Entropy-minimizing Mechanism for Differential Privacy of Discrete-time Linear Feedback Systems

Yu Wang, Zhenqi Huang, Sayan Mitra and Geir E. Dullerud.

Abstract—The concept of differential privacy stems from the study of private query of datasets. In this work, we apply this concept to metric spaces to study a mechanism which randomizes a deterministic query by adding mean-zero noise to keep differential privacy. For one-shot queries, we show that $\epsilon$-differential privacy of an $n$-dimensional input implies a lower bound $n + n \ln(2\epsilon)$ on the entropy of the randomized output, and this lower bound is achieved by Laplacian noise. We then consider the $\epsilon$-differential privacy of a discrete-time linear feedback system in which noise is added to the system output at each time. The adversary, modeled as a filter, estimates the system states from the output history. We show that, to keep the system $\epsilon$-differentially private, the output entropy is bounded below, and this lower bound is achieved by an explicit mechanism.

I. INTRODUCTION

The concept of $\epsilon$-differential privacy comes from the study of privacy-preserving queries of datasets [1]. Given a definition of adjacency between two datasets, a query $Q$ is called $\epsilon$-differentially private if for any two adjacent datasets $D_1, D_2$, the inequality,

$$\mathbb{P}[Q(D_1) \subseteq O] \leq e^\epsilon \mathbb{P}[Q(D_2) \subseteq O],$$

(1)

holds for any set of possible outputs $O$. In recent studies [2], [3], a connection between $\epsilon$-differential privacy and min-entropy, which measures the information leakage of the query, has been observed. Later, the concept of $\epsilon$-differential privacy, together with its variation $(\epsilon, \delta)$-differential privacy, was extended to the problems involving continuous state space dynamic systems, such as consensus and filtering problems [4], [5].

Before employing $\epsilon$-differential privacy, we have to first define adjacency on the set of data that we want to protect. Usually, two datasets $D_1, D_2$ are said to be adjacent if exactly one entry is different; however, in dynamic systems with states in $\mathbb{R}^n$, the adjacency of two states is defined via their distance with respect to some norm on $\mathbb{R}^n$. For example, given a threshold $c > 0$, two states $x_1, x_2$ are adjacent if $\|x_1 - x_2\| \leq c$ [6].

The common approach to protect $\epsilon$-differential privacy of some data is to mask the correct output with noise. Typically, in dynamic systems, noise hinders efficient communications between system components, especially in multi-agent systems. Therefore, there is usually a trade-off between privacy and performance: the system performance deteriorates when the privacy of system is enhanced.

It is well known that adding Laplacian noise ensures $\epsilon$-differential privacy, and adding Gaussian noise ensures $(\epsilon, \delta)$-differential privacy. However, it is recently shown that Laplace noise and Gaussian noise are not the best choices: the optimal choices are staircase distributions [7], [8], [9].

This results from the binary nature of adjacency; namely, two states are $x_1, x_2$ are not adjacent even if $\|x_1 - x_2\| - c$ is small but positive.

In this work, we use a metric version of $\epsilon$-differential privacy without adjacency [10]. With this definition, we first study an $\epsilon$-differentially private noise-adding mechanism for one-shot queries that provides the best output accuracy, which is measured by the Shannon entropy. Then, we consider the $\epsilon$-differential privacy of a discrete-time linear feedback system, which is characterized by the interplay of differential privacy and system dynamics. At each time, noise is added to the system outputs to keep the system states private against the adversary who has access to the system outputs. Since the system states and outputs at different time are related via the system dynamics, the adversary can filter the noise using the output history.

In the sequel, the preliminaries are presented in Section II. In Section III, we prove that, for a one-shot $n$-dimensional input, there is a lower bound $n + n \ln(2\epsilon)$ on the entropy of the output for an $\epsilon$-differentially private noise-adding mechanism, and the lower bound is achieved by Laplacian noise with parameter $\epsilon$. In Section IV, we introduce a discrete-time linear feedback system where the noise-adding mechanism is incorporated to keep the system state $\epsilon$-differentially private, and give the definition of $\epsilon$-differential privacy and output entropy for the system. In Section V, we show that the $\epsilon$-differential privacy of the system implies a lower bound on the output entropy, and give explicitly a mechanism that achieves the lower bound. Finally, we conclude this work in Section VI.

II. PRELIMINARIES

A. Notations

Denote respectively the set of natural numbers, positive integers, positive real numbers and real numbers by $\mathbb{N}$, $\mathbb{Z}_+$, $\mathbb{R}_+$ and $\mathbb{R}$. Let $\mathbb{R}^n$ and $\mathbb{R}^{m \times n}$ be the set of $n$-dimensional real vectors and the set of $m$ by $n$ real matrices. Denote the positive orthant in $\mathbb{R}^n$ by $\mathbb{R}_+^n$. In this work, scalars and vectors are in lower case; matrices are in upper case; random variables are in roman font.

