DIFFERENTIAL PRIVACY OVER AFFINE MANIFOLDS

Anonymous authors
Paper under double-blind review

ABSTRACT

In this paper, we study dataset processing mechanisms generated by linear queries over affine manifolds. Specifically, the input data are assumed to lie in an affine manifold. This affine manifold adds to an inherent geometry in the domain of secrets, which acts as a special constraint that may be known about the data by adversaries. To take care of the presence of this affine manifold geometry, a new neighborhood of adjacency between two databases is introduced where the dimension of the manifold has to be accounted for. We establish necessary and sufficient conditions on the possibility of achieving differential privacy via structured noise injection mechanisms where non i.i.d. Gaussian or Laplace noises are calibrated into the dataset. Next, in light of these conditions, procedures are developed by which a prescribed privacy budget can be tightly reached with a matching noise level. Finally, we show that the framework has immediate applications in differentially private cloud-based control, where the affine-manifold data dependency arises naturally from the system dynamics, and the proposed theories and procedures become effective tools in evaluating privacy levels and in the design of provably privacy-preserving algorithms.

1 INTRODUCTION

The rapid development in big data and machine learning has sparkled a revolution in various engineering disciplines during the past decade, such as manufacturing, the Internet of Things (IoT), E-commerce, healthcare, computer vision, etc. Advanced learning representations and efficient training on large datasets are uncovering the tremendous power hidden in data on a daily basis, enabled by the drastic improvements in effective data collection, storage, and processing Qiu et al. (2016). Information about individuals' identities, preferences, and activities becomes inevitably embedded, directly or indirectly, in a variety of datasets collected from various sensors and social media. The underlying privacy risks for individual users in such data-driven applications are becoming increasingly important Zhu & Blaschko (2020), especially so when the datasets involve sensitive private data such as political associations, biometric fingerprints, and healthcare records Wu et al. (2020).

The fundamental challenge in managing privacy risks of data analysis has been the tradeoff between protecting datapoint information and maintaining analysis accuracy. Classical work Denning & Denning (1979); Denning & Schlörer (1980) revealed potential privacy threats facing statistical database, where an adversary may create sequential queries to infer confidential datapoint. It was later proven that it is impossible not to reveal information about datapoint in queries unless random noise has been injected into the database Dinur & Nissim (2003). The seminar work Dwork et al. (2006a;b) brought up the idea of differential privacy in quantifying the amount of privacy risk in a randomized mechanism, where a numerical privacy budget characterizes the probabilistic similarities at the mechanism output between two datapoints of the mechanism input that are close (adjancent) with each other. In particular, differential privacy concerns with the privacy risk embedded in a *mechanism*, denoted by $\mathscr{M} : \mathcal{D} \to \mathcal{M}$, which is a randomized mapping from the input space $\mathcal{D}$ to the output space $\mathcal{M}$ Dwork et al. (2006b).

Definition 1. *(μ-*Adjacency*) For any two data $\mathbf{x}$ and $\mathbf{x}'$ drawn from the set $\mathcal{D} \subseteq \mathbb{R}^n$, they are said to be μ-adjacent with $\mu > 0$, denoted by $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu)$, if $\|\mathbf{x} - \mathbf{x}'\|_0 = 1$ and $\|\mathbf{x} - \mathbf{x}'\|_1 \leq \mu$.*

Definition 2. *(*Differential Privacy*) A randomized mechanism $\mathscr{M} : \mathcal{D} \to \mathcal{M}$ is (ϵ, δ)-differentially private under μ-adjacency for $\epsilon \geq 0, \delta \in [0, 1)$, if for all $R \subseteq range(\mathscr{M})$, there holds*

$$\mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in R\big) \leq e^{\epsilon}\mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in R\big) + \delta\,, \quad \forall (\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu). \tag{1}$$

The ideas and algorithms for differential privacy have been successfully applied in large-scale real-world dataset analysis, and to areas ranging from deep learning Shokri & Shmatikov (2015); Abadi et al. (2016); Zhao et al. (2017); Chen et al. (2020) and computer vision Zhu et al. (2020) to control and distributed computation Scaman et al. (2019); Huang et al. (2015).

1.1 Mechanisms over Affine Manifolds: A Motivating Example

We present an example to show that when the domain of secrets $\mathcal{D}$ is an affine manifold, the inherent geometric constraint can not be ignored.

Example 1. Consider a randomized mechanism

$$\mathscr{M}(x_1, x_2) := \begin{bmatrix} x_1 + \gamma_1 \\ x_2 + \gamma_2 \end{bmatrix} \tag{2}$$

where γ_1, γ_2 denote the added random noises for protecting privacy of (x_1, x_2), which are i.i.d. drawn from a Laplace distribution with zero mean and variance σ_γ^2. Let $\sigma_\gamma = \mu/\epsilon$. There are two cases.

(i) [Differential Privacy over Euclidean Space] Let $\mathcal{D} = \mathbb{R}^2$. The $\mathscr{M}$ is a standard Laplace mechanism and preserves the $(\epsilon, 0)$-differential privacy of (x_1, x_2) (Dwork et al. (2014)).

(ii) [Differential Privacy over Affine Manifolds] Let

$$\mathcal{D} := \{(x_1, x_2)^\top \in \mathbb{R}^2 : x_1 - kx_2 = b\}, , \quad \text{with } k \in \mathbb{R}_{>0} \text{ and } b \in \mathbb{R}.$$

Thus, x_1 (or x_2) is indeed released twice at the output of $\mathscr{M}$, i.e., $x_1+\gamma_1$ and $(x_1-b)/k+\gamma_2$ (or $kx_2 + b + \gamma_1$ and $x_2 + \gamma_2$). Then it can be seen that $((1 + \frac{1}{k})\epsilon, 0)$-differential privacy of x_1 and $((1 + k)\epsilon, 0)$-differential privacy of x_2 are achieved tightly, by the sequential composability property of differential privacy (Dwork et al. (2014)). This implies that the mechanism $\mathscr{M}$ is now $(g\epsilon, 0)$-differentially private with $g := \max\{1 + k, 1 + \frac{1}{k}\}$ over $\mathcal{D}$.

For Case (ii), we may design probabilistically correlated zero-mean Laplacian noises as $\gamma_1 = k\gamma_2$ with γ_2 having variance $\sigma_\gamma = \mu/\epsilon$ under which $(\epsilon, 0)$-differential privacy is also achieved. □

Example 1 is in line with the privacy frameworks of Pufferfish Kifer & Machanavajjhala (2014) and Blowfish He et al. (2014), where the privacy of data with correlation/constraint is studied. In particular, the affine manifold can be viewed as constraints about the data known by adversaries. In Blowfish He et al. (2014), added protection against adversaries who know this constraint is shown to be possible by specialized policies. It is of interest to understand how the geometry of the affine manifold can be taken into consideration for possible added protections. Besides being an interesting theoretical study, there are also practical motivations for exploring the affine-manifold geometry in differential privacy mechanisms. Particularly, in dynamical systems the system state trajectories always imply a manifold dependency from the system dynamics.

1.2 Problem Definition

We consider the following mechanism with linear queries

$$\mathscr{M}(\mathbf{x}) = \mathbf{F}\mathbf{x} + \boldsymbol{\gamma} \tag{3}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the input data, $\mathbf{F}$ is a matrix in $\mathbb{R}^{m\times n}$, and $\boldsymbol{\gamma} \in \mathbb{R}^m$ is a randomly drawn noise.

Definition 3. *Denote*

$$\mathcal{C}_d := \left\{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{D}\mathbf{x} + \mathbf{b} = 0\right\} \tag{4}$$

as an affine manifold in $\mathbb{R}^n$, where $\mathbf{D} \in \mathbb{R}^{q\times n}$ and $\mathbf{b} \in \mathbb{R}^q$. The mechanism (3) is a mechanism over $\mathcal{C}_d$ if the input data $\mathbf{x}$ is always taken from $\mathcal{C}_d$, and both $\mathbf{D}$ amd $\mathbf{b}$ are public.

