2017 IEEE 56th Annual Conference on Decision and Control (CDC)
December 12-15, 2017, Melbourne, Australia

Differential Privacy for Dynamical Sensitive Data

Fragkiskos Koufogiannis and George J. Pappas

Abstract— **We introduce the problem of protecting the privacy of *time-varying* sensitive data using differential privacy. Contrary to prior work that considers fixed private data, we wish to design a privacy-preserving mechanism that, at each time and given the observations so far, keeps the *current* state of a dynamical system private. Our work protects dynamical systems from being tracked by an adversary by providing differentially private guarantees.**

Specifically, we propose a mechanism which adds artificial noise to (i) the input of the system and (ii) the measurements which are then published. In particular, two scenarios are considered: for a scalar dynamical system under ϵ–differential privacy, we derive a mechanism that, at each time, publishes the most accurate approximation of the current state while preserving privacy. Next, for a general linear system under (ϵ, δ)–differential privacy, we propose a Gaussian–based privacy–preserving mechanism with a quadratic cost.

I. Introduction

The use of individuals' sensitive data is critical for several systems such as collaborative recommendation systems, social interaction, and social welfare maximization. This sensitive data can either be fixed over time, such as the date of birth and the health record of an individual, or can vary over time, such as a user's current position and the state of a plant. In every case, the use of sensitive data raises privacy concerns.

In particular, for time–varying sensitive data, preserving the privacy of the *current* value of the private state of a dynamical system arises in settings including network–controlled systems and sensor networks [1]. In such settings, we need to block an inference attack on the sensitive data at the current time and not necessarily the whole trajectory. Specifically, *at each time step and given the responses published so far, we need to defend against an adversary that attempts to infer the current value of private data, i.e. the current state of the system.*

Typically, privacy concerns are mitigated by mechanisms that perturb data or add noise such that an adversary who observes the published output cannot extract sensitive information. Proposed frameworks that preserve the privacy of sensitive data include: (i) an information–theoretical approach [2], where the privacy is quantified by the mutual information between the published output and the private data, (ii) a game–theoretic approach e.g. [3], where the adversary plays a game with the mechanism, and (iii) differential privacy which uses a statistical approach to provide privacy guarantees against a general and powerful adversary. In this paper, we use differential privacy for its general adversary model, concreteness of privacy guarantees, and increased popularity. According to differential privacy as introduced in [4] and surveyed in [5], artificial noise is injected to the responses before they are published such that an adversary that observes the these responses cannot confidently discern variations of the private data. The proposed privacy–preserving mechanisms in the literature approximate a variety of functions such as the mean, the median, graph–theoretic quantities, optimization problems, and filtering. As explained next in more detail, these works propose a privacy–preserving version of a function that maps a given private input (e.g. a database, a vector, a signal, a graph) to a response (e.g. a scalar value, another signal, sequences of exchanged messages) and, thus, provide a *single* privacy guarantee.

Specifically, existing work in differential privacy assumes that the private data is given and fixed in time and proposes privacy–mechanisms for a wide spectrum of applications that range from database counting functions [6] and statistics of populations [7], to distributed optimization problems [8], [9], and filtering by [10]. For example, [10] considers a signal as a private data and constructs private approximations of a desired filter. In particular, although the private data is time–varying (it is a whole signal), the privacy guarantees protect the signal as a whole and, thus, the problem reduces to a static one where filtering is viewed as a function acting on signal spaces. Similarly, [11] proposes a privacy–preserving proxy to the consensus algorithm, where the private data is the agents' initial states and the output is all exchanged messages. Again, despite the consensus dynamics, the privacy guarantee is static; the proposed mechanism is a privacy–preserving map from the private initial states to the sequences of messages. In all of these works, the private data is assumed to be an external input to the mechanism and cannot be perturbed per se. As an implication, in dynamical phenomena, these works provide only "static" privacy guarantees, i.e. the mechanism that maps private data to the responses is differentially privacy. In practice, however, privacy needs may vary over time in one or more of the following aspects: (i) the private data itself may change over time, (ii) additional responses may be published, or (iii) the strength of privacy may be either revised at a later time. More concretely, consider an individual using a location–based service and, thus, reporting her GPS location. Such an individual may wish to protect her *current* location, while not

Authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, USA.

This work was supported in part by the TerraSwarm Research Center, one of six centers supported by the STARnet phase of the Focus Center Research Program (FCRP) a Semiconductor Research Corporation program sponsored by MARCO and DARPA and in part by NSF CNS-1505799 and the Intel-NSF Partnership for Cyber-Physical Systems Security and Privacy.

978-1-5090-2873-3/17/$31.00 ©2017 IEEE

worrying about revealing her past locations. From a different point of view, an adversary may wish to track the state of a dynamical system and decide when to deploy an attack. In each case, we need to provide privacy guarantees that explicitly protect the *current* state of the system. Motivated by such applications, this work introduces time–varying privacy guarantees; at each time, the mechanism publishes additional information and the private data evolves. Then, we wish to design a privacy–preserving mechanism such that an adversary who observes the so–far responses of the mechanism cannot confidently infer the *current* private data.