For $n \in \mathbb{N}$, let $[n] = [0, 1, 2, \ldots, n]$. For $x \in \mathbb{R}$, denote the absolute value of $x$ by $|x|$. For $x \in \mathbb{R}^n$, denote the $i$-th component of $x$ by $x_i$; i.e., $x = (x_1, x_2, \ldots, x_n)$. For
If \( x, y \in \mathbb{R}^n \), let \( x \cdot y = \sum_{i=1}^{n} x_i y_i \). For a function \( f \), denote the image of \( f \) by \( \text{Im}(f) \). For a scalar function \( f(x) \) on \( \mathbb{R} \), let \( f'(x) = \frac{\partial}{\partial x} f(x) \).

A metric space is denoted by \((M, \rho, (\cdot, \cdot))\) where \( \rho, (\cdot, \cdot) \) is the metric on \( M \), \( \mathbb{R}^n \), equipped with a norm \( \| \cdot \| \) is a metric space, denoted by \((\mathbb{R}^n, \| \cdot \|)\), where \( \rho(x, y) = \|x - y\| \).

In particular, the \( l_1 \)-norm on \( \mathbb{R}^n \), which will be frequently referred to later, is defined as \( \|x\|_1 = \sum_{i=1}^{n} |x_i| \).

### B. Laplace Distribution

On \( \mathbb{R} \), a random variable \( x \) obeys Laplace distribution with parameter \( \epsilon \), denoted by \( x \sim \text{Lap}(\lambda) \), if its probability distribution function is

\[
\frac{1}{2\epsilon} \exp\left(\frac{|x|}{\epsilon}\right).
\]

On \( \mathbb{R}^n \), \( x \sim \text{Lap}_n(\lambda) \) if its probability distribution function is

\[
\frac{1}{2\epsilon^n} \exp\left(\frac{\|x\|_1}{\epsilon}\right).
\]

### C. Differential Privacy

Suppose that we have some data \( x \) taken from a metric space \((I, \rho, (\cdot, \cdot))\). To keep the query of \( x \) private, we pass \( x \) through a mechanism \( M \) which generates a randomized output \( M(x) \).

**Definition 1:** The mechanism is called \( \epsilon \)-differentially private for given \( \epsilon > 0 \), if for any two inputs \( x_1, x_2 \), and a set of possible outputs \( O \),

\[
P[M(x_1) \subseteq O] \leq \exp(\epsilon \rho(x_1, x_2)) P[M(x_2) \subseteq O],
\]

This definition is an extension of differential privacy to metric spaces. As \( \epsilon \to \infty \), any mechanism \( M \) will be \( \epsilon \)-differentially private. For \( \epsilon = 0 \), a mechanism \( M \) is \( \epsilon \)-differentially private iff the randomized output is independent of the input.

In this work, as shown in Figure 1, we let the metric space be a real normed space \((\mathbb{R}^n, \| \cdot \|)\) and the mechanism \( M \) generate the randomized output by adding noise \( n(x) \).

### D. Entropy

Let \( x \) be a random variable with probability distribution function \( f(x) \).

**Definition 2:** The entropy of \( x \) is

\[
H(x) = -\int f(x) \ln(f(x))dx
\]

Informally speaking, the entropy is a measure of the uncertainty of a random variable, either positive or negative. Conditioning reduces entropy; for two random variables \( x \) and \( y \),

\[
H(x | y) \leq H(x).
\]

with equality iff \( x \) and \( y \) are independent.

---

**III. Entropy Analysis for Differentially Private One-shot Queries**

In this section, we show that, for a single query, \( \epsilon \)-differential privacy implies a bound on the noise of the entropy and the bound is achieved by adding Laplacian noise.

Let \( M \) be a randomizing mechanism with input set \( I = (\mathbb{R}^n, \| \cdot \|) \). For any \( x \in \mathbb{R}^n \), assume that the randomized output \( x = M(x) \) has a probability density function \( f_x(y) \).

Let \( p(x, y) = f_x(y) \), which we referred to as the output distribution function. We assume that \( p(x, y) \) is absolutely continuous in \( x \) and the noise added is mean-zero; i.e.,

\[
\int_{\mathbb{R}^n} yp(x, y)dy = x.
\]

**Lemma 3:** A mechanism \( M \) with absolutely continuous output distribution function \( p(x, y) \) is \( \epsilon \)-differentially private iff for any \( x \in \mathbb{R}^n \) and unit vector \( \hat{n} \) under the norm \( \| \cdot \| \),

\[
|\hat{n} \cdot \nabla_x p(x, y)| \leq p(x, y)
\]

holds for almost every \( y \), where the gradient \( \nabla_x p(x, y) \) is taken for fixed \( y \in \mathbb{R}^n \).

**Proof:** **Necessity.** By Definition 1, for any two inputs \( x_1, x_2 \in \mathbb{R}^n \) and an output set \( O \subseteq \mathbb{R}^n \), we have \( \int_O p(x_1, y)dy \leq \exp(\epsilon \|x_1 - x_2\|) \int_O p(x_2, y)dy \).