Let $\mathrm{V} = \{1, \ldots, n\}$. Without loss of generality, we assume rank $\left(\begin{bmatrix} \mathbf{D} \\ \mathbf{e}_i^\top \end{bmatrix}\right) = q + 1$ for all $i \in \mathrm{V}$. This indeed indicates that $\mathbf{D}$ is full-row-rank, i.e., rank $(\mathbf{D}) = q$ and none of entries in data $\mathbf{x}$ is identifiable from the manifold $\mathcal{C}_d$. With rank $(\mathbf{D}) = q$, there exists a finite number, saying $l \in \mathbb{N}_+$, of sets $\mathrm{d}_j := \{d_{j,1}, \ldots, d_{j,q}\} \subseteq \mathrm{V}$, $j \in \mathcal{J} := \{1, \ldots, l\}$ such that cardinality $|\mathrm{d}_j| = q$, and matrix $\mathbf{D}_{\mathrm{d}_j} \in \mathbb{R}^{q\times q}$ is nonsingular. We denote the set $-\mathrm{d}_j = \mathrm{V}/\mathrm{d}_j$ and can rewrite the manifold $\mathcal{C}_d$ as

$$\mathcal{C}_d^j := \left\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}_{\mathrm{d}_j} = -\mathbf{D}_{\mathrm{d}_j}^{-1}\mathbf{D}_{-\mathrm{d}_j}\mathbf{x}_{-\mathrm{d}_j} - \mathbf{D}_{\mathrm{d}_j}^{-1}\mathbf{b}\right\}, \quad \forall j \in \mathcal{J}.$$

Definition 4. *(*Manifold μ-Adjacency*) We say* $\mathbf{x}$ *and* $\mathbf{x}'$ *to be* μ*-adjacent over the manifold* $\mathcal{C}_d$*, denoted by* $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$*, if there exist* $j \in \mathcal{I}$ *and* $i \in -\mathrm{d}_j$ *such that*

$$\begin{aligned} &|x_i - x_i'| \leq \mu \\ &x_k = x_k', \quad \forall k \in -\mathrm{d}_j/\{i\} \\ &\mathbf{x}_{\mathrm{d}_j} - \mathbf{x}'_{\mathrm{d}_j} = -\mathbf{D}_{\mathrm{d}_j}^{-1}\mathbf{D}_{-\mathrm{d}_j}(\mathbf{x}_{-\mathrm{d}_j} - \mathbf{x}'_{-\mathrm{d}_j})\,. \end{aligned} \tag{5}$$

Remark 1. *Note that, Definition 1 requires the databases* $\mathbf{x}, \mathbf{x}'$ *to be distinct at only one entry (i.e.,* $\|\mathbf{x} - \mathbf{x}'\|_0 = 1$*) to capture individual's contribution to the database. However, for any two databases* $\mathbf{x}, \mathbf{x}' \in \mathcal{C}_d$*, it may be impossible for them to differ at only one entry (see Example 1). With* $\mathrm{rank}(\mathbf{D}) = q$*,* $\mathbf{x}, \mathbf{x}' \in \mathcal{C}_d$ *may even differ for* $q + 1$ *entries. This means individual's contribution can no longer be looked into separately, but in groups, forming the reasoning behind Definition 4.*

Remark 2. *There has been a line of work on differential privacy over metric spaces Holohan et al. (2015); Alvim et al. (2018); Fernandes (2021). For example, Holohan et al. (2015) has considered databases with entries over a metric space* $(\mathcal{U}, \rho)$*, where* ρ *is a metric Holohan et al. (2015). In this way, an* n*-dimensional database takes values from* $\mathcal{D}^n$ *with* $\mathcal{D} \in \mathcal{U}$ *Holohan et al. (2015). Note that the affine manifold* $\mathcal{C}_d \subseteq \mathbb{R}^n$ *in the current study can not be written in the form of a Cartesian product* $\mathcal{D}^n$*. Therefore, the affine manifold is an extension, but not a special case of the differential privacy frameworks over metric spaces. With* $\mathcal{C}_d$*, now there is inherent geometry between the entries in a database, while a database in* $\mathcal{D}^n$ *has homogeneous geometry on individual entries.*

Remark 3. *Note that the adjacency in Definition 4 is defined by modifying entries in the dataset. It is also of interest to investigate the adjacency notion by adding or removing records. It can be seen that the dimension of the resulting* $\mathbf{x}'$ *by adding or removing records in* $\mathbf{x}$ *may be incompatible with matrices* $\mathbf{D}$ *and* $\mathbf{F}$*, violating the manifold constraint (4) and the mechanism (3), respectively. If the manifold* $\mathcal{C}_d$ *can be separated into several independent sub-manifolds* $\mathbf{D}_i\mathbf{x}_i + \mathbf{b}_i = 0$*, the adjacency by adding or removing records* $\mathbf{x}_i$ *can be well-defined without manifold violation, though the mechanism violation problem is still unsolved. Thus the adjacency by adding or removing records is not applicable to our scenario of linear mechanism (3) with affine manifold (4).*

Definition 5. *The randomized mechanism* $\mathscr{M}$ *in (3) is* (ϵ, δ)*-differentially over* $\mathcal{C}_d$ *for* $\epsilon \geq 0$ *and* $\delta \in [0, 1)$*, if for all* $R \subseteq range(\mathscr{M})$*,*

$$\mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in R\big) \leq e^{\epsilon}\mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in R\big) + \delta\,, \quad \forall (\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)\,. \tag{6}$$

An implication of the differentially private mechanism $\mathscr{M}$ in (3), lies in the fact that entries in the random noise $\boldsymbol{\gamma}$ may be probabilistically dependent. We introduce the following definition.

Definition 6. *The noise* $\boldsymbol{\gamma}$ *is probabilistically structured if there exists* $\Lambda \in \mathbb{R}^{m\times r}$ *with* rank $(\Lambda) = r \leq m$ *such that* $\boldsymbol{\gamma} = \Lambda\boldsymbol{\eta}$*, where the entries in* $\boldsymbol{\eta} \in \mathbb{R}^r$ *are i.i.d..*

The entries in $\boldsymbol{\eta} \in \mathbb{R}^r$ may be from a standard Gaussian or Laplace distribution, i.e., $\boldsymbol{\eta} \sim \mathcal{N}(0, 1)^r$ for Gaussian mechanism and $\boldsymbol{\eta} \sim \mathcal{L}(0, 1)^r$ for Laplace mechanism. The rank condition for Λ is without loss of generality as it guarantees a minimal value for the dimension of the random vector $\boldsymbol{\eta}$ for a concise investigation of the mechanism.

1.3 Contributions and Related Work

Contributions. Theories in necessary and sufficient conditions regarding whether a prescribed differential privacy level can be achieved with structured (non i.i.d.) noise injection are established, respectively, for Gaussian mechanism and Laplace mechanism. For any expected privacy level, systemic approaches are developed for realizing the privacy budget sufficiently and tightly with structured noise injection. As a result, a systemic framework for differential privacy over affine manifolds with linear queries has been established. The obtained theories and proposed approaches are applied to the cloud-based control problem for feedback systems. In this application, the affine manifold is shown to have naturally arisen and the framework developed in this work becomes effective in differential privacy analysis and noise injection mechanism design.

Related work. Differential privacy, proposed in Dwork et al. (2006b), is a rigorous notion for defining and preserving data privacy. In the last decades, extensive developments have been emerged

in such as the mechanism design McSherry & Talwar (2007) and applications to machine learning Chaudhuri et al. (2011) and deep learning Abadi et al. (2016), etc, advancing the differential privacy as a gold standard in data privacy. Data correlation may be an important factor influencing privacy leakage and thus cannot be ignored when designing or analyzing differentially private mechanisms Kifer & Machanavajjhala (2011); Takbiri et al. (2020). Generalizations of differential privacy have been developed for probabilistically correlated data, e.g., the Pufferfish privacy proposed in Kifer & Machanavajjhala (2012; 2014) by employing the notions of discriminative pairs of secrets and data evolution scenarios that may incorporate the data correlation. In the Pufferfish framework, domain experts are allowed, specifying the knowledge of e.g. the set of potential secrets and data correlation, which customizes the framework to the needs of given applications. Along the line of Pufferfish, many developments have been witnessed recently, such as Pufferfish privacy mechanisms for correlated data by a Bayesian network in Song et al. (2017) and the composition properties for time-series data in Song & Chaudhuri (2017). The potential interdependency between data may also be characterized as deterministic relationships. In He et al. (2014), the Blowfish privacy, as a generalization of the differential privacy and inspired by the Pufferfish framework, is established for the data of deterministic constraints/correlation on the notion of policy, specifying the secrets and constraints that may be known about the data. As a result, this enables to introduce a new notion of adjacency incorporating the deterministic constraints about the database, such as count query constraint and marginal constraint. In this respect, this paper can be regarded as a generalization of the Blowfish framework to the data with affine-manifold constraint.

Differential privacy has also been extended to metric spaces where the domain of secrets can be from any space with a metric Holohan et al. (2015), where the space can be discrete, continuous, or even functional. Our work is also an extension of this line of research but with distinctive features as highlighted in Remark 2. Calibrating noises plays a significant role in maximizing the data utility while preserving the desired differential privacy. In the seminal work Dwork et al. (2006b), the notion of sensitivity is introduced to characterize the deviation of Laplace noises ensuring the $(\epsilon, 0)$-differential privacy. Along this line, Nissim et al. (2007) shows that the utility can be improved by employing data-dependent noises calibrated to the smooth sensitivity. In Balle & Wang (2018) the necessary and sufficient condition for the Gaussian mechanism is established from the perspective of privacy loss, yielding the optimal Gaussian noises for (ϵ, δ)-differential privacy. For the differential privacy of functional data, the correct noise level is studied in Hall et al. (2013) by establishing a measure of sensitivity in the reproducing kernel Hilbert space norm. Taking the noise distribution into consideration, Geng & Viswanath (2015) shows that the staircase noise distribution is optimal for $(\epsilon, 0)$-differential privacy and the uniform noise distribution is near-optimal for $(0, \delta)$-differential privacy. In this paper, besides applying the sensitivity as in these results, we also take the noise structure as a key factor to calibrate the noises. By injecting appropriate correlated noises, it is shown that the resulting utility may be improved compared to injecting i.i.d. noises. In view of the adopted calibration tool using noise structure, this paper is closely related to the results in Li et al. (2015) where the matrix mechanism is designed to answer a workload of linear counting queries with correlated noises to improve the utility, and also in Chanyaswad et al. (2018) where the matrix-variate Gaussian mechanism is calibrated by adding a matrix-valued noise drawn from a matrix-variate Gaussian distribution. Besides, our paper can be regarded as an extension of the both works by studying the vector-valued mechanism with affine-manifold dependency data.

