plays a generic role in determining the privacy of each node's initial value, and the maximally allowed number of arbitrary disclosed nodes under privacy guarantee as a network privacy index.

The network privacy as proven to be a generic structural property, is a generalization to the classical structural observability results.

The remainder of the paper is organized as follows. Section 2 formulates the problem of interest for linear dynamical systems. In Section 3, intrinsic initial-value privacy and differential privacy are explicitly defined and studied by regarding all initial values as a whole. Then regarding the system from the network system perspective, Section 4 analyzes the intrinsic initial-value privacy of individual nodes with a public disclosure set and studies a quantitative network privacy index, from exact and generic perspectives. Finally a brief conclusion is made in Section 5. All technical proofs are omitted in this paper for page limitations, but are explicitly presented in the full version [23].

Notations. We denote by $\mathbb{R}$ the real numbers and $\mathbb{R}^n$ the real space of n dimension for any positive integer n. For a vector $\mathbf{x} \in \mathbb{R}^n$, the norm $\|\mathbf{x}\| = (\mathbf{x}^\top \mathbf{x})^{\frac{1}{2}}$. For any $\mathbf{x}_1, \ldots, \mathbf{x}_m \in \mathbb{R}^n$, we denote $[\mathbf{x}_1; \ldots; \mathbf{x}_m]$ as $\begin{bmatrix} \mathbf{x}_1^\top & \ldots & \mathbf{x}_m^\top \end{bmatrix}^\top \in \mathbb{R}^{mn}$, and $[\mathbf{x}_1, \ldots, \mathbf{x}_m]$ as a matrix of which the i-the column is $\mathbf{x}_i$, $i = 1, \ldots, m$. For any square matrix $\mathbf{A}$, let $\sigma(\mathbf{A})$ denote the set of all eigenvalues of $\mathbf{A}$, and $\sigma_M(\mathbf{A}), \sigma_m(\mathbf{A})$ denote the maximum and minimum eigenvalues, respectively. For any matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$, the norm $\|\mathbf{A}\| = \sigma_M(\mathbf{A}^\top \mathbf{A})^{\frac{1}{2}}$. The range of a matrix or a function is denoted as range$(\cdot)$, and the span of a matrix is denoted as span$(\cdot)$. We denote pdf$(\cdot)$ as the probability density function, and $\mathbf{e}_i \in \mathbb{R}^n$ as a vector whose entries are all zero except the i-th being one.

II. Problem Statement

A. Initial-Value Privacy for Linear Systems

We consider the following linear time-invariant system

$$\begin{aligned} \mathbf{x}_{t+1} &= \mathbf{A}\,\mathbf{x}_t + \boldsymbol{\nu}_t \\ \mathbf{y}_t &= \mathbf{C}\,\mathbf{x}_t + \boldsymbol{\omega}_t \end{aligned} \tag{1}$$

for $t = 0, 1, \ldots$, where $\mathbf{x}_t \in \mathbb{R}^n$ is state, $\mathbf{y}_t \in \mathbb{R}^m$ is output, $\boldsymbol{\nu}_t \in \mathbb{R}^n$ is process noise, and $\boldsymbol{\omega}_t \in \mathbb{R}^m$ is measurement noise. Throughout this paper, we assume that $\boldsymbol{\nu}_t$ and $\boldsymbol{\omega}_t$ are random variables according to some zero-mean distributions, and rank $(\mathbf{C}) > 0$.

In this paper, we suppose that initial values $\mathbf{x}_0$ are privacy-sensitive information for the system. Eavesdroppers having access to the output trajectory $(\mathbf{y}_t)_{t=0}^T$ with $T \geq n-1$ attempt to infer the initial values. To facilitate subsequent analysis, we denote the measurement vector $\mathbf{Y}_t = [\mathbf{y}_{T-t}; \mathbf{y}_{T-t+1}; \ldots; \mathbf{y}_T]$, the noise vectors $\mathbf{V}_t = [\boldsymbol{\nu}_{T-t}; \boldsymbol{\nu}_{T-t+1}; \ldots; \boldsymbol{\nu}_{T-1}]$ and $\mathbf{W}_t = [\boldsymbol{\omega}_{T-t}; \boldsymbol{\omega}_{T-t+1}; \ldots; \boldsymbol{\omega}_T]$, and let

$$\mathbf{O}_{ob} = \left[\mathbf{C}; \mathbf{CA}; \cdots; \mathbf{CA}^{n-1}\right], \ \mathbf{O}_t = \left[\mathbf{C}; \mathbf{CA}; \cdots; \mathbf{CA}^{t}\right],$$

$$\mathbf{H}_t = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \mathbf{C} & 0 & \ddots & 0 & 0 \\ \mathbf{CA} & \mathbf{C} & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{CA}^{t-2} & \mathbf{CA}^{t-3} & \ddots & \mathbf{C} & 0 \\ \mathbf{CA}^{t-1} & \mathbf{CA}^{t-2} & \cdots & \mathbf{CA} & \mathbf{C} \end{bmatrix}.$$