Noting that \( p(x, y) \) is continuous, we have \( p(x, y) \leq \exp(\epsilon \|x_1 - x_2\|) p(x_2, y) \) for any \( y \in \mathbb{R}^n \). Thus,

\[
\frac{p(x_1, y) - p(x_2, y)}{\|x_1 - x_2\|} \leq \frac{\epsilon \|x_1 - x_2\| - 1}{\|x_1 - x_2\|} p(x_2, y)
\]

By letting \( x_1 \to x_2 \), we have \( \epsilon \|x_1 - x_2\| \cdot \nabla_x p(x_2, y) \mid \leq p(x_2, y) \), for \( x \) almost everywhere, abbreviated as a.e. Since \( x_1 \) can approach \( x_2 \) in arbitrary direction, we have \( \epsilon \|\hat{n} \cdot \nabla_x p(x_2, y)\| \leq p(x_2, y) \) for arbitrary unit vector \( \hat{n} \) and \( x \) a.e.

**Sufficiency.** For any two inputs \( x_1, x_2 \in \mathbb{R}^n \), define

\[
g(z) = p(x_2 + \hat{n}z, y)
\]

where \( \hat{n} = \frac{x_1 - x_2}{\|x_1 - x_2\|} \). Clearly, \( g(0) = p(x_2, y) \) and \( g(\|x_1 - x_2\|) = p(x_1, y) \).

By (8), \( \epsilon |g'(z)| \leq g(z) \) for \( z \) a.e. Thus,

\[
|\ln p(x_1, y) - \ln p(x_2, y)| = |\ln g(\|x_1 - x_2\|) - \ln g(0)|
\]

\[
\leq \int_{0}^{\|x_1 - x_2\|} \frac{|g'(z)|}{g(z)}dz \leq \epsilon \|x_1 - x_2\|
\]

Therefore,

\[
p(x_1, y) \leq \exp(\epsilon \|x_1 - x_2\|) p(x_2, y)
\]

---

![Block Diagram for a \( \epsilon \)-Differentially Private Mechanism](image)
Finally, by integrating (12) over \( y \) on any output set \( O \subseteq \mathbb{R}^n \), we have
\[
\int_O p(x, y) dy \leq \exp (\epsilon \| x_1 - x_2 \|) \int_O p(x, y) dy.
\]

The output entropy of a mechanism \( \mathcal{M} \) is evaluated by the supremum of all possible inputs.
\[
\mathcal{H}(\mathcal{M}) = \sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} -p(x, y) \ln p(x, y) dy.
\]

From now on, we propose several properties of a mechanism \( \mathcal{M} \) that minimizes the output entropy \( \mathcal{H}(\mathcal{M}) \).

**Lemma 4:** Given any \( \epsilon \)-differentially private \( \mathcal{M} \) with output distribution function \( p(x, y) \), we can construct an \( \epsilon \)-differentially private mechanism \( \mathcal{N} \) with output probability distribution function \( q(x, y) \), such that (i) \( \mathcal{H}(\mathcal{N}) \leq \mathcal{H}(\mathcal{M}) \), and (ii) for any \( x \in \mathbb{R}^n \), \( g_a(x) = q(x, y-x) \) is even with respect to each \( y_m \) where \( m \in [n] \).

**Proof:** Without loss of generality, assume \( m = 1 \). Let
\[
\mathcal{H}^+_1(\mathcal{M}) = \sup_{x \in \mathbb{R}^n} \int_{(x_1, \infty) \times \mathbb{R}^{n-1}} -p(x, y) \ln p(x, y) dy,
\]
\[
\mathcal{H}^-_1(\mathcal{M}) = \sup_{x \in \mathbb{R}^n} \int_{(-\infty, x_1) \times \mathbb{R}^{n-1}} -p(x, y) \ln p(x, y) dy.
\]

Define
\[
q(x, y) = \begin{cases} 
p(x, y) & \text{if } y > x, \mathcal{H}^+_1(\mathcal{M}) \leq \mathcal{H}^-_1(\mathcal{M}) , \\
p(x, z) & \text{if } y < x, \mathcal{H}^+_1(\mathcal{M}) > \mathcal{H}^-_1(\mathcal{M}) , 
\end{cases}
\]
where
\[
\begin{align*}
z_i &= 2x \epsilon - y_i, & i &= 1, \\
z_i &= y_i, & i &= 2, 3, \ldots, n.
\end{align*}
\]
By definition, \( q(x, y) \geq 0 \), and \( q(x, y) \) is absolutely continuous. Since the noise is mean-zero, for any \( x \in \mathbb{R}^n \),
\[
\int_{\mathbb{R}^n} q(x, y) dy = 1 \quad \text{and} \quad \int_{\mathbb{R}^n} yq(x, y) dy = x.
\]
Noting that for fixed \( y \) and any \( x \neq y \),
\[
\nabla_x q(x, y) = \begin{cases} 
\nabla_x p(x, y) & \text{if } y \geq x, \mathcal{H}^+_1(\mathcal{M}) \leq \mathcal{H}^-_1(\mathcal{M}) , \\
\nabla_x p(x, z) & \text{if } y < x, \mathcal{H}^+_1(\mathcal{M}) > \mathcal{H}^-_1(\mathcal{M}) ,
\end{cases}
\]
where \( z \) is given in (17), we have \( c |\nabla_x q(x, y)| \leq q(x, y) \) for \( x \) a.e. By Lemma 3, \( \mathcal{N} \) is a well-defined \( \epsilon \)-differentially private mechanism.