Notation. Denote by $\mathbb{R}$ the real numbers, $\mathbb{R}^n$ the real space of n dimension and $\mathbb{N}$ the set of natural numbers. For $\mathbf{x} \in \mathbb{R}^n$, denote x_i as the i-th entry of $\mathbf{x}$, $\|\mathbf{x}\|_0$, $\|\mathbf{x}\|_1$ and $\|\mathbf{x}\|$ as the 0, 1, and 2-norm of vector $\mathbf{x}$, respectively, and for any set $\mathrm{p} \subseteq \{1, \ldots, n\}$ of l elements, $\mathbf{x}_\mathrm{p}$ a vector of dimension l with each entry as x_j with $j \in \mathrm{p}$. Denote $\mathbf{e}_i$ a basis vector of dimension n whose entries are all zero expect the i-th as one. For matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and set $\mathrm{p} := \{p_1, \ldots, p_l\} \subseteq \{1, \ldots, n\}$ of l elements, $\mathbf{A}_\mathrm{p}$ a matrix of dimension $m \times l$ with each column being the p_i-th column of $\mathbf{A}$ with $i \in \mathrm{p}$, and we denote $\mathbf{E}_\mathrm{p}$ as a matrix of dimension $m \times l$ with each column being the basis vector $\mathbf{e}_i$, $i \in \mathrm{p}$.

2 Differential Privacy Conditions

In this section, we present conditions for the mechanism $\mathscr{M}$ in (3) to achieve differential privacy, under Gaussian and Laplacian mechanisms, respectively.

2.1 GAUSSIAN MECHANISM

First of all, we study the case with the entries in $\boldsymbol{\eta} \in \mathbb{R}^r$ drawn from a standard Gaussian distribution, i.e. $\boldsymbol{\eta} \sim \mathcal{N}(0,1)^r$. We introduce $\Lambda^\dagger = (\Lambda^\top \Lambda)^{-1}\Lambda^\top$, $\Psi_j = \mathbf{I}_n - \mathbf{E}_{\mathrm{d_j}}(\mathbf{D}_{\mathrm{d_j}})^{-1}\mathbf{D}$ for $j \in \mathcal{I}$, and $\Delta_i^{\mathcal{N}} = \max_{j\in\mathcal{I}} \|\Lambda^\dagger \mathbf{F}\Psi_j \mathbf{E}_{-\mathrm{d_j}}\mathbf{E}_{-\mathrm{d_j}}^\top \mathbf{e}_i\|$ for $i \in \mathrm{V}$. Denote $\mathbf{D}^\perp \in \mathbb{R}^{n\times(n-q)}$ as a matrix such that $\mathbf{D}\mathbf{D}^\perp = 0$ and rank $\left(\begin{bmatrix}\mathbf{D}^\top & \mathbf{D}^\perp\end{bmatrix}\right) = n$, and define $\Delta_{\mathcal{N}} := \max_{i\in\mathrm{V}}\{\Delta_i^{\mathcal{N}}\}$ and $\Phi(s) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{s} e^{-\tau^2/2} d\tau$. We present the following result.

Theorem 1. *The mechanism $\mathscr{M}$ in (3) with $\boldsymbol{\eta} \sim \mathcal{N}(0,1)^r$ achieves (ϵ,δ)-differential privacy under μ-adjacency over $\mathcal{C}_d$ if and only if there hold*

$$\text{rank }(\Lambda) = \text{rank }\left(\begin{bmatrix}\Lambda & \mathbf{F}\mathbf{D}^\perp\end{bmatrix}\right) = r\,; \tag{7}$$

$$\Phi\left(\frac{\mu\Delta_{\mathcal{N}}}{2} - \frac{\epsilon}{\mu\Delta_{\mathcal{N}}}\right) - e^{\epsilon}\Phi\left(-\frac{\mu\Delta_{\mathcal{N}}}{2} - \frac{\epsilon}{\mu\Delta_{\mathcal{N}}}\right) \le \delta\,. \tag{8}$$

2.2 LAPLACE MECHANISM

Next, we consider the Laplace mechanism with $\boldsymbol{\eta} \backsim \mathcal{L}(0,1)^r$, and study the $(\epsilon,0)$-differential privacy of the mechanism $\mathscr{M}(\mathbf{x})$ over $\mathcal{C}_d$. To this end, define $\Delta_{\mathcal{L}} := \max_{i\in\mathrm{V}}\{\Delta_i^{\mathcal{L}}\}$ with $\Delta_i^{\mathcal{L}} = \max_{j\in\mathcal{I}} \|\Lambda^\dagger \mathbf{F}\Psi_j \mathbf{E}_{-\mathrm{d_j}}\mathbf{E}_{-\mathrm{d_j}}^\top \mathbf{e}_i\|_1$ for $i \in \mathrm{V}$.

Theorem 2. *The mechanism $\mathscr{M}$ in (3) with $\boldsymbol{\eta} \sim \mathcal{L}(0,1)^r$ achieves $(\epsilon,0)$-differential privacy under μ-adjacency over $\mathcal{C}_d$ if and only if there hold* (7) *and*

$$\Delta_{\mathcal{L}} \le \epsilon/\mu\,. \tag{9}$$

In Theorems 1 and 2, (7) implies the least amount of independent standard Gaussian/Laplace noises (i.e., $r \ge$ rank $(\mathbf{F}\mathbf{D}^\perp)$) and provides a structural property of the noise matrix Λ for the differential privacy; (8) and (9) quantify the privacy levels that can be achieved by the amount of injected Gaussian and Laplace noises, respectively.

3 STRUCTURED MECHANISM DESIGN

In this section, we show how the conditions in Theorem 1 and Theorem 2 will inform efficient structured randomization design.

3.1 STRUCTURED GAUSSIAN MECHANISM DESIGN

In this subsection, we discuss for a fixed Gaussian mechanism $\mathscr{M}$ and a given privacy budget (ϵ,δ), how we can design the structured randomization Λ so that $\mathscr{M}$ achieves (ϵ,δ)-differential privacy under μ-adjacency over the affine manifold $\mathcal{C}_d$.

In Algorithm 1, we present a design approach of the Gaussian noise $\boldsymbol{\gamma} := \Lambda\boldsymbol{\eta}$ by applying the two conditions (7) and (8) in Theorem 1 such that the $\mathscr{M}$ achieves the prescribed (ϵ,δ)-differential privacy under μ-adjacency over $\mathcal{C}_d$.

Algorithm 1: Structured Gaussian Noise Design Algorithm

Input: Privacy levels $\epsilon \ge 0, \delta > 0, \mu > 0$.

1. Let $\boldsymbol{\eta} \sim \mathcal{N}(0,1)^r$ with $r =$ rank $(\mathbf{F}\mathbf{D}^\perp)$;
2. Let $\bar{\Lambda} \in \mathbb{R}^{n\times r}$ be a matrix sharing the same column space with $\mathbf{F}\mathbf{D}^\perp$;
3. Let σ be such that

$$\Phi\Big(\frac{\mu\bar{\Delta}_{\mathcal{N}}}{2\sigma} - \frac{\epsilon\sigma}{\mu\bar{\Delta}_{\mathcal{N}}}\Big) - e^{\epsilon}\Phi\Big(-\frac{\mu\bar{\Delta}_{\mathcal{N}}}{2\sigma} - \frac{\epsilon\sigma}{\mu\bar{\Delta}_{\mathcal{N}}}\Big) \le \delta \tag{10}$$

with

$$\bar{\Delta}_{\mathcal{N}} := \max_{(i,j)\in\mathrm{V}\times\mathcal{I}} \|\bar{\Lambda}^\dagger \mathbf{F}\Psi_j \mathbf{E}_{-\mathrm{d_j}}\mathbf{E}_{-\mathrm{d_j}}^\top \mathbf{e}_i\|\,. \tag{11}$$

Output: $\boldsymbol{\gamma} = \sigma\bar{\Lambda}\boldsymbol{\eta}$.

It is clear that with $\Lambda := \sigma\bar{\Lambda}$ following the steps 2 and 3 in Algorithm 1, the rank constraint (7) and the inequality (8) are both satisfied. Regarding the design of $\bar{\Lambda}$ at the step 2, it can be easily achieved by computing the basis vectors spanning the column space of $\mathbf{F}\mathbf{D}^{\perp}$ and then assigning each column of $\bar{\Lambda}$ with a basis vector. As for the design of σ from (10), one can either adopt the numerical design algorithm in Balle & Wang (2018), or disregard the second negative term on the left side of (10), and use an analytical but less tight lower bound as $\sigma \geq (\mu\bar{\Delta}_{\mathcal{N}})/(\sqrt{\Phi^{-2}(\delta) + 2\epsilon} + \Phi^{-1}(\delta))$.

3.2 Structured Laplace Mechanism Design

We now turn to the Laplace mechanism $\mathscr{M}$, and propose a design approach of the random perturbation $\boldsymbol{\gamma} := \Lambda\boldsymbol{\eta}$ in Algorithm 2 for $(\epsilon, 0)$-differential privacy by Theorem 2.

Algorithm 2: Structured Laplace Noise Design Algorithm

Input: Privacy levels $\epsilon \geq 0, \delta = 0, \mu > 0$.