Our work deviates from the literature by considering *time–varying* differentially private guarantees. For an underlying dynamical system, we formulate and solve the problem of designing a mechanism that provides the following privacy guarantees. At each time step, an adversary that has observed the outputs of the mechanism so far cannot confidently infer the current state of the system. The time–varying sense of a privacy statement stems from the fact that the mechanism publishes new outputs with every step, thus, offering to the adversary additional knowledge. Moreover, the private data that needs to be protected is the current state which changes over time. On a technical note, we also allow the privacy level, i.e. the strength of the privacy guarantees, at each time step to vary as well; either increase or decrease at each time step. Our contributions are both conceptual and technical. Conceptually, we extend differential privacy for the case where the private is not a fixed quantity. Also, the proposed mechanism overcomes the problem of *"depletion of privacy budget"* by changing the private data itself and, thus, allows for infinite horizons while maintaining meaningful privacy guarantees and accuracy of the responses. Additionally, contrary to existing privacy–preserving mechanisms that inject noise only in the published responses, our mechanism consists of two noise sources: aside from corrupting the published responses with noise, the mechanism perturbs the private data itself. Regarding our technical contributions, we design a Gaussian–based privacy–preserving mechanism for a linear system that provides (ϵ, δ)–differential privacy. Additionally, for scalar system under a ϵ–differential privacy, we provide an efficient privacy–preserving mechanism that, at each time, publishes the most accurate but private approximation of the current state. Both mechanisms consist of two parts: the sensor part which misreports the current private data by publishing only noisy versions of it and is typically used in differential privacy literature, and the controller part which injects noise to the system and corrupts the private data itself.

The paper is structured as follows. Section II briefly revisits the notion differential privacy, formulates the problem of designing a dynamically private mechanism that protects the *current* state of a dynamical system, and provides a technical comparison to existing work. Section III-A considers the case of a linear system under (ϵ, δ)–differential privacy while Section III-B proposes a Laplace–based mechanism for a linear scalar system under ϵ–differential privacy. We conclude this work with Section IV which discusses future directions.

II. Problem Formulation

In this section, we introduce the problem of designing mechanisms that provide differentially private guarantees for time–varying private data. After motivating and informally formulating the general problem in Subsection II-A, Subsection II-B provides a brief reviews of the framework of differential privacy, and Subsection II-C derives a concrete formulation for linear systems.

A. Time–varying Private Data

As a motivating example, we consider a swarm of mobile agents collaboratively monitoring a quantity of interest —e.g. a target's position or a temperature field— and publishing an estimate of this quantity. Additionally, the agents themselves do not want to be tracked and, thus, have privacy needs for their *current* state —e.g. an adversary may try to localize and attack them. Since the agents' positions may be inferred from the published responses, we wish to design a privacy–preserving mechanism that publishes accurate information while guaranteeing the agents' privacy. Another example, considers a vehicle traveling on a highway segment and reporting its position for traffic–monitoring purposes. However, due to privacy concerns, the vehicle does not wish to be accurately localized on the highway at any time.

A key observation, to be exploited later, is that if the privacy requirements cannot be satisfied by solely perturbing the published responses, the agent noisily perturbs its private data. This observation deviates from the assumptions in the differential privacy literature where the private data are assumed given and fixed over time and the mechanism cannot tamper with them. In practice, although some private data such as health records cannot be altered, in several scenarios, private data including sensor locations, leadership tokens, and a dynamical system's state can be updated by a mechanism.

We will introduce our problem in its general form and, later, we will focus on linear instances of it. Formally, we consider a dynamical system with state x_t and open–loop dynamics $x_{t+1} = f(x_t, u_t)$. For each time $t \in \{1, \ldots, T\}$, we wish to publish the observations $y_t = g(x_t)$. However, due to privacy concerns, at time t, we wish to protect the privacy of the current state x_t by appropriately injecting noise. Importantly, the privacy constraints are time–varying; the data that needs to remain private is not always the same but it evolves with time. Moreover, the adversary's knowledge changes as additional observations are published and, thus, past noisy observations potentially can harm the privacy of the current private data. To this end, we wish to design a privacy–preserving mechanism such that, at time t, the mechanism that maps the current state x_t to the so–far published observations $\{y_s : s \in 1 : t\}$ is ϵ_t–differentially private. The sequence of privacy levels $[\epsilon_t]_{t=1}^T$ is assumed to be given.

Contrary to existing privacy–preserving mechanisms that only perturb the published responses, the approach proposed

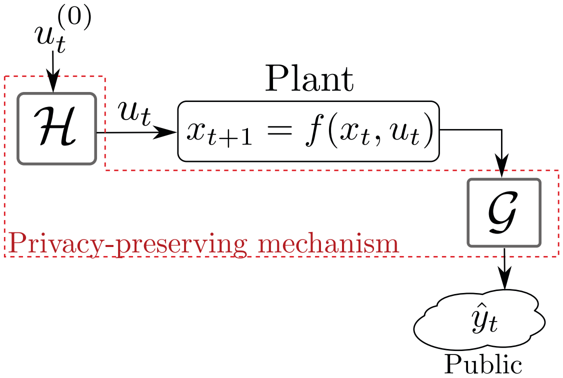

Fig. 1: We wish to design a privacy–preserving mechanism $(\mathcal{H}, \mathcal{G})$ such that, at time t and given the published observations $\{\hat{y}_i\}_{i=1}^t$, the current state x_t is ϵ_t–differentially private.

in this paper considers mechanisms that inject noise both in the sensor and the controller, as depicted in Figure 1.

i. *Sensor noise:* instead of the exact measurement y_t, the mechanism only publishes noisy versions of it $\hat{y}_t \sim \mathcal{G}(x_t)$; for example, $\hat{y}_t = y_t + V_t$, where V_t is suitable privacy-preserving noise. Intuitively, noise V_t protects the current state x_t from an adversary that knows the current observation $\hat{y}_t$. This type of noise is similar to the noise added by existing privacy–preserving mechanisms.

ii. *Controller noise:* the mechanism injects noise to the input of the system, $u_t \sim \mathcal{H}(u_t^{(0)})$, where $\mathcal{H}$ is a suitable privacy-preserving mechanism; for example $u_t = u_t^{(0)} + W_t$, where $u_t^{(0)}$ is an external control input —for simplicity we assume $u^{(0)} \equiv 0$— and W_t is appropriate noise. In words, if past observations $\hat{y}_1, \dots, \hat{y}_{t-1}$ can be used to accurately infer the next state x_t, then, the injected noise perturbs the system's state itself to enforce privacy.