Here $\mathbf{O}_{ob}$ is observability matrix, and $\mathbf{O}_t$ denotes *extended* observability matrix for $t \geq n$ and $\mathbf{H}_t$ is a lower block triangular Toeplitz matrix. Thus, the mapping from initial state $\mathbf{x}_0$ to the output trajectory $\mathbf{Y}_T$ as $\mathcal{M} : \mathbb{R}^n \to \mathbb{R}^{m(T+1)}$ can be described by

$$\mathbf{Y}_T = \mathcal{M}(\mathbf{x}_0) := \mathbf{O}_T \mathbf{x}_0 + \mathbf{H}_T \mathbf{V}_T + \mathbf{W}_T\,. \tag{2}$$

The system (1) may be implemented or run independently for multiple times with the same initial state $\mathbf{x}_0$. When all resulting output trajectories are eavesdropped, the eavesdropper may derive an estimate of $\mathbf{x}_0$ by statistical inference methods such as maximum likelihood estimation (MLE).

The resulting estimate accuracy may converge to zero as the number of eavesdropped output trajectories converges to infinity, leading to initial-value privacy risks. In view of this, we consider a requirement that

(R1) the initial values should not be uniquely recoverable by an eavesdropper having an infinite number of output trajectories.

To address the requirement (R1), we define intrinsic initial-value privacy as below.

Definition 1: The system (1) preserves *intrinsic initial-value privacy* if the initial state $\mathbf{x}_0$ is statistically non-identifiable from observing $(\mathbf{y}_t)_{t=0}^T$, i.e., for any $\mathbf{x}_0 \in \mathbb{R}^n$, there exists a $\mathbf{x}_0' \neq \mathbf{x}_0 \in \mathbb{R}^n$ such that

$$\text{pdf}\left(\mathbf{Y}_T|\mathbf{x}_0\right) = \text{pdf}\left(\mathbf{Y}_T|\mathbf{x}_0'\right) . \tag{3}$$

In Definition 1, equality (3) indicates that there exist other values $\mathbf{x}_0'$, yielding the same output trajectory distribution as that of the initial value $\mathbf{x}_0$. This in turn guarantees that the system preserving the intrinsic initial-value privacy satisfies the requirement (R1).

Remark 1: The intrinsic initial-value privacy guarantees the initial state $\mathbf{x}_0$ indistinguishable from the $\mathbf{x}_0'$ satisfying (3), which is related to the notion of undetectable attacks in the secure control literature, e.g. [24], where the attacker tries to inject signals that are indistinguishable.

On the other hand, when a finite N output trajectories are eavesdropped, the eavesdroppers may infer the initial values under which there is a probability of generating these output trajectories. In view of this, we consider a requirement that

(R2) any inference about the true initial value from the eavesdroppers can be denied by supplying any value within a range to the inference with a similar probability of generating the eavesdropped N output trajectories.

This property is referred to as plausible deniability in the literature [25]. With this in mind, we denote the eavesdropped output trajectories as $\mathbf{Y}_T^1, \ldots, \mathbf{Y}_T^N$. The mapping from initial state $\mathbf{x}_0$ to $\mathbf{Y}_T^1, \ldots, \mathbf{Y}_T^N$ is a concatenation of N mappings $\mathcal{M}(\mathbf{x}_0)$, i.e.,

$$\begin{bmatrix} \mathbf{Y}_T^1 \\ \cdots \\ \mathbf{Y}_T^N \end{bmatrix} = \mathcal{M}^N(\mathbf{x}_0) := \begin{bmatrix} \mathcal{M}(\mathbf{x}_0) \\ \cdots \\ \mathcal{M}(\mathbf{x}_0) \end{bmatrix} . \tag{4}$$

We then define differential initial-value privacy as below [9], [10].

Definition 2: We define two initial values $\mathbf{x}_0, \mathbf{x}_0' \in \mathbb{R}^n$ as d-adjacent if $\|\mathbf{x}_0 - \mathbf{x}_0'\| \leq d$. The system (1) preserves (ϵ, δ)-*differential privacy of initial values* for some privacy budgets $\epsilon > 0, 0.5 > \delta > 0$ under d-adjacency if for all $R \subset \text{range}(\mathcal{M}^N)$,

$$\mathbb{P}(\mathcal{M}^N(\mathbf{x}_0) \in R) \leq e^\epsilon \cdot \mathbb{P}(\mathcal{M}^N(\mathbf{x}_0') \in R) + \delta \tag{5}$$

holds for any two d-adjacent initial values $\mathbf{x}_0, \mathbf{x}_0' \in \mathbb{R}^n$.

In Definition 2, inequality (5) indicates that the system can plausibly deny any guess from the eavesdroppers having N output trajectories, using any value from its d-adjacency. Namely, the system preserving the differential initial-value privacy satisfies the requirement (R2).