1. Let $\boldsymbol{\eta} \sim \mathcal{L}(0,1)^r$ with $r = \text{rank } (\mathbf{F}\mathbf{D}^{\perp})$;
2. Let $\bar{\Lambda} \in \mathbb{R}^{n\times r}$ be a matrix sharing the same column space with $\mathbf{F}\mathbf{D}^{\perp}$;
3. Let σ be such that
$$\sigma \geq \mu\bar{\Delta}_{\mathcal{L}}/\epsilon \tag{12}$$
with
$$\bar{\Delta}_{\mathcal{L}} := \max_{(i,j)\in \mathrm{V}\times\mathcal{J}} \|\bar{\Lambda}^{\dagger}\mathbf{F}\Psi_j\mathbf{E}_{-\mathrm{d}_j}\mathbf{E}_{-\mathrm{d}_j}^{\top}\mathbf{e}_i\|_1 \,. \tag{13}$$

Output: $\boldsymbol{\gamma} = \sigma\bar{\Lambda}\boldsymbol{\eta}$.

Following Algorithm 2, it can be verified that the resulting noise matrix $\Lambda := \sigma\bar{\Lambda}$ fulfills both requirements (7) and (9) in Theorem 2, achieving the desired $(\epsilon, 0)$-differential privacy over $\mathcal{C}_d$.

4 Application: Differentially Private Control

Emerging applications in cyber-physical systems such as smart girds and intelligent transportations have inspired cloud-based control system paradigms, where a dynamical system sends its output to a cloud, and receives feedback control decisions from the cloud Tanaka et al. (2017). The benefit of cloud-based control is promise in improved control accuracy and system performance since the cloud holds more information about the environment and other systems; one particular cost of cloud-based control is leak of private state trajectories to adversaries eavesdropping the communication between the system and the cloud.

4.1 Cloud-based Control Systems

Consider the cloud-based control systems of the form

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) \end{aligned} \tag{14}$$

where the *privacy-sensitive* system state $\mathbf{x}(t) \in \mathbb{R}^{n_x}$, the control input $\mathbf{u}(t) \in \mathbb{R}^{n_u}$, the system output/measurement $\mathbf{y}(t) \in \mathbb{R}^{n_y}$. In the cloud-enabled setup, the system transmits $\mathbf{y}(t)$ to the cloud which computes a feedback control signal $\mathbf{u}(t)$ and then sends it back to the system for implementation. When the communication networks between the system and the cloud are eavesdropped by adversaries, information about $\mathbf{x}(t)$ might be inferred.

We propose to perturb the output with random noises $\gamma(t)$ and submit to the cloud the perturbed output

$$\hat{\mathbf{y}}(t) := \mathbf{C}\mathbf{x}(t) + \gamma(t) \,. \tag{15}$$

See Fig. 1 for the considered cloud-based control scheme. In the following, our goal is to design the random noises $\gamma(t)$ by employing the established results in Section 2 to achieve the desired differential privacy of system states, while the readers of interest in the remainder controller design can refer to, e.g. linear–quadratic–Gaussian (LQG) control Åström (2012) and neural network control Ge et al. (2013). As Gaussian noises are more convenient and common for tackling in these control

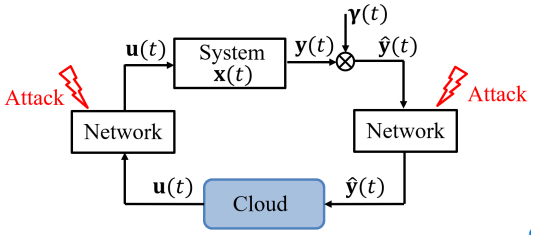

Figure 1: Differentially private cloud-based control scheme (e.g., Tanaka et al. (2017)).

techniques, we choose the random noises $\gamma(t)$ as structured Gaussian noises and thus pursue the (ϵ, δ)-differential privacy with $\delta > 0$.

4.2 Differentially Private Cloud-based Control

Let the system running iterations $T \geq n_x$ and denote the observability matrix

$$\mathbf{O}_T := [\mathbf{C};\, \mathbf{CA};\, \cdots;\, \mathbf{CA}^{T-1}]\,.$$

We impose the following assumption on the system (14).

Assumption 1. *(i). The system (14) is observable, i.e.,* rank $(\mathbf{O}_T) = n_x$*; (ii) The system matrix* $\mathbf{A}$ *in (14) is nonsingular, i.e.,* rank $(\mathbf{A}) = n_x$.

In the following, we apply the previous results to design the injected random noises $\gamma(t)$ such that the differential privacy of the system states trajectory $(\mathbf{x}(t))_{t=0}^{T-1}$ is achieved under the desired privacy level (ϵ, δ, μ).

Affine-Manifold Constraint from Dynamics. We define $n := Tn_x$ and the private data $\mathbf{x} := \big[\mathbf{x}(0); \mathbf{x}(1); \ldots; \mathbf{x}(T-1)\big] \in \mathbb{R}^n$, and can obtain the mechanism $\mathscr{M}$ as

$$\mathscr{M}(\mathbf{x}) = \mathbf{F}\mathbf{x} + \boldsymbol{\gamma} \tag{16}$$

where $\mathbf{F} = \mathbf{I}_T \otimes \mathbf{C}$ and $\boldsymbol{\gamma} = \big[\gamma(0); \gamma(1); \ldots; \gamma(T-1)\big]$. Note that the private data $\mathbf{x}$ is naturally and deterministically correlated by the $\mathbf{x}$-dynamics of (14), i.e., subject to the affine-manifold constraint

$$\mathcal{C}_d = \{\mathbf{x} : \mathbf{D}\mathbf{x} + \mathbf{b} = 0\} \tag{17}$$

with $\mathbf{b} = (\mathbf{I}_{T-1} \otimes \mathbf{B})\mathbf{u}$, $\mathbf{u} := \big[\mathbf{u}(0); \mathbf{u}(1); \ldots; \mathbf{u}(T-2)\big]$, and

$$\mathbf{D} = \begin{bmatrix} \mathbf{A} & -\mathbf{I}_{n_x} & & \\ & \ddots & \ddots & \\ & & \mathbf{A} & -\mathbf{I}_{n_x} \end{bmatrix} \in \mathbb{R}^{(n-n_x)\times n}\,. \tag{18}$$

Here $\mathbf{b}$ is known by the adversaries as the communication messages (i.e., $\hat{\mathbf{y}}(t)$ and $\mathbf{u}(t)$) between the system and the cloud are eavesdropped. Moreover, there holds rank $(\mathbf{D}) = n - n_x$, and by Assumption 1,

$$\text{rank}\left(\begin{bmatrix}\mathbf{D} \\ \mathbf{e}_i^\top\end{bmatrix}\right) = n - n_x + 1\,, \quad \forall i = 1, \ldots, n\,.$$

Structured Noise Mechanism. With the mechanism (16) and the affine manifold (17), we next apply Algorithms 1 to design the random noises $\boldsymbol{\gamma}$ for a (ϵ, δ)-differentially private Gaussian mechanism $\mathscr{M}$. We denote

$$\mathbf{D}^\perp = \big[\mathbf{I}_{n_x};\, \mathbf{A};\, \cdots;\, \mathbf{A}^{T-1}\big]\,, \tag{19}$$

satisfying $\mathbf{D}\mathbf{D}^\perp = 0$ and rank $\left(\begin{bmatrix}\mathbf{D}^\top & \mathbf{D}^\perp\end{bmatrix}\right) = n$. We define

$$\mathcal{O}_{ij} = \mathbf{O}_T \mathbf{A}^{1-i} \mathbf{v}_j\,, \quad (i,j) \in [1,T] \times [1,n_x]\,,$$

where $\mathbf{v}_j$ is a vector of dimension n_x with entries being zero except the j-th being one. Letting $\boldsymbol{\eta} \sim \mathcal{N}(0,1)^{n_x}$ and $\bar{\Lambda} = \mathbf{F}\mathbf{D}^{\perp} = \mathbf{O}_T$, and we design σ such that

$$\Phi\Big(\frac{\mu\bar{\Delta}_{\mathcal{N}}}{2\sigma} - \frac{\epsilon\sigma}{\mu\bar{\Delta}_{\mathcal{N}}}\Big) - e^{\epsilon}\Phi\Big(-\frac{\mu\bar{\Delta}_{\mathcal{N}}}{2\sigma} - \frac{\epsilon\sigma}{\mu\bar{\Delta}_{\mathcal{N}}}\Big) \leq \delta \tag{20}$$

where $\bar{\Delta}_{\mathcal{N}} = \max\limits_{(j,k)\in[1,T]\times[1,n_x]} \big\|\mathbf{A}^{1-j}\mathbf{v}_k\big\|$. See the discussions before Section 3.2 for more details on how to derive such σ from (20). As a result, we design the random noise as $\boldsymbol{\gamma} = \sigma\mathbf{O}_T\boldsymbol{\eta}$, i.e., $\gamma(t) = \sigma\mathbf{C}\mathbf{A}^t\boldsymbol{\eta}\,, \quad$ with $\boldsymbol{\eta} \sim \mathcal{N}(0,1)^{n_x}$ for $t = 0, 1, \ldots, T-1$. Then the following result can be concluded by Theorem 1.