Regarding performance of the system, we wish to minimize the amount of injected noise; increased sensor noise V_t renders the measurements $\hat{y}_t$ uninformative, whereas increased noise W_t changes the control input from the nominal one and, thus, degrades the performance of the plant.

B. Differential Privacy

Differential privacy, which was introduced in [12] and surveyed in [5], provides concrete privacy guarantees whenever private data is accessed. Specifically, artificial noise is injected such that a curious adversary that observes the output of a mechanism cannot confidently infer the initial private data. Differential privacy is defined as in Definition 1.

Definition 1 (Differential Privacy)**.** *Let $\epsilon > 0$ be a privacy level, $\mathcal{U}$ be the space of private data, and $\mathcal{A} \subseteq \mathcal{U} \times \mathcal{U}$ be an adjacency relation. The mechanism $Q : \mathcal{U} \to \Delta(\mathcal{Y})$ is ϵ-differentially private if:*

$$\mathbb{P}(Q(u) \in \mathcal{S}) \leq e^{\epsilon}\, \mathbb{P}(Q(u') \in \mathcal{S}) + \delta,$$

for all $\mathcal{S} \subseteq \mathcal{Y}$ and all adjacent inputs $(u, u') \in \mathcal{A}$.

The case of $\delta = 0$ is referred to as *ϵ–differential privacy*, whereas, the general case is termed *(ϵ, δ)–differential privacy*. Moreover, differential privacy is resilient to post-processing which allows the use of the Laplace and Gaussian mechanisms in constructing privacy-preserving mechanisms.

Theorem 2 (Laplace/Gaussian Mechanism)**.** *Consider the mechanism $Q : \mathcal{U} \to \Delta(\mathbb{R}^n)$ that adds noise to the result of query $q : \mathcal{U} \to \mathbb{R}^n$:*

$$Q(u) = q(u) + V.$$

Then,

- *Laplace mechanism: for $n = 1$, if $V \sim Lap\left(\frac{\|\Delta q\|_1}{\epsilon}\right)$, the mechanism Q is ϵ–differentially private;*
- *Gaussian mechanism: if $V \sim \mathcal{N}\left(0, \frac{\|\Delta q\|_2^2}{\kappa^2(\epsilon,\delta)}\right)$, the mechanism Q is (ϵ, δ)–differentially private;*

where $\|\Delta q\|_1 = \max_{(u,u') \in \mathcal{A}} \|q(u) - q(u')\|_1$, $\|\Delta q\|_2 = \max_{(u,u') \in \mathcal{A}} \|q(u) - q(u')\|_2$, Lap is the Laplace distribution, $\mathcal{N}$ is the normal distribution, $\kappa(\epsilon, \delta) = \frac{2\epsilon}{K + \sqrt{K^2 + 2\epsilon}}$, and $K = \mathcal{Q}^{-1}(\delta)$, where $\mathcal{Q}$ is the tail probability of the normal distribution.

As mentioned earlier, literature provides an impressive line of differentially private mechanisms. In these settings, the private data u is assumed to be a given input and cannot be altered by the mechanism. In such settings, noise in injected whenever a response y is published, such that privacy is preserved. Although the response is noisy and, thus, inaccurate, the private data remains unchanged. Instead, in our problem, we wish to protect the current state of a dynamical system where the mechanism can change the private data itself besides publishing responses corrupted with noise.

C. Differentially Private State

Finally, we concretely formulate the problem of designing a mechanism that, at each time, guarantees the privacy of the current state of a dynamical system. In this work, we consider linear dynamical systems

$$\begin{aligned} x_{t+1} &= A_t\, x_t + B_t\, u_t, \\ y_t &= C_t\, x_t \end{aligned}$$

Given a nominal input $u_t^{(0)}$, we wish to design a privacy-preserving mechanism $(\mathcal{H}, \mathcal{G})$ that sets $u_t = \mathcal{H}(u_t^{(0)})$ and $\hat{y}_t = \mathcal{G}(y_t)$ as depicted in Figure 1. Specfically, this mechanism is formulated in Problem 1.

Problem 1. *Given a sequence of privacy levels $[(\epsilon_t, \delta_t)]_{t=1}^T$, design a mechanism $(\mathcal{H}, \mathcal{G})$ such that*

- *Privacy: at time t, state x_t is (ϵ_t, δ_t)-private, i.e.*

$$\mathbb{P}(\hat{y}_1, \dots, \hat{y}_t | x_t) \leq e^{\epsilon_t}\, \mathbb{P}(\hat{y}_1, \dots, \hat{y}_t | x_t') + \delta_t,$$

for adjacent $(x_t, x_t') \in \mathcal{A}$;

- *Performance: the amount of injected noise is minimized*

$$\frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[C_1(0, u_t)\right] + \mathbb{E}\left[C_2(y_t, \hat{y}_t)\right],$$

where we assume that $u_t^{(0)} = 0$ and C_1 and C_2 are cost functions that penalize excessive noise.

In the rest of the paper, we provide privacy–preserving mechanisms for two instances of Problem 1: (i) a linear system with a finite horizon under (ϵ, δ)–privacy, where we prove that protecting the least–squares estimator suffices to provide differential privacy and (ii) a scalar system under ϵ–privacy, which models scenarios such as a car driving on a highway segment. Additionally, our results allow for time–varying privacy levels, i.e. for given sequences of (ϵ_t, δ_t), we require (ϵ_t, δ_t)–privacy at time time t.