**Lemma 5:** Given any \( \epsilon \)-differentially private \( \mathcal{M} \) with output distribution function \( p(x, y) \), we can construct an \( \epsilon \)-differentially private mechanism \( \mathcal{N} \) with output probability distribution function \( q(x, y) \), such that (i) \( \mathcal{H}(\mathcal{N}) \leq \mathcal{H}(\mathcal{M}) \), and (ii) for any \( x \in \mathbb{R}^n \) and \( m \in [n] \), fixing \( \{ x_i | i \neq m, i \in [n] \} \) and \( \{ y_i | i \neq m, i \in [n] \} \), the equation \( q(x_m, y) = q(2a - x_m, 2a - y_m) \) holds.

**Proof:** Without loss of generality, assume \( m = 1 \). Let \( L^+ = \{ x \in \mathbb{R}^n | x_1 \geq t \} \) and \( L^- = \{ x \in \mathbb{R}^n | x_1 < t \} \). Define
\[
\mathcal{H}^+_1(\mathcal{M}) = \sup_{L^+} \int_{\mathbb{R}^n} -p(x, y) \ln p(x, y) dy,
\]
\[
\mathcal{H}^-_1(\mathcal{M}) = \sup_{L^-} \int_{\mathbb{R}^n} -p(x, y) \ln p(x, y) dy.
\]
For fixed \( \{ x_i | i \neq 1, i \in [n] \} \) and \( \{ y_i | i \neq 1, i \in [n] \} \), if \( \mathcal{H}^+_1(\mathcal{M}) \leq \mathcal{H}^-_1(\mathcal{M}) \), then define
\[
q(x_1, y_1) = \begin{cases} 
p(x_1, y_1), & x_1 \in L^+, \\
p(2a - x_1, 2a - y_1), & x_1 \in L^-.
\end{cases}
\]
otherwise, define
\[
q(x_1, y_1) = \begin{cases} 
p(2a - x_1, 2a - y_1), & x_1 \in L^+, \\
p(x_1, y_1), & x_1 \in L^-.
\end{cases}
\]
By definition, \( q(x_1, y_1) = q(2a - x_1, 2a - y_1) \). By Lemma 4, let \( p(a, y_1) = p(a, 2a - y_1) \), then \( q(x_1, y_1) \) is continuous at \( x_1 = a \), hence continuous with respect to \( x_1 \) and \( y_1 \).

Clearly, \( q(x, y) \geq 0 \), \( \int_{\mathbb{R}^n} q(x, y) dy = 1 \) and \( \int_{\mathbb{R}^n} yq(x, y) dy = x \). By definition, \( q(x, y) \) is absolutely continuous, and \( \| \nabla x q(x, y) \| = \| \nabla x p(x, y) \| \). Thus \( c | \nabla_x q(x, y) | \leq q(x, y) \) for \( x \) a.e. By Lemma 3, \( \mathcal{N} \) is a well-defined \( \epsilon \)-differentially private mechanism. Furthermore,
\[
\mathcal{H}(\mathcal{N}) = \min \{ \mathcal{H}^+_1(\mathcal{M}), \mathcal{H}^-_1(\mathcal{M}) \} \quad \text{and} \quad \mathcal{H}(\mathcal{M}) = \mathcal{H}(\mathcal{N}).
\]
The equality holds iff \( \mathcal{H}^+_1(\mathcal{M}) = \mathcal{H}^-_1(\mathcal{M}) \).

**Lemma 6:** There exists an \( \epsilon \)-differentially private mechanism \( \mathcal{M} \), such that (i) \( \mathcal{M} \) minimizes the output entropy \( \mathcal{H}(\mathcal{M}) \), and (ii) the noise added is independent of the input \( x \in \mathbb{R}^n \).

**Proof:** Let \( q(x, y) = p(x, y - x) \). Then \( q(x, y) \) is the probability distribution function of the noise \( n(x) \) for input \( x \). By Lemma 5, for any \( t \in \mathbb{R} \) and \( m \in [n] \), fixing \( \{ x_i | i \neq m, i \in [n] \} \) and \( \{ y_i | i \neq m, i \in [n] \} \), we can have \( q(x_m, y_m) = q(2t - x_m, -y_m) \). By Lemma 4, we can have \( q(x_m, y_m) = q(2t - x_m, -y_m) = q(2t - x_m, y_m) \). Therefore, \( q(x, y) \) is independent of \( x_m \), hence independent of \( x \).