Remark 2: The requirement (R1) indeed can be understood from a perspective of denability. Namely,

(R1′) **[Deniability from non-identifiability]** any inference $\hat{\mathbf{x}}_0$ about the true initial value from the eavesdroppers having an infinite number of output trajectories, can be denied by supplying any other value $\mathbf{x}_0'$ satisfying

$$\text{pdf}\left(\mathbf{Y}_T|\hat{\mathbf{x}}_0\right) = \text{pdf}\left(\mathbf{Y}_T|\mathbf{x}_0'\right) . \tag{6}$$

Note that the derivation of such $\mathbf{x}_0'$ needs extra computation such that (6) is fulfilled and the resulting $\mathbf{x}_0'$ may be very close to or far away from the inference $\hat{\mathbf{x}}_0$. In contrast, the $\mathbf{x}_0'$ used to deny $\hat{\mathbf{x}}_0$ in (R2) is arbitrarily selected within a range to $\hat{\mathbf{x}}_0$. In view of this, the plausible deniability in (R2) provides the system with a more convenient denial mechanism. On the other hand, it can be seen that (6) indicates that (5) holds with $(\epsilon, \delta) = (0, 0)$, yielding that the eavesdroppers cannot distinguish between $\hat{\mathbf{x}}_0$ and $\mathbf{x}_0'$ with probability one. The above analysis thus demonstrates that intrinsic initial-value privacy and differential initial-value privacy are not inclusive mutually.

III. Initial-Value Privacy of General Linear Systems

In this section, both intrinsic and differential initial-value privacy of systems (1) are analyzed. We first present the following result on the equivalence of intrinsic initial-value privacy and unobservability.

Proposition 1: The system (1) preserves intrinsic initial-value privacy *if and only if* $(\mathbf{A}, \mathbf{C})$ is not observable, i.e., rank $(\mathbf{O}_{ob}) < n$.

Remark 3: Observability has been extensively studied in the fields of estimation [26] and feedback control [27]. In [16], for a linear control system the input observability is also explored to preserve differential privacy of control inputs and initial states. In Proposition 1, the intrinsic initial-value privacy and observability are bridged for linear systems (1).

Next, the differential privacy of initial values for (1) is studied. As in [15], we define $\mathcal{Q}(w) := \frac{1}{\sqrt{2\pi}} \int_w^\infty e^{-\frac{v^2}{2}} dv$, and $\kappa(\epsilon, \delta) := \frac{\mathcal{Q}^{-1}(\delta) + \sqrt{(\mathcal{Q}^{-1}(\delta))^2 + 2\epsilon}}{2\epsilon}$.

Theorem 1: Suppose that $(\mathbf{V}_T; \mathbf{W}_T)$ are random variables according to $(\mathbf{V}_T; \mathbf{W}_T) \backsim \mathcal{N}(0, \Sigma_T)$. Then the dynamical system (1) preserves (ϵ, δ)-differential privacy of initial state under d-adjacency, with $\epsilon > 0$ and $0.5 > \delta > 0$, if

$$\begin{aligned} \sigma_m\left(\begin{bmatrix}\mathbf{H}_T & \mathbf{I}_{m(T+1)}\end{bmatrix} \Sigma_T \begin{bmatrix}\mathbf{H}_T & \mathbf{I}_{m(T+1)}\end{bmatrix}^\top\right) \\ \geq d^2 N \|\mathbf{O}_T\|^2 \kappa(\epsilon, \delta)^2 . \end{aligned} \tag{7}$$

Remark 4: In Theorem 1, $\boldsymbol{\nu}_t, \boldsymbol{\omega}_t$ are assumed to admit Gaussian distributions. This renders the mapping (4) to be a Gaussian mechanism [10], [15], resulting in the (ϵ, δ)-differential privacy. One may wonder if assuming Laplacian noise $\boldsymbol{\nu}_t, \boldsymbol{\omega}_t$ would lead to a stronger $(\epsilon, 0)$-differential privacy, as in [14]. However, we note that the resulting mapping (4) is not a Laplace mechanism, because there is no guarantee that $\mathbf{H}_T\mathbf{V}_T + \mathbf{W}_T$ is still Laplacian, even if $\boldsymbol{\nu}_t, \boldsymbol{\omega}_t$ are Laplacian variables.

Remark 5: Though only initial values of system (1) are treated as private information, we note that the results in Theorem 1 can be extended to the case that all states $\mathbf{x}_t$ are sensitive. For dynamical systems (1), the outputs $\mathbf{y}_k$ for all $k \geq t$ in $\mathbf{Y}_T$ contain the information of $\mathbf{x}_t$, rendering a mapping from $\mathbf{x}_t$ to $\mathbf{Y}_{T-t}$ as

$$\mathbf{Y}_{T-t} = \mathcal{M}_t(\mathbf{x}_t) := \mathbf{O}_{T-t}\mathbf{x}_t + \mathbf{H}_{T-t}\mathbf{V}_{T-t} + \mathbf{W}_{T-t}\,.$$

By combining all $\mathcal{M}_t(\mathbf{x}_t)$, $t = 0, 1, \ldots, T$ together, one then can establish a mapping from the state trajectory $(\mathbf{x}_t)_{t=0}^T$ to the output trajectory $\mathbf{Y}_T$. For such a combined mapping, following the arguments in the proof of Theorem 1, one then can establish (ϵ, δ)-differential privacy of the state trajectory $(\mathbf{x}_t)_{t=0}^T$ with some $\epsilon > 0$ and $0.5 > \delta > 0$. In this way, our framework can be further applied to solve the problems in [15], [17], where the state trajectory is private information.