III. Linear Systems with Private State

We now solve two instances of Problem 1. Subsection III-A considers a linear system under (ϵ, δ)–differentially private guarantees and, next, shows that our technique allows for time–varying privacy levels. Subsection III-B considers a scalar linear time–varying system under ϵ–differential privacy and provides a simple privacy–preserving mechanism that allows infinite horizon.

A. Linear System under (ϵ, δ)–Differential Privacy

We now design a Gaussian–based privacy–preserving mechanism that solves Problem 1 for a linear system in an LQG–like setting. Specifically, we consider the following linear, for simplicity time invariant, system.

$$x_{t+1} = A\, x_t + B\, u_t, \quad \text{and} \quad y_t = C\, x_t. \tag{1}$$

We assume that the system parameters A, B, and C are publicly known, that the system starts at $t = 0$, but the first observation published is y_1, and that (ϵ, δ) is a given privacy level. Our results remain applicable for time–varying privacy levels, i.e. a given sequence $[(\epsilon_t, \delta_t)_{t=1}^T$ of privacy levels, where T is a time horizon. Finally, we consider the adjacency relation $(x_t, x_t') \in \mathcal{A} \Leftrightarrow \|x_t - x_t'\|_2 \leq 1$ and the quadratic cost

$$C_T = \frac{1}{T}\left[\sum_{t=0}^{T-1} \mathbb{E}\,(u_t - u^{(0)})^T\, R\,(u_t - u^{(0)}) + \sum_{t=1}^{T} \mathbb{E}\,(\hat{y}_t - y_t)^T\, Q\,(\hat{y}_t - y_t)\right],$$

where we assume $R \succeq 0$ and $Q \succeq 0$ are positive semi–definite matrices that penalize control and output noise, respectively. Additionally, we assume that the nominal input $u_t^{(0)}$ is publicly known and, thus, we can ignore it by assuming $u_t^{(0)} = 0$. Since the input signal may be computed based on public information or be inferred by past executions of the system, we cannot argue about the privacy of the nominal input. Thus, following the dogma of differential privacy for a powerful adversary, we assume that this signal is publicly known.

In order to design a mechanism that, at time t, guarantees (ϵ, δ)–privacy of the current state x_t with respect to the adjacency relation $\mathcal{A}$, we design a privacy–preserving mechanism of the form

$$u_t = u_t^{(0)} + W_t = W_t, \quad \text{and} \quad \hat{y} = y_t + V_t. \tag{2}$$

Then, Problem 1 is stated as in Problem 2.

Problem 2. *Design the stochastic processes $[W_t]_{t=1}^T$ and $[V_t]_{t=1}^T$ such that the privacy–preserving mechanism that inputs $u_t = W_t$ and publishes $\hat{y}_t = y_t + V_t$ satisfies the following properties.*

- *At time t, the current state x_t is (ϵ, δ)–differentially private.*
- *The quadratic cost C_T is minimized; i.e. the processes W and V are not unnecessarily noisy.*

For this problem, we consider only zero–mean Gaussian–based privacy–preserving mechanisms of the form shown in Equation (2). Specifically, we design the covariance matrix

$$\begin{pmatrix} E_t \\ W_t \\ V_{t+1} \end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \mathbf{\Sigma}_t & 0 & \mathbf{X}_t \\ 0 & \mathbf{W}_t & \mathbf{Y}_t \\ \mathbf{X}_t^T & \mathbf{Y}_t^T & \mathbf{Z}_t \end{bmatrix}\right), \tag{3}$$

where $E_t = \hat{x}_t - x_t$ and $\hat{x}_t$ is the least–squares estimator of x_t, given the responses $\{\hat{y}_i\}_{i=1}^t$.

In the structure of the correlation matrix in Equation (3), we allow for correlation between the input and the output noise. However, we chose not to allow for any correlation between the input noise and the error of the least–squares estimator. These properties are similar to the mechanism presented in Subsection III-B for a scalar system under ϵ–differential privacy.

The covariance matrix in Equation (3) is derived from the following convex optimization problem.

$$\begin{aligned} \underset{\{\mathbf{\Sigma}_t, \mathbf{W}_t, \mathbf{Z}_t, \mathbf{X}_t, \mathbf{Y}_t\}_{t=1}^T}{\text{minimize}} \quad & \sum_{t=0}^{T-1} \left[\text{tr}(R\,\mathbf{W}_t) + \text{tr}(Q\,\mathbf{Z}_t)\right] \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{\Sigma}_t & 0 & \mathbf{X}_t \\ 0 & \mathbf{W}_t & \mathbf{Y}_t \\ \mathbf{X}_t^T & \mathbf{Y}_t^T & \mathbf{Z}_t \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} M_t - \mathbf{\Sigma}_{t+1} & N_t + M_t C^T \\ * & C M_t C^T + Z_t + \text{sym}(C\,N_t) \end{bmatrix} \succeq 0 \\ & \mathbf{\Sigma}_t \succeq \kappa^{-2}(\epsilon, \delta)\, I, \quad \forall t, \end{aligned} \tag{4}$$

where matrices M_t and N_t are linear functions of the variables defined in the proof of Theorem 3. The first constraint requires that the covariance block–matrix is well-defined, whereas the second constraint recursively couples the covariance matrices across different times. Lastly, the third inequality enforces the privacy constraint.

The following result proves that, for a feasible solution of program in Equation (4), the current state x_t is (ϵ, δ)–private.