By Lemma 4 and Lemma 6, if a mechanism \( \mathcal{M} \) minimizes the output entropy \( \mathcal{H}(\mathcal{M}) \), the output distribution function will be in the form \( p(x, y) = f(y-x) \), where \( f \) is an absolutely continuous probability distribution function. For each \( i \in [n] \), \( f \) is a even function with respect to \( x_i \). For now, we assume a scalar input, that is \( n = 1 \). Then the problem of minimizing (13) becomes
Problem 1 (Scalar Case):

Minimize: \( H(f) = -\int_{[0,\infty)} f(x) \ln f(x) dx \),

subject to: \( f(x) \) is absolutely continuous,  
\( f(x) \geq 0 \),  
\( \epsilon |f'(x)| \leq f(x) \) a.e.,  
\( \int_{[0,\infty)} f(x) dx = \frac{1}{2} \).

Lemma 7: \( f(x) \) is nonincreasing if it solves Problem 1.

Proof: Let \( f(x) \) be a function that solves Problem 1. Let \( g(x) = \sup_{y \geq x} f(y) \). Clearly, \( g(x) \geq f(x) \) for \( x \geq 0 \). Assume that \( f(x) \) is not non-increasing. Then for some \( x^* > 0 \), \( g(x^*) > f(x^*) \). By continuity of \( f \), there exists a “largest” non-empty interval \((a, b)\) containing \( x^* \), on which \( g(x) > f(x) \). Note that \( b \) is finite since \( f(x) > 0 \) and \( \lim_{x \to \infty} f(x) = 0 \). In addition, \( g(b) = f(b) \). Let  
\[ d = \frac{1}{f(a)} \int_a^b f(x) dx. \]  
(25)

By definition, \( d \in [0, b - a] \).

There are two cases on the value of \( a \). If \( a > 0 \), then \( f(a) = g(a) = f(b) = g(b) \). Define  
\[ h(x) = \begin{cases} f(x), & x \in [0, a], \\ f(b), & x \in [a, a + d], \\ f(x + b - a - d), & x \in [a + d, \infty], \end{cases} \]  
(26)

Otherwise, \( a = 0 \). Define  
\[ h(x) = \begin{cases} f(b), & x \in [0, d], \\ f(x + b - d), & x \in [d, \infty]. \end{cases} \]  
(27)

In both cases, \( h(x) \) satisfies the constraints in Problem 1 and \( H(h) < H(f) \). This is in contradiction with the assumption.

Theorem 8: For a \( \epsilon \)-differentially private mechanism \( \mathcal{M} \) with input set \( \mathcal{I} = (\mathbb{R}, | \cdot |) \), the output entropy satisfies  
\[ H(\mathcal{M}) \geq 1 + \ln(2\epsilon). \]  
(28)

The equality holds if the output distribution function is  
\[ p(x, y) = \frac{1}{2\epsilon} \exp(-\frac{\epsilon}{2\epsilon}). \]  
(29)

Proof: It suffices to show that in the above minimization Problem 1, \( H \geq \frac{1}{2} + \frac{1}{2} \ln(2\epsilon) \) and the minimum is achieved by \( f(x) = \frac{1}{2\epsilon} \exp(-\frac{\epsilon}{2\epsilon}) \).

Let \( F(x) = \int_x^{\infty} f(y) dy \) and note that \( f(\infty) = \lim_{x \to \infty} f(x) = 0 \). By definition of \( \epsilon \)-differential privacy,  
\[ F(x) \geq \int_x^{\infty} \epsilon |f'(x)| dy \geq \epsilon \int_x^{\infty} f'(x) dy = \epsilon |f(\infty) - f(x)| = \epsilon f(x) \]  
(30)
The first equality holds iff \( |f'(x)| = f(x) \) for \( x \) a.e. The second equality holds iff \( f(x) \) is a monotone increasing function. Considering that \( f(x) > 0 \) and \( f(\infty) = 0 \), \( f(x) \) is monotonously decreasing. In sum, \( F(x) = \epsilon f(x) \) iff \( f'(x) = -f(x) \). In particular, \( f(0) \geq \frac{F(0)}{\epsilon} = \frac{1}{2\epsilon} \).

By Lemma 7, \( f'(y) \leq 0 \) a.e. Thus,  
\[ H(f) = -\int_0^{\infty} f(x) \ln f(x) dx \]  
\[ = -\int_0^{\infty} f(x) \left( \ln f(0) + \int_0^x f'(y) dy \right) dx \]  
\[ = -\frac{1}{2} \ln f(0) - \int_0^{\infty} f'(y) \left( \int_0^x f(x) dx \right) dy \]  
\[ = -\frac{1}{2} \ln f(0) - \int_0^{\infty} f'(y) F(y) dy \]  
\[ \geq -\frac{1}{2} \ln f(0) - \int_0^{\infty} \epsilon f'(y) dy \]  
\[ = \epsilon f(0) - \frac{1}{2} \ln f(0) \]  
(31)
The equality holds iff \( \epsilon f'(x) = -f(x) \).