We note that given any covariance matrix $\Sigma_T > 0$, there always exist $\epsilon > 0$ and $0.5 > \delta > 0$, depending on the norm of extended observability matrix $\mathbf{O}_T$ such that (7) is satisfied, yielding the (ϵ, δ)-differential initial-value privacy. Thus, the (ϵ, δ)-differential initial-value privacy and the intrinsic initial-value privacy are mutually independent, with the latter determined by the unobservability of systems (1), i.e., rank $(\mathbf{O}_{ob}) < n$ by Proposition 1. This will be further explained in the subsequent Example 1.

If the noise can be designed, then there always exists a sufficiently large covariance matrix Σ_T such that (7) holds for any privacy budgets $\varepsilon > 0, 0.5 > \delta > 0$. To have a better view of this, we consider a particular case that $\boldsymbol{\nu}_t$ and $\boldsymbol{\omega}_t$ are i.i.d. random variables. The following corollary can be easily derived by verifying the condition (7).

Corollary 1: Suppose $\boldsymbol{\nu}_t$ and $\boldsymbol{\omega}_t$, $t = 0, 1, \ldots, T$ are i.i.d. random variables according to $\boldsymbol{\nu}_t \backsim \mathcal{N}(0, \sigma_{\boldsymbol{\nu}}^2 \mathbf{I}_n)$ and $\boldsymbol{\omega}_t \backsim \mathcal{N}(0, \sigma_{\boldsymbol{\omega}}^2 \mathbf{I}_m)$. Then for any $\epsilon > 0$, $0.5 > \delta > 0$, and all $\sigma_{\boldsymbol{\nu}} \geq 0$ and $\sigma_{\boldsymbol{\omega}} \geq d\sqrt{N}\|\mathbf{O}_T\|\kappa(\epsilon, \delta)$, the dynamical system (1) preserves (ϵ, δ)-differential privacy of initial state under d-adjacency.

Remark 6: Though in Corollary 1 arbitrary (ϵ, δ)-differential privacy can be achieved by choosing a sufficiently large $\sigma_{\boldsymbol{\omega}}$, this doesn't mean that the process noise $\boldsymbol{\nu}_t$ does not contribute to the differential privacy. In fact, simple calculations following the proof of Theorem 1 can lead to a less restrictive condition as

$$\|\mathbf{O}_T^\top(\sigma_{\boldsymbol{\nu}}^2 \mathbf{H}_T \mathbf{H}_T^\top + \sigma_{\boldsymbol{\omega}}^2 \mathbf{I}_m)^{-1}\mathbf{O}_T\| \leq \frac{1}{d^2 N \kappa(\epsilon, \delta)^2}\,,$$

from which it can be seen that $\sigma_{\boldsymbol{\nu}}$ also plays a role in achieving arbitrary differential privacy of initial values.

Example 1. Consider system (1) with $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}$. Let $T = 1$, and $\boldsymbol{\nu}_t, \boldsymbol{\omega}_t$ be i.i.d. random Gaussian noises with variances being $\sigma_{\boldsymbol{\nu}}^2 \mathbf{I}_2 > 0$ and $\sigma_{\boldsymbol{\omega}}^2 > 0$, respectively. In the following, we will respectively consider two kinds of outputs.

(a) Let output $\mathbf{y}_t = \mathbf{C}_1\mathbf{x}_t$ with $\mathbf{C}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$. We then obtain that $(\mathbf{A}, \mathbf{C}_1)$ is not observable as rank $(\mathbf{O}_T) = 1$. It is clear that the intrinsic initial-value privacy is preserved by Proposition 1. Regarding the differential privacy, one can find that $\|\mathbf{O}_T\| = 1$ and the differential initial-value privacy is preserved with some privacy budgets (ε, δ) by Theorem 1.

(b) Let output $\mathbf{y}_t = \mathbf{C}_2\mathbf{x}_t$ with $\mathbf{C}_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}$. We then obtain that $(\mathbf{A}, \mathbf{C}_2)$ is observable as rank $(\mathbf{O}_T) = 2$. By Proposition 1 the intrinsic initial-value privacy is disclosed, while the system preserves the differential initial-value privacy with some privacy budgets (ε, δ) by Theorem 1 as $\|\mathbf{O}_T\| = 1$.