Theorem 3. *Consider the linear system* (1) *with $u_t = W_t$ and $\hat{y}_t = y_t + V_t$ as defined in* (2) *and a privacy level $[(\epsilon, \delta)]_{t=1}^T$. If the covariance matrix satisfies the constraints of the optimization problem in* (4)*, then, at time t and given* *the observations $[\hat{y}_i]_{i=1}^{t}$, the current state x_t of the system is (ϵ, δ)–differentially private.*

The proof of this result is included in the Appendix. The derivation of the result follows the steps of Kalman estimation, in particular, as in [14] and, then, invoking the Gaussian mechanism from differential privacy. However, here we allow for correlation between the control noise, the sensor noise, and the estimation error and, thus, the exact expression is different. Specifically, we guarantee (ϵ, δ)–differential privacy for x_t at time t, if the least–squares estimator $\hat{x}_t$ can be written in the form of a Gaussian mechanism

$$\hat{x}_t = x_t + E_t,$$

where $\mathrm{Var}(E_t) \geq \kappa^2(\epsilon, \delta)\, I$.

The following result provides sufficient conditions for the feasibility of the optimization problem.

Proposition 4. *If the matrix $[A; B]$ has full row rank, then, the problem in Equation 4 is feasible.*

Sampling for the privacy–preserving noises $[W_t]_{t=0}^{T-1}$ and $[V_t]_{t=1}^{T}$ can be done as follows.

- The sensor part initializes the Kalman estimator by choosing $E_0 \sim \mathcal{N}(0, \mathbf{\Sigma}_0)$.
- At each time t, the controller part draws $W_t \sim \mathcal{N}(0, \mathbf{W}_t)$.
- At time t, the sensor part measures the state x_{t+1}, infers W_t and E_t, draws V_{t+1} conditioned on W_t and E_t, and publishes the response $\hat{y}_{t+1} = C\, x_{t+1} + V_{t+1}$.

Finally, we highlight that our technique allows for different privacy levels at each time step t, which captures the scenario where the privacy of part of the trajectory needs to be better protected. Specifically, given a sequence of privacy levels $[(\epsilon_t, \delta_t)]_{t=1}^{T}$, we can replace the last constraint in Equation (4) with the time–dependent expression

$$\mathbf{\Sigma}_t \succeq \kappa^{-2}(\epsilon_t, \delta_t)\, I, \quad \forall t \in \{1, \ldots, T\}.$$

B. Scalar System under ϵ–Differential Privacy

In this section, we consider a scalar system and ϵ–differential privacy and we provide a simple Laplace–based privacy–preserving mechanism that, at time t, protects the current state x_t with a privacy level ϵ_t. Specifically, we consider a noiseless scalar system with state $x_t \in \mathbb{R}$ and publicly known dynamics

$$x_{t+1} = a_t\, x_t + u_t \quad \text{and} \quad y_t = x_t, \tag{5}$$

a sequence of privacy levels $[\epsilon_t]_{t=1}^{T}$, where $T \in \mathbb{N} \cup \{\infty\}$ is a, possibly infinite, time horizon, and the adjacency relation $\mathcal{A}$ defined as

$$(x_t, x_t') \in \mathcal{A} \Leftrightarrow |x_t - x_t'| \leq 1.$$

At time t, the value of ϵ_t captures the strength of the privacy guarantees. Importantly, we do not make any assumptions on the monotonicity of the sequence of privacy levels and, thus, we allow for both privacy relaxation and tightening over time. For constant private data $x_t = x$, $\forall t$, the problem of *relaxing* privacy (increasing sequences of ϵ_t) has been explored in earlier work [13] but in the case of fixed private data, whereas, privacy tightening is conceptually impossible; once a response is published it is impossible to recall it. In our setting, we overcome this limitation by allowing for the privacy–preserving mechanism to noisily change the private data itself. Specifically, given the system in Equation (5), we consider a mechanism of the form

$$u_t = W_t \quad \text{and} \quad \hat{y}_t = y_t + V_t,$$

where $[W_t]_{t=1}^{T}$ and $[V_t]_{t=1}^{T}$ are appropriate privacy–preserving stochastic processes. As mentioned earlier, the input noise W_t changed the private data and, thus, at time $t+1$, we need to protect the new private data $a_t\, x_t + W_t$. Contrary to W_t which becomes part of the private data, the output noise V_t is logistic; essentially the mechanism misreports its state. Regarding accuracy, we consider a cost that penalizes inaccurate published data

$$C_T = \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\,(\hat{y}_t - y_t)^2 = \frac{1}{T}\sum_{t=1}^{T} \mathbb{E}\, V_t^2.$$

Then, Problem 1 takes the more specific form of Problem 3.

Problem 3. *Design the stochastic processes $[W_t]_{t=1}^{T}$ and $[V_t]_{t=1}^{T}$ such that*

- *for each time t and given the current state x_t, the mechanism that publishes $\{\hat{y}_i\}_{i=1}^{t}$ is ϵ_t–differentially private; and*
- *the published responses $\hat{y}_t$ accurately approximate the desired output x_t; i.e. minimizes the cost C_T.*

Theorem 5 solves Problem 3 and hints to an efficient algorithm that draws a sample from the stochastic processes W_t and V_t. In order to state Theorem 5, we define the following probability densities, where $\epsilon_2 \geq \epsilon_1 > 0$:

$$\ell_{\epsilon_1}(v) = \frac{\epsilon_1}{2} e^{-\epsilon_1 |v|},$$

$$\ell_{\epsilon_2|\epsilon_1}(v_2; v_1) = \left[\left(\frac{\epsilon_1}{\epsilon_2}\right)^2 \delta(v_1 - v_2) + \left(1 - \left(\frac{\epsilon_1}{\epsilon_2}\right)^2\right) \right.$$
$$\left. \ell_{\epsilon_1}(v_1 - v_2) \right] \frac{\ell_{\epsilon_2}(v_2)}{\ell_{\epsilon_1}(v_1)},$$

$$\ell_{\epsilon_1|\epsilon_2}(v) = \left(\frac{\epsilon_1}{\epsilon_2}\right)^2 \delta(v) + \left(1 - \left(\frac{\epsilon_1}{\epsilon_2}\right)^2\right) \ell_{\epsilon_1}(v),$$

where $\delta(\cdot)$ is Dirac's delta function. We refer the reader to earlier work [13] on the properties of these distributions. The proof of the following theorem, which proposes a mechanism and proves its privacy guarantees, can be found in the Appendix.