Theorem 3. *Given any privacy levels $\epsilon, \delta, \mu > 0$, let Assumption 1 and the random noise $\gamma(t) = \sigma\mathbf{C}\mathbf{A}^t\boldsymbol{\eta}$ with $\boldsymbol{\eta} \sim \mathcal{N}(0,1)^{n_x}$ and σ satisfying (20) for $t = 0, 1, \ldots, T-1$. Then the system (14) with the perturbed output (15) preserves the (ϵ, δ)-differential privacy of the state trajectories under manifold μ-adjacency.*

5 Numerical Validations

In this section, we provide a numerical example to illustrate the effectiveness of the privacy-preserving clould-based control approach based on the manifold differential privacy.

System setup. Consider the cloud-based control of autonomous vehicles with the dynamical model Hoh et al. (2011); Yazdani et al. (2018) as

$$\begin{array}{rcl} \mathbf{x}(t+1) & = & A\mathbf{x}(t) + B\mathbf{u}(t) \\ \mathbf{y}(t) & = & C\mathbf{x}(t) \end{array} \tag{21}$$

where $\mathbf{x}(t) = [p(t); v(t)]$ with $p(t), v(t)$ as the *private* position and the velocity, respectively, and the system matrices

$$A = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} T_s^2/2 \\ T_s \end{bmatrix}, \quad C = [1 \quad 0]$$

with the sampling period $T_s = 0.1$. In the cloud-based setup, for privacy concern the vehicle sends perturbed output $\hat{\mathbf{y}}(t) = \mathbf{y}(t) + \gamma(t)$ to the cloud, which then delivers to the vehicle for implementation a control signal, of the form [1]

$$\begin{array}{rcl} \mathbf{u}(t) & = & -K_P^{\top}(\hat{\mathbf{x}}(t) - \mathbf{x}_r(t)) \\ \hat{\mathbf{x}}(t+1) & = & A\hat{\mathbf{x}}(t) + B\mathbf{u}(t) + L(\hat{\mathbf{y}}(t) - C\hat{\mathbf{x}}(t)) \end{array} \tag{22}$$

with $K_P = [3.4240; 4.3095]$, $L = [0.8266; 0.6973]$, and the reference trajectory $\mathbf{x}_r(t) := [p_r(t); v_r(t)] = \big[tanh(t); 1 - |tanh(t-9)|\big]$.

Experiments. We note that the $v(t)$-dynamics in (21) is dependent of $v(t), \mathbf{u}(t)$ but independent of the private state $p(t)$. In other words, the privacy-sensitive positions $p(t)$ are affine-manifold-dependent on the $p(t)$-dynamics. In view of this, we apply the previous Gaussian mechanism design in Subsection 4.2 with $\mathbf{A} = 1$ and $\mathbf{C} = 1$, and by Algorithm 1, take $\gamma(t) = \sigma\boldsymbol{\eta}$ with $\boldsymbol{\eta} \sim \mathcal{N}(0,1)$ and σ satisfying

$$\Phi\Big(\frac{\mu}{2\sigma} - \frac{\epsilon\sigma}{\mu}\Big) - e^{\epsilon}\Phi\Big(-\frac{\mu}{2\sigma} - \frac{\epsilon\sigma}{\mu}\Big) \leq \delta\,.$$

In simulations, we let the running time $T = 100$ and the desired privacy levels $(\epsilon, \delta, \mu) = (1, 10^{-2}, 1)$, and then design $\gamma(t) = \sigma\eta$ with the standard Gaussian noise $\boldsymbol{\eta} \sim \mathcal{N}(0,1)$ and $\sigma = 2.5244$.

[1] Here for simplicity we choose a classical output-feedback controller. However, it is noted that other methods such as LQG control Åström (2012) and neural network control Ge et al. (2013) are also applicable with no influence on the design of noises $\boldsymbol{\gamma}(t)$ for privacy preservation.

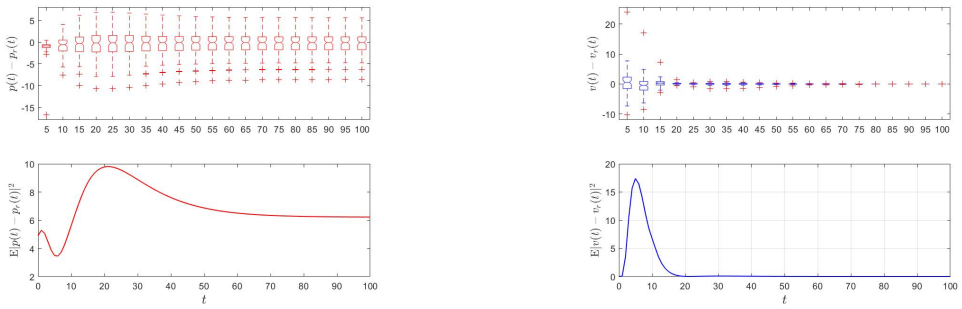

Figure 2: Tracking errors $p(t) - p_r(t)$ (left) and $v(t) - v_r(t)$ (right).

Results. The resulting tracking errors are presented in Fig. 2, where the tracking velocity error is asymptotically vanishing, while the tracking position error is not vanishing in the mean-square sense due to presence of noises for privacy preservation. We also show the relationship between the tracking performance and the privacy requirements. The simulation results are presented in Fig. 3. It is clear that a higher privacy level (i.e., a smaller ϵ) leads to larger tracking errors of both position and velocity in the mean square senses, demonstrating the trade-off between the data utility and the differential privacy guarantee. Besides, in Fig. 3 we also compare the tracking performance by calibrating the noises via the proposed structured noise injection approach ($r = 1$) to the common approach of using i.i.d. noises ($r = T$). It can be seen that the calibrated noises following Algorithm 1 shows better tracking performances under the same privacy requirement, implying that the proposed structured noise injection approach may improve the data utility.

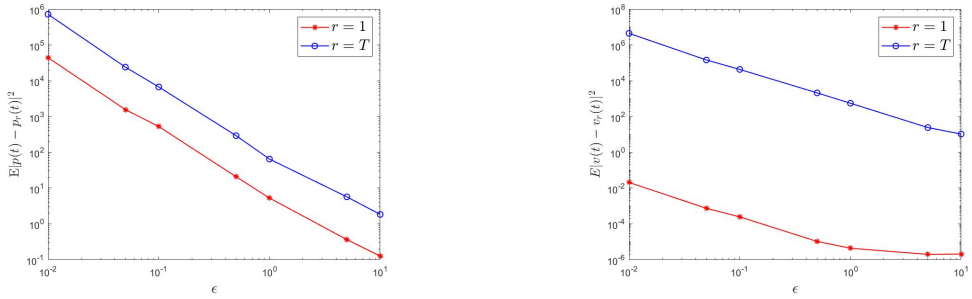

Figure 3: Mean-square tracking errors $p(t) - p_r(t)$ (left) and $v(t) - v_r(t)$ (right) under different privacy requirements ϵ and structured noises ($r = 1$ for the structured noises following Algorithm 1 and $r = T$ for i.i.d. noises).

Reproduction of the results. The code used for producing this numerical example is provided in the supplementary material.

6 Conclusions

We have studied differential privacy for mechanisms generated by linear queries over affine manifolds. We established necessary and sufficient conditions on whether differential privacy can be achieved when the affine manifolds encode geometry of the entries in a dataset. The derived framework was applied to differentially private cloud-based control, for which the affine-manifold data dependency has been encoded in the systems themselves. In future work, it would be interesting to further investigate extending these results to general differentially private mechanisms with nonlinear queries, and the resulting theories will then have direct applicability in applications such as differentially private deep learning.

REFERENCES

Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pp. 308–318, 2016.

Mário Alvim, Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Anna Pazii. Local differential privacy on metric spaces: optimizing the trade-off with utility. In *2018 IEEE 31st Computer Security Foundations Symposium (CSF)*, pp. 262–267. IEEE, 2018.

Karl J Åström. *Introduction to stochastic control theory*. Courier Corporation, 2012.

Borja Balle and Yu-Xiang Wang. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *International Conference on Machine Learning*, pp. 394–403. PMLR, 2018.

Thee Chanyaswad, Alex Dytso, H Vincent Poor, and Prateek Mittal. Mvg mechanism: Differential privacy under matrix-valued query. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pp. 230–246, 2018.

Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(3), 2011.

Dingfan Chen, Tribhuvanesh Orekondy, and Mario Fritz. Gs-wgan: A gradient-sanitized approach for learning differentially private generators. *Advances in Neural Information Processing Systems (NeurIPS 2020)*, 2020.

Dorothy E Denning and Peter J Denning. The tracker: A threat to statistical database security. *ACM Transactions on Database Systems (TODS)*, 4(1):76–96, 1979.

Dorothy E Denning and Jan Schlörer. A fast procedure for finding a tracker in a statistical database. *ACM Transactions on Database Systems (TODS)*, 5(1):88–102, 1980.

Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 202–210, 2003.

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pp. 486–503. Springer, 2006a.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pp. 265–284. Springer, 2006b.

Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

Natasha Fernandes. *Differential privacy for metric spaces: information-theoretic models for privacy and utility with new applications to metric domains*. PhD thesis, École Polytechnique Paris; Macquarie University, 2021.

Shuzhi Sam Ge, Chang C Hang, Tong H Lee, and Tao Zhang. *Stable adaptive neural network control*, volume 13. Springer Science & Business Media, 2013.