Therefore, it can be seen that whether the system preserves the intrinsic initial-value privacy is independent of the differential initial-value privacy.

IV. Intrinsic Initial-Value Privacy of Networked Linear Systems

The system (1) can also be understood from a network system perspective, e.g., [28]. Let $x_{i,t}$ be the i-th entry of $\mathbf{x}_t$. If each $x_{i,t}$ is viewed as the dynamical state of a node, the matrix $\mathbf{A}$ would indicate a graph of interactions among the nodes. If each entry of $\mathbf{y}_t$ is viewed as the measurement of a sensor, then the matrix $\mathbf{C}$ would indicate a graph of interactions between the nodes and the sensors.

In view of this, we consider a network consisting of n network nodes and m sensing nodes, leading to a network node set $\mathrm{V} = \{1, \ldots, n\}$ and a sensing node set $\mathrm{V_S} = \{s_1, \ldots, s_m\}$ [1], respectively. Define the interaction graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$ with edge set $\mathrm{E} \subset \mathrm{V} \times \mathrm{V}$, and the sensing graph $\mathrm{G_S} = (\mathrm{V}, \mathrm{V_S}, \mathrm{E_S})$ with edge set $\mathrm{E_S} \subset \mathrm{V} \times \mathrm{V_S}$. Let $\mathbf{A} = [a_{ij}] \in \mathbb{R}^{n \times n}$ and $\mathbf{C} = [c_{ij}] \in \mathbb{R}^{m \times n}$.

To this end, this section aims to study how topological effects affect the privacy analysis of the networked system (1) with $(\mathbf{A}, \mathbf{C})$ being a configuration complying with the graphs $\mathrm{G}, \mathrm{G_S}$, i.e., if $a_{ij} = 0$ for $(j, i) \notin \mathrm{E}$ and $c_{ij} = 0$ for $(j, s_i) \notin \mathrm{E_S}$.

A. Intrinsic privacy of individual initial values

It is noted that in Definition 1 regarding intrinsic initial-value privacy, the initial state vector $\mathbf{x}_0$ is considered as a whole, and we suppose that the eavesdroppers have no prior knowledge of any individual initial values. In the following, we present several definitions that refine the notion in Definition 1 to dynamical networked systems (1) by studying intrinsic privacy of individual initial values, i.e., $x_{i,0}$, against eavesdroppers having knowledge of the whole sensor measurements (i.e., $\mathbf{y}_t$) and initial values of some network nodes. For convenience, we term the set of nodes whose initial values are prior knowledge to eavesdroppers as a *public disclosure set*.

Definition 3: For any given configuration $(\mathbf{A}, \mathbf{C})$ complying with graphs $\mathrm{G}, \mathrm{G_S}$, take $i \in \mathrm{V}$ and let $\mathrm{P} \subset \mathrm{V}$. The networked system (1) preserves *intrinsic initial-value privacy of node* i w.r.t. public disclosure set P if for any initial state $\mathbf{x}_0 \in \mathbb{R}^n$, there exists an $\mathbf{x}_0' = [x_{1,0}'; \ldots; x_{n,0}'] \in \mathbb{R}^n$ such that $x_{i,0} \neq x_{i,0}'$, $x_{j,0} = x_{j,0}'$ for all $j \in \mathrm{P}$, and

$$\mathrm{pdf}\,(\mathbf{Y}_T|\mathbf{x}_0) = \mathrm{pdf}\,(\mathbf{Y}_T|\mathbf{x}_0')\,. \tag{8}$$

[1] To be distinguished with notations for nodes in the interaction graph G, we use s_i to denote the i-th sensing node whose measurement is y_i.

Remark 7: The equality (8) indicates that even if the initial values of some nodes $j \in \mathrm{P}$ are public, the initial value $x_{i,0}$ cannot be identified from trajectories of $\mathbf{y}_t$, even with an infinite number of realizations of the dynamic networked system (1).

Remark 8: If the eavesdroppers have no prior knowledge of any node initial states, the above definition is also applicable with $\mathrm{P} = \emptyset$. In this case, according to Proposition 1, the notion of intrinsic initial-value privacy of node i is related to state variable unobservability of state $x_{i,t}$, that is a dual notion of state variable uncontrollability in [29].

Let $l = |\mathrm{P}|$ and $\mathrm{P} = \{p_1, \ldots, p_l\} \subset \mathrm{V}$. Define $\bar{\mathrm{P}} := \{\bar{p}_1, \ldots, \bar{p}_{n-l}\} = \mathrm{V} \backslash \mathrm{P}$. For convenience, we further let $\mathbf{E}_{\mathrm{P}} = [\mathbf{e}_{p_1}, \ldots, \mathbf{e}_{p_l}] \in \mathbb{R}^{n \times l}$ and $\mathbf{E}_{\bar{\mathrm{P}}} = [\mathbf{e}_{\bar{p}_1}, \ldots, \mathbf{e}_{\bar{p}_{n-l}}] \in \mathbb{R}^{n \times (n-l)}$, and $\mathbf{K}_j^{ob}$ be the j-the column of matrix $\mathbf{O}_{ob}$.