Theorem 5. *Given the sequence of privacy levels $[\epsilon_t]_{t=1}^{T}$, define the processes $[W_t]_{t=1}^{T}$ and $[V_t]_{t=1}^{T}$ such that $V_1 \sim \ell_{\epsilon_1}$ and, for $t \geq 2$,*

- *if* $\epsilon_t > |a_t|\,\epsilon_{t+1}$*, set*
$$W_t \sim \ell_{\epsilon_{t+1}|\frac{\epsilon_t}{|a_t|}} \quad \textit{and} \quad V_{t+1} = a_t\,V_t - W_t;$$
- *if* $\epsilon_t \leq |a_t|\,\epsilon_{t+1}$*, set*
$$W_t = 0 \quad \textit{and} \quad V_{t+1}|a_t\,V_t \sim \ell_{\epsilon_{t+1}|\frac{\epsilon_t}{|a_t|}}.$$

Then,

- *at time* t *and given the responses* $\{\hat{y}_i\}_{i=1}^t$*, the current state* x_t *is* ϵ_t*–private.*
- *the cost* C_T *is minimized; i.e.* $C_T = \frac{1}{T}\sum_{t=1}^T \frac{2}{\epsilon_t^2}$.

The proof of this result can be found in the Appendix. Theorem 5 suggests a practical online algorithm. Specifically, at time t, the samples W_t and V_{t+1} depend only on the current level ϵ_t and the next one ϵ_{t+1}. Additionally, the controller and sensor part of the privacy–preserving mechanism do not need to communicate —the sensor part can infer the noises the controller injects. At each time step, Theorem 5 performs one of the following actions.

- If the current privacy level is tighter than the next one ($\epsilon_t \leq |a_t|\,\epsilon_{t+1}$), then, the sensor performs gradual release of private data according to [13], and there is no need to inject any noise to the system.
- If the current privacy level is looser than the next one ($\epsilon_t > |a_t|\,\epsilon_{t+1}$), then, the released information $\hat{y}_t$ can be used to infer the next state x_{t+1} and, thus, violating the privacy guarantees. Theorem 5 enforces privacy by injecting noise and driving the next state of the system x_t away from the predicted one $a_t\,\hat{y}_t$.

At each time, Theorem 5 publishes accurate responses $\hat{y}_t$ of the current state $y_t = x_t$. Specifically, any other proxy with smaller expected squared error $\mathbb{E}(\hat{y}_t - y_t)^2$ would violate the privacy constraints. Nonetheless, the algorithm does not penalize the use of input noise and, therefore, minimizes the quadratic cost C_T which penalizes inaccurate responses. However, the cost C_T does not penalize the noise W_t added to the private data.

Moreover, Theorem 5 is amenable to an infinite horizon setting since, intuitively, the privacy budget is regenerated by the input noise W_t.

IV. Discussion and Future Work

In this work, we introduced the idea of time–varying differentially private guarantees. Specifically, we focused on protecting the privacy of the current state of dynamical system. After formulating the general problem, we considered two cases, a scalar time-varying system and a linear system, and proposed two privacy-preserving mechanisms. These mechanisms deviated from the ones proposed in the literature of differential privacy in that they include both a controller part which drives the private data and a sensor part which publishes the responses. Besides the technical contributions, the mechanisms proposed here conceptually differ from existing ones in differential privacy in that they consist of two parts. The controller part randomly shifts the private time over time while the sensor part resembles the existing privacy-preserving mechanisms. In particular, our mechanisms circumvent the problem of *"depleted privacy budget"*, where after long enough time, either the privacy guarantees cease to exist or the responses become unboundedly noisy. Future work includes extending the techniques and provide privacy guarantees for *current* state in nonlinear scenarios.

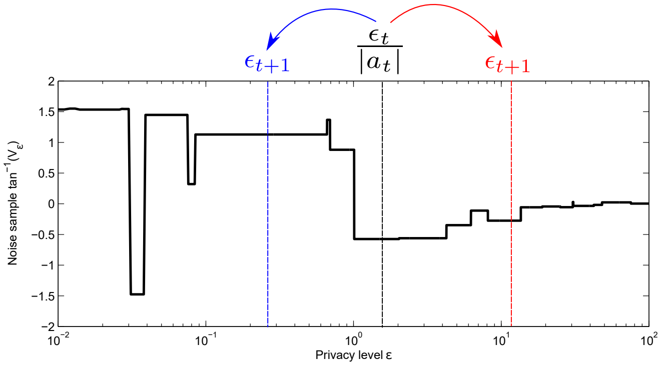

Fig. 2: Theorem 5 can be understood in terms of the stochastic process introduced in [13]. At each time step, we either perform gradual release of private data (denoted by red) and publish a more accurate reponse, or we tighten the privacy by perturbing the private data itself (denoted by blue).

References

[1] L Atzori, A Iera, and G Morabito. The internet of things: A survey. *Computer networks*, 54(15):2787–2805, 2010.