Since \( f(0) \in (0, \frac{1}{2\epsilon}) \) and \( \epsilon f(0) - \frac{1}{2} \ln f(0) \) is decreasing on this interval, we have \( H \geq \frac{1}{2} + \frac{1}{2} \ln(2\epsilon) \). Again, the equality holds if \( \epsilon f'(x) = -f(x) \) a.e.

Remark 9: If we replace the restriction \( \int_0^{\infty} f(x) dx = \frac{1}{2} \) in Problem 1 with \( \int_0^{\infty} f(x) dx = \frac{c}{2} \), where \( c > 0 \) is a real constant, then the above proof will still work. In this case, we have \( H \geq \frac{1}{2} + \frac{1}{2} \ln(2\epsilon) + \frac{1}{2} \ln c \) and the minimum is achieved by  
\[ f(x) = \frac{c}{2\epsilon} \exp(-\frac{x}{\epsilon}). \]  
(33)

For multidimensional inputs, \( n \geq 2 \), noting that we have taken the \( \ell_1 \)-norm on \( \mathbb{R}^n \), the problem of minimizing (13) becomes

Problem 2 (Multi-dimensional Case):

Minimize: \( H(f) = -\int_{\mathbb{R}^n_+} f \ln f dx \),

subject to: \( f(x_1, x_2, \ldots, x_n) \) is absolutely continuous,  
\( f(x) \geq 0 \),  
\( \epsilon |\frac{\partial f}{\partial x_i}| \leq f, \forall i \in [n] \) a.e.,  
\( \int_{\mathbb{R}^n} f dx = \frac{1}{2^n} \).

Theorem 10: For a \( \epsilon \)-differentially private mechanism \( \mathcal{M} \) with input set \( \mathcal{I} = (\mathbb{R}^n, | \cdot |) \), the output entropy satisfies  
\[ H(\mathcal{M}) \geq n + n \ln(2\epsilon). \]  
(34)
The equality holds if the output distribution function is  
\[ p(x, y) = \left(\frac{1}{2\epsilon}\right)^n \exp\left(-\frac{\epsilon}{2\epsilon}\|y - x\|_1\right). \]  
(35)

Proof: It suffices to show that in the above minimization Problem 2, \( H \geq \frac{n}{2} + \frac{1}{2} \ln(2\epsilon) \) and the minimum is achieved by \( f(x_1, x_2, \ldots, x_n) = \left(\frac{1}{2\epsilon}\right)^n \exp\left(-\frac{\epsilon}{2\epsilon}\|x_1 + x_2 + \ldots + x_n\|_1\right) \).
For each fixed \( x_2, x_3, \ldots, x_n \), let
\[
g_{x_2, x_3, \ldots, x_n}(x_1) = f(x_1, x_2, \ldots, x_n),
\]
then we have \( g_{x_2, x_3, \ldots, x_n}(x_1) \geq 0 \), \( \epsilon|g_{x_2, x_3, \ldots, x_n}(x_1)| \leq g_{x_2, x_3, \ldots, x_n}(x_1) \) and
\[
H(f) = -\int_{\mathbb{R}_+^{n-1}} \left( \int_{[0, \infty)} g(x_1) \ln g(x_1) \, dx_1 \right) \, dx_2 \, dx_3 \ldots \, dx_n.
\]
By Theorem 8 together with Remark 9, we learn that, to minimize \( H \),
\[
f(x_1, x_2, \ldots, x_n) = g_{x_2, x_3, \ldots, x_n}(x_1) = \exp(-\frac{x_1}{\epsilon}) h(x_2, x_3, \ldots, x_n)
\]
where \( h(x_2, x_3, \ldots, x_n) \) is some function of \( x_2, x_3, \ldots, x_n \). Note that the above result also holds for \( x_2, x_3, \ldots, x_n \), thus we have
\[
f(x_1, x_2, \ldots, x_n) = k \exp(-\frac{x_1 + x_2 + \ldots + x_n}{\epsilon})
\]
where \( k \) is some constant. By the restriction \( \int_{\mathbb{R}_+^n} f \, dx = \frac{1}{\epsilon} \), we have \( k = \left( \frac{1}{\epsilon} \right)^n \). By plugging (39) back to Problem 2, we have \( H_{\text{min}} = \frac{n}{2} + \frac{n}{2} \ln(2\epsilon) \).