Theorem 2: Let the dynamical networked system (1) be equipped with configuration $(\mathbf{A}, \mathbf{C})$ complying with graphs $\mathrm{G}, \mathrm{G_S}$. Let $i \in \mathrm{V}$ and $\mathrm{P} \subset \mathrm{V}$ with $i \notin \mathrm{P}$. The following statements are equivalent.

a). The networked system (1) preserves *intrinsic initial-value privacy of node* i w.r.t. P.

b). rank $(\mathbf{O}_{ob}\mathbf{E}_{\bar{\mathrm{P}}})$ = rank $([\mathbf{K}_{i_1}^{ob}, \ldots, \mathbf{K}_{i_{n-l-1}}^{ob}])$ with $\{i_1, \ldots, i_{n-l-1}\} = \mathrm{V} \backslash (\mathrm{P} \cup \{i\})$.

c). rank $\left(\begin{bmatrix} \mathbf{O}_{ob} \\ \mathbf{e}_i^\top \end{bmatrix} \mathbf{E}_{\bar{\mathrm{P}}} \right)$ = rank $(\mathbf{O}_{ob}\mathbf{E}_{\bar{\mathrm{P}}}) + 1$.

In Theorem 2, explicit rank conditions are proposed to determine whether the intrinsic initial-value privacy of individual nodes is preserved, with respect to any given public disclosure set P. On the other hand, for a networked system, one may naturally ask what is the maximum allowable disclosure such that there always exists at least one node whose initial-value privacy is preserved. To address this issue, the network privacy index is introduced below.

Definition 4: The networked system (1) achieves level-l network privacy, if for any public disclosure set $\mathrm{P} \subset \mathrm{V}$ with $|\mathrm{P}| = l$, there exists a node $i \in \mathrm{V} \backslash \mathrm{P}$ whose intrinsic initial-value privacy is preserved w.r.t. P. The network privacy index of (1), denoted as $\mathbf{I_{rp}}$, is defined as the maximal value of l such that level-l relative privacy is achieved.

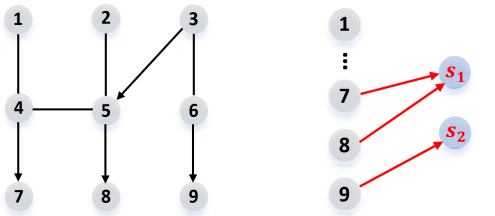
Proposition 2: The network privacy index of networked system (1) is $\mathbf{I}_{rp} = n - \text{rank}\,(\mathbf{O}_{ob}) - 1$.

Remark 9: By Definition 4, the full initial value is not disclosed irrespective of which $\mathbf{I}_{rp}$ nodes are public. It is clear that a larger $\mathbf{I}_{rp}$ means a stronger privacy-preservation ability of the networked system (1). According to Proposition 2, this further implies that a networked system possesses a better privacy-preservation ability, if the dimension of its unobservable subspace (i.e., $n - \text{rank}\,(\mathbf{O}_{ob})$) is higher.

Example 2. Consider a networked system (1) with $(\mathbf{A}, \mathbf{C})$ complying with the graphs $\mathrm{G}, \mathrm{G_S}$ in Figure 1, in which each edge is assigned with the same weight 1.

Ir can be seen that rank $(\mathbf{O}_{ob}) = 6$. According to Proposition 2, the network privacy index $\mathbf{I}_{rp} = 9 - \text{rank}\,(\mathbf{O}_{ob}) - 1 = 2$. Taking intrinsic privacy of individual initial values into consideration, simple calculations show that rank $([\mathbf{O}_{ob}; \mathbf{e}_j^\top]) = 7$ for $j = 1, 2, 4, 5, 7, 8$ and rank $([\mathbf{O}_{ob}; \mathbf{e}_j^\top]) = 6$ for $j = 3, 6, 9$. This indicates that the initial values of all nodes except nodes $3, 6, 9$ are private in the sense of Definition 3 using Theorem 2 with the public disclosure set $\mathrm{P} = \emptyset$.

Let the public disclosure set $\mathrm{P} = \{4, 8\}$. It can be verified that rank $(\mathbf{O}_{ob}\mathbf{E}_{\bar{\mathrm{P}}}) = 6$, and rank $\left(\begin{bmatrix} \mathbf{O}_{ob} \\ \mathbf{e}_j^\top \end{bmatrix} \mathbf{E}_{\bar{\mathrm{P}}} \right) = 7$ for $j = 1, 2$ and rank $\left(\begin{bmatrix} \mathbf{O}_{ob} \\ \mathbf{e}_j^\top \end{bmatrix} \mathbf{E}_{\bar{\mathrm{P}}} \right) = 6$ for $j = 3, 5, 6, 7, 9$. Thus, the intrinsic initial-value privacy of only nodes $1, 2$ is still preserved w.r.t. $\mathrm{P} = \{4, 8\}$. In fact, for any P with two elements, it can be verified that there always exist network nodes whose intrinsic initial-value privacy is preserved. On the other hand, if $\mathrm{P} = \{1, 4, 8\}$, one then can see that the intrinsic initial-value privacy of all nodes is disclosed. This in turn is consistent with the fact that the network privacy index $\mathbf{I}_{rp} = 2$.