Quan Geng and Pramod Viswanath. Optimal noise adding mechanisms for approximate differential privacy. *IEEE Transactions on Information Theory*, 62(2):952–969, 2015.

Rob Hall, Alessandro Rinaldo, and Larry Wasserman. Differential privacy for functions and functional data. *The Journal of Machine Learning Research*, 14(1):703–727, 2013.

Xi He, Ashwin Machanavajjhala, and Bolin Ding. Blowfish privacy: Tuning privacy-utility tradeoffs using policies. In *Proceedings of the 2014 ACM SIGMOD international conference on Management of data*, pp. 1447–1458, 2014.

Baik Hoh, Toch Iwuchukwu, Quinn Jacobson, Daniel Work, Alexandre M Bayen, Ryan Herring, Juan-Carlos Herrera, Marco Gruteser, Murali Annavaram, and Jeff Ban. Enhancing privacy and accuracy in probe vehicle-based traffic monitoring via virtual trip lines. *IEEE Transactions on Mobile Computing*, 11(5):849–864, 2011.

Naoise Holohan, Douglas J Leith, and Oliver Mason. Differential privacy in metric spaces: Numerical, categorical and functional data under the one roof. *Information Sciences*, 305:256–268, 2015.

Zhenqi Huang, Sayan Mitra, and Nitin Vaidya. Differentially private distributed optimization. In *Proceedings of the 2015 International Conference on Distributed Computing and Networking*, pp. 1–10, 2015.

Daniel Kifer and Ashwin Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, pp. 193–204, 2011.

Daniel Kifer and Ashwin Machanavajjhala. A rigorous and customizable framework for privacy. In *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGAI symposium on Principles of Database Systems*, pp. 77–88, 2012.

Daniel Kifer and Ashwin Machanavajjhala. Pufferfish: A framework for mathematical privacy definitions. *ACM Transactions on Database Systems (TODS)*, 39(1):1–36, 2014.

Chao Li, Gerome Miklau, Michael Hay, Andrew McGregor, and Vibhor Rastogi. The matrix mechanism: optimizing linear counting queries under differential privacy. *The VLDB journal*, 24(6): 757–781, 2015.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pp. 94–103. IEEE, 2007.

Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pp. 75–84, 2007.

Junfei Qiu, Qihui Wu, Guoru Ding, Yuhua Xu, and Shuo Feng. A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, 2016(1):1–16, 2016.

Kevin Scaman, Francis Bach, Sébastien Bubeck, Yin Lee, and Laurent Massoulié. Optimal convergence rates for convex distributed optimization in networks. *Journal of Machine Learning Research*, 20:1–31, 2019.

Reza Shokri and Vitaly Shmatikov. Privacy-preserving deep learning. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, pp. 1310–1321, 2015.

Shuang Song and Kamalika Chaudhuri. Composition properties of inferential privacy for time-series data. In *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 814–821. IEEE, 2017.

Shuang Song, Yizhen Wang, and Kamalika Chaudhuri. Pufferfish privacy mechanisms for correlated data. In *Proceedings of the 2017 ACM International Conference on Management of Data*, pp. 1291–1306, 2017.

Nazanin Takbiri, Amir Houmansadr, Dennis L Goeckel, and Hossein Pishro-Nik. Privacy of dependent users against statistical matching. *IEEE Transactions on Information Theory*, 66(9): 5842–5865, 2020.

Takashi Tanaka, Mikael Skoglund, Henrik Sandberg, and Karl Henrik Johansson. Directed information and privacy loss in cloud-based control. In *2017 American Control Conference (ACC)*, pp. 1666–1672. IEEE, 2017.

Nan Wu, Farhad Farokhi, David Smith, and Mohamed Ali Kaafar. The value of collaboration in convex machine learning with differential privacy. In *2020 IEEE Symposium on Security and Privacy (SP)*, pp. 304–317. IEEE, 2020.

Kasra Yazdani, Austin Jones, Kevin Leahy, and Matthew Hale. Differentially private lq control. *arXiv preprint arXiv:1807.05082*, 2018.

Jun Zhao, Junshan Zhang, and H Vincent Poor. Dependent differential privacy for correlated data. In *2017 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–7. IEEE, 2017.

Junyi Zhu and Matthew B Blaschko. R-gap: Recursive gradient attack on privacy. In *International Conference on Learning Representations*, 2020.

Yuqing Zhu, Xiang Yu, Manmohan Chandraker, and Yu-Xiang Wang. Private-knn: Practical differential privacy for computer vision. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11854–11862, 2020.

A PROOFS

A.1 PROOF OF THEOREM 1

A.1.1 A TECHNICAL LEMMA

In this subsection, we present the necessary and sufficient condition from the perspective of privacy loss Balle & Wang (2018) for (ϵ, δ)-differential privacy of the mechanism (3) over the affine manifold $\mathcal{C}_d$.

Given any $\mathbf{x}, \mathbf{x}'$, we denote the random variables $\mathbf{Y} = \mathscr{M}(\mathbf{x})$ and $\mathbf{Y}' = \mathscr{M}(\mathbf{x}')$, whose probability density function are given by $g(\mathbf{y} - \mathbf{F}\mathbf{x})$ and $g(\mathbf{y}' - \mathbf{F}\mathbf{x}')$, respectively. Then, we define the privacy loss random variables $L_{\mathbf{x},\mathbf{x}'} = \ell_{\mathbf{x},\mathbf{x}'}(\mathbf{Y})$ and $L_{\mathbf{x}',\mathbf{x}} = \ell_{\mathbf{x}',\mathbf{x}}(\mathbf{Y}')$, where $\ell_{\mathbf{x},\mathbf{x}'}$ is the privacy loss function of mechanism $\mathscr{M}$ on a pair of adjacent $\mathbf{x}, \mathbf{x}'$ as

$$\ell_{\mathbf{x},\mathbf{x}'}(\mathbf{y}) = \log\left(\frac{g(\mathbf{y} - \mathbf{F}\mathbf{x})}{g(\mathbf{y} - \mathbf{F}\mathbf{x}')}\right).$$

In view of the previous definitions, the mechanism output random variable $\mathbf{Y} = \mathscr{M}(\mathbf{x})$ is transformed into the privacy loss random variable $L_{\mathbf{x},\mathbf{x}'}$. Moreover, following (Balle & Wang, 2018, Theorem 5), the (ϵ, δ)-differential privacy over $\mathcal{C}_d$ can be rewritten in the form of the privacy loss random variable.

Lemma 1. *The mechanism $\mathscr{M}$ is (ϵ, δ)-DP under μ-adjacency over the affine manifold $\mathcal{C}_d$ if and only if*

$$\mathbb{P}\big(L_{\mathbf{x},\mathbf{x}'} \geq \epsilon\big) - e^{\epsilon}\mathbb{P}\big(L_{\mathbf{x}',\mathbf{x}} \leq -\epsilon\big) \leq \delta \tag{23}$$

holds for every pair $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$.

A.1.2 PROOF OF NECESSITY

In this following, we suppose that the mechanism $\mathscr{M}$ is (ϵ, δ)-differentially private with privacy levels ϵ, δ, μ, and prove that both (7) and (8) hold, respectively.

Necessity of (7). We first apply the contradiction method to show that (7) holds. We suppose that (7) is not satisfied, i.e.,

$$\text{rank }(\Lambda) \neq \text{rank }\left(\begin{bmatrix}\Lambda & \mathbf{F}\mathbf{D}^{\perp}\end{bmatrix}\right). \tag{24}$$

Then we let $\Lambda_{\mathrm{ker}} \in \mathbb{R}^{(m-r)\times m}$ be such that $\Lambda_{\mathrm{ker}}\Lambda = 0$ and rank $\big([\Lambda^{\dagger}; \Lambda_{\mathrm{ker}}]\big) = m$. Thus, we have $\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{D}^{\perp} \neq 0$, i.e., there exist $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$ such that

$$\mathbf{D}(\mathbf{x} - \mathbf{x}') = 0\,, \qquad \Lambda_{\mathrm{ker}}\mathbf{F}(\mathbf{x} - \mathbf{x}') \neq 0\,.$$

With the pair $(\mathbf{x}, \mathbf{x}')$ being the case, we let $\mathcal{M}_* \subset \mathbb{R}^m$ such that $\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{x} \in \Lambda_{\mathrm{ker}}\mathcal{M}_*$, $\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{x}' \notin \Lambda_{\mathrm{ker}}\mathcal{M}_*$ and $\Lambda^{\dagger}\mathcal{M}_* = \mathbb{R}^r$. Then it is observed that

$$\begin{aligned}
& \mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in \mathcal{M}_*\big) \\
= \;& \mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in \mathcal{M}_*\big) \\
= \;& \mathbb{P}\big(\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{x} \in \Lambda_{\mathrm{ker}}\mathcal{M}_*, \Lambda^{\dagger}\mathbf{F}\mathbf{x} + \Lambda^{\dagger}\Lambda\boldsymbol{\eta} \in \Lambda^{\dagger}\mathcal{M}_*\big) \\
= \;& \mathbb{P}\big(\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{x} \in \Lambda_{\mathrm{ker}}\mathcal{M}_*\big)\mathbb{P}\big(\Lambda^{\dagger}\mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathbb{R}^r\big) \\
= \;& 1
\end{aligned}$$

and similarly,

$$\begin{aligned}
& \mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in \mathcal{M}_*\big) \\
= \;& \mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in \mathcal{M}_*\big) \\
= \;& \mathbb{P}\big(\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{x}' \in \Lambda_{\mathrm{ker}}\mathcal{M}_*\big)\mathbb{P}\big(\Lambda^{\dagger}\mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \mathbb{R}^r\big) \\
= \;& 0
\end{aligned}$$

This indicates that there is no (ϵ, δ) such that (1) is satisfied, contradicting with the fact that $\mathscr{M}$ is differentially privat with privacy levels (ϵ, δ, μ). Therefore, (24) does not hold, which completes the necessity proof of statement (i).