IV. DIFFERENTIAL PRIVACY OF DISCRETE-TIME LINEAR FEEDBACK SYSTEM

In the following, we assume that the probability distribution functions exist when necessary. In [11], the authors discussed the differential privacy of a distributed control system. Here, we simplify the system to a discrete-time linear feedback system, as shown in Figure 2. The system input \( u(t) \) here corresponds to the “system preference” in the setup of [11]. In the following, we set \( u(t) = 0 \). Thus
\[
v(t) = u(t) + y(t) = y(t).
\]
The system output \( v(t) \) is fed back to the input of the plant \( \mathcal{P} \) through the controller \( \mathcal{C} \). Here, we assume \( \mathcal{C} \) to be a proportional controller with unit gain; that is,
\[
v(t) = y(t) = z(t).
\]

Similar analysis can be applied to systems with linear feedback control. In the plant \( \mathcal{P} \), the state \( x(t) \in \mathbb{R}^n \) is updated via
\[
x(t + 1) = Ax(t) + Bv(t),
\]
where \( A, B \in \mathbb{R}^{n \times n} \). From (42), we know for \( t \geq 1 \),
\[
x(t) = A^t x(0) + \sum_{i=0}^{t-1} A^{t-i-1} B z(i).
\]
The state \( x(t) \) of the plant \( \mathcal{P} \), referred to as the system state, is passed directly to the mechanism \( \mathcal{L} \) which generates an output \( z(t) \), referred to as the system output, by adding a time-dependent mean-zero noise \( n(t) \),
\[
z(t) = x(t) + n(t).
\]

According to Lemma 6, we assume that the noise is independent of the system state. The noise \( \{u(t) \mid t \in \mathbb{N}\} \) may be correlated. Therefore, “mean-zero” here indicates that the noise \( n(t) \) conditioning on the initial state \( x(0) \) and a history of noise \( \{n(i) \mid i \in [t - 1]\} \) is mean-zero. That is,
\[
\mathbb{E}[n(t) \mid x(0), n(0), n(1), \ldots, n(t - 1)] = 0.
\]
The system output \( z(t) \) is accessible to the adversary. As shown in (43), due to the feedback setting, given the past history of outputs \( \{z(t) \mid t \in \mathbb{N}\} \), we can derive the trajectory of system states \( \{x(t) \mid t \in \mathbb{N}\} \) from the initial system state \( x(0) \). Therefore, protecting the \( \epsilon \)-differential privacy of the initial system state is equivalent to protecting the \( \epsilon \)-differential privacy of the whole system state trajectory.

For \( t \geq 0 \), the adversary \( \mathcal{A} \) estimates the initial system state from the past history of output \( \{z(i) \mid i \in [t]\} \). The best estimation of \( x(0) \) at time \( t \) is a random variable such that
\[
\tilde{x}(t) = \mathbb{E}[x(0) \mid z(0), z(1), \ldots, z(t)].
\]
We denote the probability density function of \( \tilde{x} \) by \( \tilde{h}_t \). Since the adversary gets cumulating outputs, the entropy of \( \tilde{x}(t) \) is nonincreasing.

Definition 11: The mechanism \( \mathcal{L} \) is \( \epsilon \)-differentially private at time \( t \in \mathbb{N} \), if for any pair of initial states \( x_1, x_2 \in \mathbb{R}^n \),
\[
\tilde{h}_t(x_1) \leq \exp(\epsilon \|x_1 - x_2\|) \tilde{h}_t(x_2).
\]
Roughly speaking, the shape of distribution of \( \tilde{x}_t \) is flat enough to mask two different guesses \( x_1, x_2 \) of the initial state. Finally, we define the output entropy for the system.

Definition 12: The output entropy \( H(\mathcal{L}, t) \) of mechanism \( \mathcal{L} \) at time \( t \in \mathbb{N} \) is \( H(\mathcal{L}, t) = H(\tilde{x}(t)) \).

V. ENTROPY ANALYSIS

In this section, we consider the \( \epsilon \)-differentially private mechanisms and derive the one that gives the most accurate output up to time \( t \geq 0 \); i.e., a mechanism \( \mathcal{L} \) that minimizes the output entropy \( H(\mathcal{L}, i) \) at each time \( i \in [t] \) while preserving \( \epsilon \)-differential privacy.

First we assume that the norm used in this section is the \( \ell_1 \)-norm \( \| \cdot \|_1 \). By combining (40) (41), we have
\[
x(t + 1) = (A + B)x(t) + Bn(t),
\]
which implies that
\[ x(t) = (A + B)^t x(0) + \sum_{i=0}^{t-1} (A + B)^{t-i-1} B n(i). \] (49)

Thus, \( z(0) = x(0) + n(0) \) and for \( t \geq 1 \)
\[ z(t) = (A + B)^t x(0) + \sum_{i=0}^{t-1} (A + B)^{t-i-1} B n(i) + n(t) \] (50)

On the other hand, by (43) and (44), we have \( n(0) = z(0) - x(0) \) and for \( t \geq 1 \),
\[ n(t) = z(t) - x(t) \]
\[ = z(t) - A^t x(0) - \sum_{i=0}^{t-1} A^{t-i-1} B z(i) \] (51)