[2] L Sankar, SR Rajagopalan, and HV Poor. Utility-privacy tradeoffs in databases: An information-theoretic approach. *IEEE Transactions on Information Forensics and Security*, 8(6):838–852, 2013.

[3] MH Manshaei, Q Zhu, T Alpcan, T Bacşar, and JP Hubaux. Game theory meets network security and privacy. *ACM Computing Surveys (CSUR)*, 45(3):25, 2013.

[4] C Dwork. Differential privacy: A survey of results. In *Theory and Applications of Models of Computation*, 2008.

[5] C Dwork and A Roth. The algorithmic foundations of differential privacy. *Theoretical Computer Science*, 2013.

[6] A Ghosh, T Roughgarden, and M Sundararajan. Universally utility-maximizing privacy mechanisms. *SIAM Journal on Computing*, 41(6):1673–1693, 2012.

[7] K Nissim, S Raskhodnikova, and A Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 75–84. ACM, 2007.

[8] J Hsu, Z Huang, A Roth, and ZS Wu. Jointly private convex programming. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 580–599. SIAM, 2016.

[9] S Han, U Topcu, and GJ Pappas. Differentially private convex optimization with piecewise affine objectives. In *IEEE Conference on Decision and Control*, 2014.

[10] J Le Ny and GJ Pappas. Differentially private filtering. *Automatic Control, IEEE Transactions on*, 2014.

[11] Z Huang, S Mitra, and G Dullerud. Differentially private iterative synchronous consensus. In *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society*, pages 81–90. ACM, 2012.

[12] C Dwork. Differential privacy. In *Automata, Languages and Programming*, 2006.

[13] F Koufogiannis, S Han, and GJ Pappas. Gradual release of sensitive data under differential privacy. *arXiv preprint arXiv:1504.00429*, 2015.

[14] T Tanaka, KK Kim, PA Parrilo, and SK Mitter. Semidefinite programming approach to gaussian sequential rate-distortion trade-offs. *arXiv preprint arXiv:1411.7632*, 2014.

[15] Y Wang, Z Huang, S Mitra, and GE Dullerud. Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems. In *IEEE Conference on Decision and Control*, 2014.

[16] F Koufogiannis, S Han, and GJ Pappas. Optimality of the laplace mechanism in differential privacy. *arXiv preprint arXiv:1504.00065*, 2015.

APPENDIX

Here, we provide proofs for the two main theorems presented in this work.

Proof of Theorem 3. We will prove the theorem by assuming that the mechanism initially publishes a noisy version $\hat{x}_0$ of the initial state x_0, where

$$\hat{x}_0 = x_0 + E_0,$$

where E_0 is artificial noise and we are going to prove the privacy guarantees for such a mechanism that publishes $\hat{x}_0$ and, then, sequentially, $\hat{y}_t$. The post–processing theorem states that the privacy guarantees carry over for the mechanism that does not publish the initial response $\hat{x}_0$.

At time t, given the initial estimator $\hat{x}_0$ and the published responses $[\hat{y}_i]_{i=1}^t$, we denote with $\hat{x}_t$ the least–squares estimator of the current state x_t. For any time t, it suffices to prove that the mechanism that, given x_t as a private data, publishes the least–squares estimator $\hat{x}_t$ is (ϵ_t, δ_t)-differentially private. Indeed, given x_t, the (randomized) function that maps the estimator to the published responses

$$\hat{x}_t \to (\hat{x}_0, \hat{y}_1, \ldots, \hat{y}_t)$$

is a post–processing that is independent of the privacy–preserving mechanism that maps the private state to its least–squares estimator

$$x_t \to \hat{x}_t.$$

At time $t+1$, for a fixed x_{t+1}, the least–squares estimator $\hat{x}_{t+1}$ is derived as a linear combination of the last estimator $\hat{x}_t = x_t + E_t$ and the last published response $\hat{y}_{t+1} = C\, x_{t+1} + V_{t+1}$. Specifically, letting

$$\begin{aligned} M_t &= A\, \mathbf{\Sigma}_t\, A^T + B\, \mathbf{W}_t\, B^T \text{ and} \\ N_t &= B\, \mathbf{Y}_t - A\, \mathbf{X}_t. \end{aligned}$$

the least–squares estimator $\hat{x}_{t+1}$ is computed to be

$$\hat{x}_{t+1} = x_{t+1} + K\, V_{t+1} + (I - K\, \mathbf{C})\, (A\, E_t - B\, W_t),$$

where $K = (M_t\, C^T + N_t)\, (C\, M_t\, C^T + Z_t + \text{sym}(C\, N_t))^{-1}$ and $\text{sym}(A) = A + A^T$. The covariance of the estimation error $E_{t+1} = \hat{x}_{t+1} - x_{t+1}$ is then

$$\begin{aligned} \mathbf{\Sigma}_{t+1} = M_t - (M_t\, C^T + N_t)^T & \\ \left(C\, M_t C^T + \text{sym}(C\, N_t) + Z_t\right)^{-1} & \, (M_t\, C^T + N_t). \end{aligned}$$

Next, we relax this equality as follows. The direction of the inequality can be interpreted as the mechanism publishing a more accurate least–squares estimator than the one computed from $\hat{x}_t$ and $\hat{y}_t$. Later, we will demand that this "tighter" estimator meets our privacy requirements.

$$\begin{aligned} \mathbf{\Sigma}_{t+1} \preceq M_t - (M_t\, C^T + N_t)^T & \\ \left(C\, M_t C^T + \text{sym}(C\, N_t) + Z_t\right)^{-1} & \, (M_t\, C^T + N_t) \end{aligned}$$