Necessity of (8). In view of the previous analysis, for any $\mathbf{x} \in \mathbb{R}^n$ and any $\mathcal{M} \subseteq \mathbb{R}^m$ it follows that

$$\mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in \mathcal{M}\big) = \mathbb{P}\big(\Lambda_{\mathrm{ker}}\mathbf{F}\mathbf{x} \in \Lambda_{\mathrm{ker}}\mathcal{M}\big)\mathbb{P}\big(\Lambda^{\dagger}\mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \Lambda^{\dagger}\mathcal{M}\big)\,. \tag{25}$$

With this in mind, we suppose that $\mathscr{M}$ over $\mathcal{C}_d$ is (ϵ, δ)-differentially private and (7) is satisfied, and proceed to show the inequality (8) is also satisfied.

We first show

$$\mathbb{P}\big(\Lambda_{\ker}\mathbf{F}\mathbf{x} \in \mathcal{M}_1\big) = \mathbb{P}\big(\Lambda_{\ker}\mathbf{F}\mathbf{x}' \in \mathcal{M}_1\big) \in \{0, 1\} \tag{26}$$

for all $\mathcal{M}_1 \subseteq \mathbb{R}^{m-r}$ and all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$. By the definition of $\mathcal{C}_d$, given any $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$ it is clear that $\mathbf{D}(\mathbf{x} - \mathbf{x}') = 0$. On the other hand, by (7) we have $\Lambda_{\ker}\mathbf{F}\mathbf{D}^{\perp} = 0$. Thus, it can be easily verified that $\Lambda_{\ker}\mathbf{F}(\mathbf{x} - \mathbf{x}') = 0$ for all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$. This immediately yields (26) by recalling that the event $\Lambda_{\ker}\mathbf{F}\mathbf{x} \in \mathcal{M}_1$ is deterministic.

With (26), we now proceed to show (8). By combining (25) and (26), given any $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$ we have

$$\begin{aligned} &\mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in \mathcal{M}\big) \leq e^{\epsilon}\mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in \mathcal{M}\big) + \delta\,, \quad \forall \mathcal{M} \subseteq \mathbb{R}^m \\ \iff \quad &\mathbb{P}\big(\Lambda^{\dagger}\mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big) \leq e^{\epsilon}\mathbb{P}\big(\Lambda^{\dagger}\mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \mathcal{M}_2\big) + \delta\,, \quad \forall \mathcal{M}_2 \subseteq \mathbb{R}^r \end{aligned} \tag{27}$$

It is noted that the lower inequality of (27) indicates that the mechanism $\mathscr{M}_2(\mathbf{x}) := \Lambda^{\dagger}\mathbf{F}\mathbf{x} + \boldsymbol{\eta}$ over $\mathcal{C}_d$ is (ϵ, δ)-differentially private.

Next, we apply this fact and Lemma 1 to the mechanism $\mathscr{M}_2(\mathbf{x})$ to show (8). It can be verified that the random variables $\mathbf{Y} := \Lambda^{\dagger}\mathbf{F}\mathbf{x} + \boldsymbol{\eta} \backsim \mathcal{N}(\Lambda^{\dagger}\mathbf{F}\mathbf{x}, \mathbf{I}_r)$, $\mathbf{Y}' := \Lambda^{\dagger}\mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \backsim \mathcal{N}(\Lambda^{\dagger}\mathbf{F}\mathbf{x}', \mathbf{I}_r)$ and the privacy loss function

$$\ell_{\mathbf{x},\mathbf{x}'}(\mathbf{y}) = (\mathbf{y} - \Lambda^{\dagger}\mathbf{F}\mathbf{x})^{\top}\Lambda^{\dagger}\mathbf{F}(\mathbf{x} - \mathbf{x}') + \frac{1}{2}\|\Lambda^{\dagger}\mathbf{F}(\mathbf{x} - \mathbf{x}')\|^2$$

Then, we can obtain both the privacy loss random variables $L_{\mathbf{x},\mathbf{x}'}$ and $L_{\mathbf{x}',\mathbf{x}}$ share the same distributions as

$$L_{\mathbf{x},\mathbf{x}'} \backsim \mathcal{N}(\eta/2, \eta)$$

with $\eta = \|\Lambda^{\dagger}\mathbf{F}(\mathbf{x} - \mathbf{x}')\|^2$. According to Lemma 1, the (ϵ, δ)-differential privacy of $\mathscr{M}_2$ is equivalent to saying

$$\mathbb{P}\big(L_{\mathbf{x},\mathbf{x}'} \geq \epsilon\big) - e^{\epsilon}\mathbb{P}\big(L_{\mathbf{x}',\mathbf{x}} \leq -\epsilon\big) \leq \delta$$

i.e.,

$$\Phi\Big(-\frac{\epsilon}{\sqrt{\eta}} + \frac{\sqrt{\eta}}{2}\Big) - e^{\epsilon}\Phi\Big(-\frac{\epsilon}{\sqrt{\eta}} - \frac{\sqrt{\eta}}{2}\Big) \leq \delta \tag{28}$$

for all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$. We note that the left side of the above inequality increases as η increases by Balle & Wang (2018). Thus, there must hold

$$\Phi\Big(-\frac{\epsilon}{\sqrt{\eta_{\max}}} + \frac{\sqrt{\eta_{\max}}}{2}\Big) - e^{\epsilon}\Phi\Big(-\frac{\epsilon}{\sqrt{\eta_{\max}}} - \frac{\sqrt{\eta_{\max}}}{2}\Big) \leq \delta \tag{29}$$

where we have defined

$$\eta_{\max} := \sup_{(\mathbf{x},\mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)} \|\Lambda^{\dagger}\mathbf{F}(\mathbf{x} - \mathbf{x}')\|^2$$

We now proceed to show that $\eta_{\max} = \mu^2\Delta_{\mathcal{N}}^2$. For any $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$, we recall Definition 4 and observe that for any $j \in \mathcal{J}$

$$\begin{aligned} &\mathbf{x} - \mathbf{x}' \\ &= \mathbf{E}_{\mathrm{d_j}}(\mathbf{x}_{\mathrm{d_j}} - \mathbf{x}'_{\mathrm{d_j}}) + \mathbf{E}_{-\mathrm{d_j}}(\mathbf{x}_{-\mathrm{d_j}} - \mathbf{x}'_{-\mathrm{d_j}}) \\ &= -\mathbf{E}_{\mathrm{d_j}}\mathbf{D}_{\mathrm{d}_j}^{-1}\mathbf{D}_{-\mathrm{d}_j}(\mathbf{x}_{-\mathrm{d_j}} - \mathbf{x}'_{-\mathrm{d_j}}) + \mathbf{E}_{-\mathrm{d_j}}(\mathbf{x}_{-\mathrm{d_j}} - \mathbf{x}'_{-\mathrm{d_j}}) \\ &\overset{(a)}{=} [\mathbf{I}_n - \mathbf{E}_{\mathrm{d_j}}\mathbf{D}_{\mathrm{d}_j}^{-1}\mathbf{D}]\mathbf{E}_{-\mathrm{d_j}}(\mathbf{x}_{-\mathrm{d_j}} - \mathbf{x}'_{-\mathrm{d_j}}) \\ &= \Psi_j\mathbf{E}_{-\mathrm{d_j}}(\mathbf{x}_{-\mathrm{d_j}} - \mathbf{x}'_{-\mathrm{d_j}}) \end{aligned}$$

where the third equation of (5) has been used to obtain the equation (a). Then, for any $i \in \mathrm{V}$ with $|x_i - x'_i| \leq \mu$, by using the first two equations of (5) there always exists j such that $i \in \mathrm{p}_j$ and

$$\mathbf{x} - \mathbf{x}' = \Psi_j\mathbf{E}_{-\mathrm{d_j}}(\mathbf{x}_{-\mathrm{d_j}} - \mathbf{x}'_{-\mathrm{d_j}}) = \Psi_j\mathbf{e}_i(x_i - x'_i) \tag{30}$$

yielding

$$\|\Lambda^{\dagger}\mathbf{F}(\mathbf{x} - \mathbf{x}')\| \leq \|\Lambda^{\dagger}\mathbf{F}\Psi_j\mathbf{e}_i\|\mu \leq \mu\Delta_i^{\mathcal{N}}$$

Thus, we have $\eta_{\max} = \mu^2\Delta_{\mathcal{N}}^2$, proving (8) by (29).

A.1.3 PROOF OF SUFFICIENCY

With (7) and (8), we now prove that the mechanism $\mathscr{M}$ preserves (ϵ, δ)-differential privacy over $\mathcal{C}_d$.