(a) Interaction graph G (b) Sensing graph $\mathrm{G_S}$

Fig. 1. Network topologies.

B. Generic intrinsic initial-value privacy

In the previous subsection, the intrinsic initial-value privacy of individual nodes w.r.t. the public disclosure set P of networked systems (1) is studied, and a network privacy index $\mathbf{I}_{rp}$ is proposed to quantify the privacy of networked system (1). In the following, we turn to study the effect of network structure $(\mathrm{G}, \mathrm{G_S})$ to the intrinsic privacy and the network privacy index. To be precise, we demonstrate that these properties are indeed generic, i.e., are fulfilled for almost all edge weights under any network structure $(\mathrm{G}, \mathrm{G_S})$.

Theorem 3: Let $\mathrm{P} \subset \mathrm{V}$ and $i \in \mathrm{V}$. Then the intrinsic initial-value privacy of node i w.r.t. P is generically determined by the network topology. To be precise, exactly one of the following statements holds for any non-trivial network structure $(\mathrm{G}, \mathrm{G_S})$.

(i) The intrinsic initial-value privacy of node i is preserved generically, i.e., for almost all configurations $(\mathbf{A}, \mathbf{C})$ complying with the network structure $(\mathrm{G}, \mathrm{G_S})$.

(ii) The intrinsic initial-value privacy of node i is lost generically, i.e., for almost all configurations $(\mathbf{A}, \mathbf{C})$ complying with the network structure $(\mathrm{G}, \mathrm{G_S})$.

Theorem 3 demonstrates that given any network structure $(\mathrm{G}, \mathrm{G_S})$ and $\mathrm{P} \subset \mathrm{V}$, the intrinsic initial-value privacy of node i is either preserved or lost generically. We note that if there exists a configuration $(\mathbf{A}, \mathbf{C})$ complying with $(\mathrm{G}, \mathrm{G_S})$ such that the intrinsic initial-value privacy of node i is preserved (or lost), there is no guarantee that such property is preserved (or lost) generically. This is different from other common generic properties like structural controllability [32], for

which if there exists a configuration such that a linear system is controllable, then it must be controllable for almost all configurations, i.e., structurally controllable.

Similar to Theorem 3, the network privacy index is also generically determined by the network structure $(\mathrm{G}, \mathrm{G_S})$.

Theorem 4: The network privacy index is generically determined by the network topology. Namely, for almost all configurations $(\mathbf{A}, \mathbf{C})$ complying with the network structure $(\mathrm{G}, \mathrm{G_S})$, the network privacy index $\mathbf{I}_{rp} = n - n_{ob}^g - 1$ with n_{ob}^g given by the maximal rank of $\mathbf{O}_{ob}$.

Remark 10: We note that rank $(\mathbf{O}_{ob}) = n_{ob}^g$ holds for almost all configurations $(\mathbf{A}, \mathbf{C})$ complying with $(\mathrm{G}, \mathrm{G_S})$. According to [30] and the duality principle between controllability and observability, the n_{ob}^g indeed is given by the maximal number of edges in the set of stem-cycle disjoint graphs [29], [30].

V. Conclusions

In this paper, we have studied the intrinsic initial-value privacy and differential initial-value privacy of linear dynamical systems with random process and measurement noises. We proved that the intrinsic initial-value privacy is equivalent to unobservability, while the differential initial-value privacy can be achieved for a privacy budget depending on an extended observability matrix of the system and the covariance of the noises. Next, by regarding the considered linear system as a network system, we proposed necessary and sufficient conditions on the intrinsic initial-value privacy of individual nodes, in the presence of some nodes whose initial-value privacy is public. A quantitative network privacy index was also proposed using the largest number of arbitrary public nodes such that the whole initial values are not fully exposed. In addition, we showed that both the intrinsic initial-value privacy and the network privacy index are generically determined by the network structure. In future works, topological conditions (see e.g. [31], [32]) will be explored for generic intrinsic initial-value privacy of individual nodes, and the considered privacy metrics will be utilized to develop privacy-preservation approaches for linear dynamical systems.

References

[1] J. P. Hespanha, P. Naghshtabrizi, Y. Xu, "A survey of recent results in networked control systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 138-162, 2007.

[2] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer Networks*, vol. 54, no. 15, pp. 2787-2805, 2010.

[3] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645-1660, 2013.

[4] M. Kolhe, "Smart grid: Charting a new energy future: Research, development and demonstration," *The Electricity Journal*, vol. 25, pp. 88–93, 2012.