For simplicity, let \( w_0 = z_0 \), and for \( t \geq 1 \),
\[ w_t = z_t - \sum_{i=0}^{t-1} A^{t-i-1} B z_i. \] (52)

**Lemma 13**: Given the initial state \( x(0) \) and a history of outputs \( \{ z(i) \mid i \in [t-1] \} \), we have
\[ \mathbb{E}[n(t) \mid x(0), z(0), z(1), \ldots, z(t-1)] = 0. \] (53)

**Proof**: By (50) and (51), there is an one-to-one linear map between \( \{ x(0), x(0), u(0), u(1), \ldots, u(t-1) \} \) and \( \{ x(0), x(0), z(0), z(1), \ldots, z(t-1) \} \). Then by (45), (53) holds. \( \blacksquare \)

**Lemma 14**: Let \( \tilde{\alpha}_t = \mathbb{E}[\tilde{x}(t)] \). Then we have
\[ A^t \tilde{\alpha}_t = w_t. \] (54)

**Proof**: In (51), take expectation conditioning on \( \{ z(i) \mid i \in [t] \} \). \( \blacksquare \)

**Theorem 15**: For an \( \epsilon \)-differential privacy mechanism \( \mathcal{L} \),
\[ H(\mathcal{L}, t) \geq n + n \ln(2\epsilon), \] (55)
for any time \( t \geq 0 \). The equality holds when \( n(0) \sim \text{Lap}(\epsilon) \), and for \( t \geq 1 \), \( n(t) = \text{An}(t-1) \).

**Proof**: Denote the joint probability distribution function of \( \{ n(i) \mid i \in [t] \} \) by \( p_t(n_0, n_1, \ldots, n_t) \). By (51),
\[ \hat{h}_0(x) = c_0 p_0(w_0 - x) \]
\[ \hat{h}_1(x) = c_1 p_1(w_0 - x, w_1 - A x) \]
\[ \ldots \]
\[ \hat{h}_t(x) = c_t p_t(w_0 - x, \ldots, w_t - A^t x) \]
\[ \ldots \] (56)

where \( \{ c_t \mid t \in \mathbb{N} \} \) are unifying constants. To minimize the entropy of \( \tilde{x} \) at each time \( t \in \mathbb{N} \) while ensuring \( \epsilon \)-differential privacy, by Theorem 8, we should have
\[ \hat{h}_t(x) = \frac{1}{2\epsilon} \exp(-\frac{\| x - \tilde{\alpha}_t \|_1}{\epsilon}). \] (57)

Plug (54) (57) into (56). Since the new equalities holds for arbitrary \( x \) and \( \{ w_t \mid t \in \mathbb{N} \} \), by comparing the arguments, we have
\[ p_0(x_0) = d_1 \exp(-\frac{\| x_0 \|_1}{\epsilon}) \]
\[ p_1(x_0, Ax_1) = d_2 \exp(-\frac{\| x_1 \|_1}{\epsilon}) \]
\[ \ldots \]
\[ p_t(x_0, \ldots, A^t x_t) = d_t \exp(-\frac{\| x_t \|_1}{\epsilon}) \]
\[ \ldots \] (58)

where \( \{ x_t \mid t \in \mathbb{N} \} \subseteq \mathbb{R} \) and \( \{ d_t \mid t \in \mathbb{N} \} \) are unifying constants. A mechanism that satisfies (58) is \( n(0) \sim \text{Lap}(\epsilon) \) and for \( t \geq 1 \), \( n(t) = \text{An}(t-1) \). The noise \( n(1) \) is determined by \( n(0) \); \( n(2) \) is determined by \( n(0), n(1) \); and so on.

In this case, the output entropy is independent of time \( t \),
\[ H(\mathcal{L}, 1) = H(\mathcal{L}, 2) = \ldots = H(\mathcal{L}, t) = n + n \ln(2\epsilon). \] (59)

This is because all the randomness is brought to the system at the initial time. \( \blacksquare \)

VI. CONCLUSION

In this work, we introduced the entropy-minimizing mechanism for differential privacy. In particular, we studied a query mechanism which randomizes a deterministic quantity by adding mean-zero noise to keep differential privacy. For one-shot queries, we showed that for \( n \)-dimensional input, \( \epsilon \)-differential privacy implies a lower bound \( n + n \ln(2\epsilon) \) on the entropy of the randomized output of the mechanism, and this lower bound is achieved by adding Laplacian noise with parameter \( \epsilon \). Then we studied a discrete-time linear feedback system that simplifies the system in [11], where the adversary is modeled as a filter on the system output. We demonstrate that there is a lower bound on the entropy if the mechanism ensures the \( \epsilon \)-differential privacy of the system. This lower bound is achievable by \( n(0) \sim \text{Lap}_n(\epsilon) \), and for \( t \geq 1 \), \( n(t) = \text{An}(t-1) \).

REFERENCES