We apply Schur complement to retrieve the second inequality in the constraints of (4). We complete the proof by invoking the Gaussian mechanism and requiring

$$\mathbf{\Sigma}_{t+1} \succeq \kappa^{-2}(\epsilon_t, \delta_t)\, I.$$

□

Proof of Proposition 4. In order to prove feasibility, we need to prove that, for a proper choice of the decision variables, $\mathbf{\Sigma}_t$ has full rank. Then, we can scale any such solution in order to satisfy the privacy constraint. For $\mathbf{Z}_t$ arbitrarily large, i.e. $\mathbf{Z}_t \to \infty$, the second constraint, as stated in the form of Equation reduces to

$$\mathbf{\Sigma}_t \preceq A\, \mathbf{\Sigma}_t\, A^T + B\, \mathbf{W}_t\, B^T.$$

It suffices to prove that the right hand side of the inequality is full rank. Indeed, let $v \in \mathbb{R}^n$ be such that $v^T\, \left(A\, \mathbf{\Sigma}_t\, A^T + B\, \mathbf{W}_t\, B^T\right)\, v = 0$. Then, $v^T\, A \mathbf{\Sigma}_t^{\frac{1}{2}} = 0$ and $v^T\, B\, \mathbf{W}_t^{\frac{1}{2}} = 0$ and, thus, $v^T\, A = 0$ and $v^T\, B = 0$. Since $[A; B]$ has rank n, this implies that $v = 0$ and this completes the proof. □

Proof of Theorem 5. For simplicity, we assume that $a_t \neq 0$. First, we observe that $\mathbb{E}(\hat{y}_t - y_t)^2 = \mathbb{E}\, V_t^2 \geq \frac{2}{\epsilon_t^2}$ due to the optimality of the Laplace mechanism [15], [16]. On the other hand, we use induction on t and prove that $V_t \sim \ell_{\epsilon_t}$. For $t = 1$, it holds that $V_1 \sim \ell_{\epsilon_1}$. For $t+1$, we consider two cases.

- If $\epsilon_t > |a_t|\, \epsilon_{t+1}$, since $V_t \sim \ell_{\epsilon_t}$ and $W_t \sim \ell_{\epsilon_{t+1}|\frac{\epsilon_t}{|a_t|}}$ and are independent, it follows that $V_{t+1} = a_t\, V_t - W_t \sim \ell_{\epsilon_{t+1}}$.
- If $\epsilon_t \leq |a_t|\, \epsilon_{t+1}$, then, by integrating out V_t we get that $V_{t+1} \sim \ell_{\epsilon_{t+1}}$.

Therefore, the minimum cost is achieved and this proves the second part of Theorem 5.

Next, we prove the privacy guarantees using induction on t. We abuse notation by using the symbol $\mathbb{P}$ for probability densities and re-use the same symbol for the random variable and its value Specifically, we prove that, at time t and given the current state x_t, the likelihood probability of the past responses is of the form

$$\mathbb{P}(\hat{y}_1, \ldots, \hat{y}_t) = \ell_{\epsilon_t}(\hat{y}_t - x_t)\, h(\hat{y}_1, \ldots, \hat{y}_t), \tag{6}$$

for some function h. Note that the density in Equation (6) does not depend on past states $\{x_i\}_{i<t}$. For $t = 1$ and given x_1, it holds

$$\mathbb{P}(\hat{y}_1 = z_1) = \mathbb{P}(V_1 = z_1 - x_1) = \ell_{\epsilon_1}(z_1 - x_1).$$

For $t+1$, we consider two cases.

- If $\epsilon_t > |a_t|\, \epsilon_{t+1}$, we condition on the W_t and from the induction hypothesis we get, given x_{t+1}

$$\begin{aligned}&\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_t|W_t=w)\\&\quad=\ell_{\epsilon_t}\left(\hat{y}_t-\frac{x_{t+1}-w}{a_t}\right)\,h(\hat{y}_1,\ldots,\hat{y}_t).\end{aligned}$$

Since $\hat{y}_{t+1}=a_t\,\hat{y}_t$, we compute

$$\begin{aligned}&\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_{t+1})\\&=\int_{\mathbb{R}}\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_t|W_t=w)\,\mathbb{P}(W_t=w)\,dw\\&=\ell_{\epsilon_{t+1}}(\hat{y}_{t+1}-x_{t+1})\,h_1(\hat{y}_1,\ldots,\hat{y}_{t+1}),\end{aligned}$$

for a function h_2.

- If $\epsilon_t\leq|a_t|\,\epsilon_{t+1}$, given x_{t+1}

$$\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_t)=\ell_{\epsilon_t}\left(\hat{y}_t-\frac{x_{t+1}}{a_t}\right)\,h(\hat{y}_1,\ldots,\hat{y}_t).$$

Then, given x_{t+1}

$$\begin{aligned}&\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_{t+1})\\&=\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_t)\,\mathbb{P}(\hat{y}_{t+1}|\hat{y}_t)\\&=\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_t)\\&\qquad\mathbb{P}\left(V_{t+1}=\hat{y}_{t+1}-x_{t+1}|V_t=\hat{y}_t-\frac{x_{t+1}}{a_t}\right)\\&=\ell_{\epsilon_{t+1}}(\hat{y}_{t+1}-x_{t+1})\,h_2(\hat{y}_1,\ldots,\hat{y}_{t+1}),\end{aligned}$$

for a function h_2.

We finish the proof by noting that the log–likelihood function of the responses is ϵ_t-Lipschitz in x_t

$$\left|\frac{d}{dx_t}\ln\mathbb{P}(\hat{y}_1,\ldots,\hat{y}_t)\right|=\epsilon_t.$$

□