[5] P. Papadimitratos, A. D. La Fortelle, K. Evenssen, R. Brignolo and S. Cosenza, "Vehicular communication systems: Enabling technologies, applications, and future outlook on intelligent transportation," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 84-95, 2009.

[6] J. Zhang, F. Wang, K. Wang, W. Lin, X. Xu and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1624-1639, 2011.

[7] H. Sandberg, G. Dan, and R. Thobaben, "Differentially private state estimation in distribution networks with smart meters," in *Proc. 54th IEEE Conference on Decision and Control*, pp. 4492-4498, 2015.

[8] J. Ma, J. Qin, T. Salsbury, and P. Xu, "Demand reduction in building energy systems based on economic model predictive control," *Chemical Engineering Science*, vol. 67, no.1, pp. 92-100, 2012.

[9] C. Dwork, F. McSherry, K. Nissim, and A. Smith, "Calibrating noise to sensitivity in private data analysis," in *Proc. 3rd Theory of Cryptography Conference*, pp. 265-284, 2006.

[10] C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor, "Our data, ourselves: Privacy via distributed noise generation," *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 486-503, 2006.

[11] A. Roth, "New algorithms for preserving differential privacy," Ph.D., Carnegie Mellon Univ., Pittsburgh, PA, USA, 2010.

[12] Z. Huang, S. Mitra, and G. Dullerud, "Differentially private iterative synchronous consensus," in *Pro. 2012 ACM Workshop on Privacy in the Electronic Society*, pp. 81–90, ACM, 2012.

[13] S. Han, U. Topcu, and G. J. Pappas, "Differentially private distributed constrained optimization," *IEEE Transactions on Automatic Control*, vol. 62, no. 1, pp. 50–64, 2017.

[14] E. Nozari, P. Tallapragada, and J. Cortes, "Differentially private distributed convex optimization via functional perturbation," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 1, pp. 395-408, 2018.

[15] J. Le Ny and G. J. Pappas, "Differentially private filtering," *IEEE Transactions on Automatic Control*, vol. 59, no. 2, pp. 341–354, 2014.

[16] Y. Kawano and M. Cao, "Design of privacy-preserving dynamic controllers," *IEEE Transactions on Automatic Control*, DOI:10.1109/TAC.2020.2994030.

[17] M. Hale, A. Jones, K.Leahy, "Privacy in feedback: The differentially private LQG," in *Pro. IEEE American Control Conference*, pp. 3386-3391, 2018.

[18] S. Han and G. J. Pappas, "Privacy in control and dynamical systems," *Annual Review of Control, Robotics, and Autonomous Systems*, vol.1, pp. 309-332, 2018.

[19] J. Cortes, G. E. Dullerud, S. Han, J. Le Ny, S. Mitra, and G. J. Pappas, "Differential Privacy in control and network systems," in *Pro. 55th IEEE Conference on Decision and Control*, pp. 4252-4272, 2016.

[20] Y. Mo, R. M. Murray, "Privacy preserving average consensus," *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 753-765, 2017.

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

[22] Y. Xu, W. Liu, and J. Gong, "Stable multi-agent-based load shedding algorithm for power systems," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2006-2014, 2011.

[23] L. Wang, I. Manchester, J. Trumpf, and G. Shi, "Initial-Value Privacy of Linear Dynamical Systems," *arXiv:2008.10193*, 2020.

[24] F. Pasqualetti, F. Dorfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems" *IEEE Control Systems Magazin*, vol. 35, no. 1, pp. 110-127, 2015.

[25] C. Dwork, A. Roth. *The Algorithmic Foundations of Differential Privacy*. Foundations and Trends in Theoretical Computer Science, 2014.

[26] K. A. Clanents, G. R. Krutnpholz, P. W. Davis, "Power system state estimation with measurement deficiency: An observability/measurement placement algorithm," *IEEE Transactions on Power Apparatus and Systems*, vol.7, pp. 2012-2020, 1983.

[27] E. D. Sontag, *Mathematical Control Theory*, Texts in Applied Mathematics, 1998.

[28] M. Mesbahi, M. Egerstedt. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, 2010.

[29] L. Blackhall, and D. Hill, "On the structural controllability of networks of linear systems," in *Proc. 2nd IFAC Workshop on Distributed Estimation and Control in Networked Systems*, pp.245-250, 2010.

[30] S. Hosoe, "Determination of generic dimensions of controllable subspaces and its application," *IEEE Transactions on Automatic Control*, vol.25, no.6, pp.1192-1196, 1981.

[31] J. M. Hendrickx, M. Gevers, and A. S. Bazanella, "Identifiability of dynamical networks with partial node measurements," *IEEE Transactions on Automatic Control*, vol. 64, no. 6, pp. 2240–2253, 2019.

[32] C. T. Lin, "Structural controllability," *IEEE Transactions on Automatic Control*, vol. 19, no. 3, pp. 201-208, 1974.