We first show that under (8) the mechanism $\mathscr{M}_2(\mathbf{x}) := \Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta}$ over $\mathcal{C}_d$ is (ϵ, δ)-differentially private. According to Lemma 1, this is equivalent to proving (28) for all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$. Then, by recalling the fact that $\eta_{\max} = \mu^2 \Delta^2_{\mathcal{N}}$, one can immediately conclude from the (8) that (28) is true for all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$, i.e.,

$$\mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big) \leq e^\epsilon \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \mathcal{M}_2\big) + \delta\,, \quad \forall \mathcal{M}_2 \subseteq \mathbb{R}^r\,.$$

Thus, by (25) and (26) we further have

$$\begin{array}{rl} & \mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in \mathcal{M}\big) \\ = & \mathbb{P}\big(\Lambda_{\ker}\mathbf{F}\mathbf{x} \in \Lambda_{\ker}\mathcal{M}\big)\mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \Lambda^\dagger \mathcal{M}\big) \\ \leq & \mathbb{P}\big(\Lambda_{\ker}\mathbf{F}\mathbf{x}' \in \Lambda_{\ker}\mathcal{M}\big)\Big(e^\epsilon \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \Lambda^\dagger \mathcal{M}\big) + \delta\Big) \\ \leq & e^\epsilon \mathbb{P}\big(\Lambda_{\ker}\mathbf{F}\mathbf{x}' \in \Lambda_{\ker}\mathcal{M}\big)\mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \Lambda^\dagger \mathcal{M}\big) + \delta \\ = & e^\epsilon \mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in \mathcal{M}\big) + \delta \end{array}$$

for all $\mathcal{M} \subseteq \mathbb{R}^m$. Therefore, the proof is completed.

A.2 PROOF OF THEOREM 2

A.2.1 PROOF OF NECESSITY

Necessity of (7). The proof is the same as the necessity proof of (7) in Appendix A.1.2, and is thus omitted for simplicity.

Necessity of (9). Similar to the arguments in the necessity proof of (8), we can obtain (27) and thus the $(\epsilon, 0)$-differential privacy of the Laplace mechanism $\mathscr{M}_2(\mathbf{x}) := \Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta}$ w.r.t. $\mathcal{C}_d$ for $\boldsymbol{\eta} \sim \mathcal{L}(0, 1)^r$, by using (7) and the fact that $\mathscr{M}(\mathbf{x})$ is $(\epsilon, 0)$-differentially private over $\mathcal{C}_d$. With this in mind, we then proceed to show (9) by contradiction.

We suppose that $\Delta_{\mathcal{L}} > \epsilon/\mu$. This, by the definition of $\Delta_{\mathcal{L}}$ and (30), indicates that there exists $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$ such that

$$\|\Lambda^\dagger \mathbf{F}(\mathbf{x} - \mathbf{x}')\|_1 = \mu \Delta_{\mathcal{L}} > \epsilon$$

With such $(\mathbf{x}, \mathbf{x}')$ being the case, we then can always construct a non-empty set $\mathcal{S} \subset \mathbb{R}^r$ such that

$$\begin{array}{rcl} \|\mathbf{u} - \Lambda^\dagger \mathbf{F}\mathbf{x}\|_1 & = & \|\mathbf{u} - \Lambda^\dagger \mathbf{F}\mathbf{x}' - \Lambda^\dagger \mathbf{F}(\mathbf{x} - \mathbf{x}')\|_1 \\ & = & \|\mathbf{u} - \Lambda^\dagger \mathbf{F}\mathbf{x}'\|_1 - \|\Lambda^\dagger \mathbf{F}(\mathbf{x} - \mathbf{x}')\|_1 \end{array}$$

for all $\mathbf{u} \in \mathcal{S}$. We then let $\mathcal{M}_2 = \{\mathbf{y} \in \mathbb{R}^r : \mathbf{y} = \mathbf{u} - \Lambda^\dagger \mathbf{F}\mathbf{x}, \mathbf{u} \in \mathcal{S}\}$, and observe that

$$\begin{array}{rl} & \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big) \\ = & \frac{1}{2^r} \int_{\mathcal{M}_2} e^{-\|\boldsymbol{\eta}\|_1} d\boldsymbol{\eta} \\ = & \frac{1}{2^r} \int_{\mathcal{S}} e^{-\|\mathbf{u} - \Lambda^\dagger \mathbf{F}\mathbf{x}\|_1} d\mathbf{u} \\ = & \frac{1}{2^r} \int_{\mathcal{S}} e^{-\|\mathbf{u} - \Lambda^\dagger \mathbf{F}\mathbf{x}'\|_1 + \|\Lambda^\dagger \mathbf{F}(\mathbf{x} - \mathbf{x}')\|_1} d\mathbf{u} \\ = & e^{\|\Lambda^\dagger \mathbf{F}(\mathbf{x} - \mathbf{x}')\|_1} \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \mathcal{M}_2\big) \\ > & e^\epsilon \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \mathcal{M}_2\big), \end{array}$$

which clearly contradicts with the fact that $\mathscr{M}_2$ is $(\epsilon, 0)$-differentially private, i.e., for all $\mathcal{M}_2 \subseteq \mathbb{R}^r$ and all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$, there holds

$$\mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big) \leq e^\epsilon \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \mathcal{M}_2\big)\,.$$

Therefore, we have $\Delta_{\mathcal{L}} \leq \epsilon/\mu$, proving (9).

A.2.2 PROOF OF SUFFICIENCY

With (7) and (9), we now prove that the mechanism $\mathscr{M}$ preserves $(\epsilon, 0)$-differential privacy over the affine manifold $\mathcal{C}_d$.

We first show that under (9) the Laplace mechanism $\mathscr{M}_2(\mathbf{x}) := \Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta}$ over $\mathcal{C}_d$ is $(\epsilon, 0)$-differentially private. We first observe from the definition of $\Delta_{\mathcal{L}}$ and (30) that

$$\|\Lambda^\dagger \mathbf{F}(\mathbf{x} - \mathbf{x}')\|_1 \leq \mu \Delta_{\mathcal{L}} \leq \epsilon$$

for all $(\mathbf{x}, \mathbf{x}') \in \mathrm{Adj}(\mu, \mathcal{C}_d)$. Then, we note that for all $\mathcal{M}_2 \subseteq \mathbb{R}^r$,

$$\begin{array}{rcl} \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big) & \leq & e^{\|\Lambda^\dagger \mathbf{F}(\mathbf{x}-\mathbf{x}')\|_1} \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big) \\ & \leq & e^{\epsilon} \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \mathcal{M}_2\big)\,, \end{array}$$

proving $(\epsilon, 0)$-differential privacy of $\mathscr{M}_2$ over the affine manifold $\mathcal{C}_d$.

Thus, by (25) and (26) we further have

$$\begin{array}{rcl} & & \mathbb{P}\big(\mathscr{M}(\mathbf{x}) \in \mathcal{M}\big) \\ & = & \mathbb{P}\big(\Lambda_{\mathrm{ker}} \mathbf{F}\mathbf{x} \in \Lambda_{\mathrm{ker}} \mathcal{M}\big) \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x} + \boldsymbol{\eta} \in \Lambda^\dagger \mathcal{M}\big) \\ & \leq & \mathbb{P}\big(\Lambda_{\mathrm{ker}} \mathbf{F}\mathbf{x}' \in \Lambda_{\mathrm{ker}} \mathcal{M}\big) \Big(e^{\epsilon} \mathbb{P}\big(\Lambda^\dagger \mathbf{F}\mathbf{x}' + \boldsymbol{\eta} \in \Lambda^\dagger \mathcal{M}\big)\Big) \\ & = & e^{\epsilon} \mathbb{P}\big(\mathscr{M}(\mathbf{x}') \in \mathcal{M}\big)\,. \end{array}$$

for all $\mathcal{M} \subseteq \mathbb{R}^m$. Therefore, the proof is completed.

A.3 PROOF OF THEOREM 3

To prove Theorem 3, we first need to further elaborate the expressions of $\bar{\Delta}_{\mathcal{N}}$ in (11) with the mechanism $\mathscr{M}$ in (16) and the manifold $\mathcal{C}_d$ in (17). It is observed that with $\mathbf{D}$ in (18), there exist T sets d_j, $j \in \mathcal{J} := \{1, \ldots, T\}$ such that $\mathbf{D}_{\mathrm{d}_j} \in \mathbb{R}^{(n-n_x)\times(n-n_x)}$ is nonsingular and $-\mathrm{d}_j := \mathrm{V}/\mathrm{d}_j = \{(j-1)n_x + 1, \ldots, (j-1)n_x + n_x\}$. It then follows that

$$\begin{array}{rcl} \Psi_j \mathbf{E}_{-\mathrm{d}_j} & = & \big(\mathbf{I}_n - \mathbf{E}_{\mathrm{d}_j} (\mathbf{D}_{\mathrm{d}_j})^{-1} \mathbf{D}\big) \mathbf{E}_{-\mathrm{d}_j} \\ & = & \mathbf{E}_{-\mathrm{d}_j} - \mathbf{E}_{\mathrm{d}_j} (\mathbf{D}_{\mathrm{d}_j})^{-1} \mathbf{D}_{-\mathrm{d}_j} \\ & = & \mathbf{D}^\perp \mathbf{A}^{1-j} \end{array}$$

where the last equality is obtained by using $(\mathbf{D}_{\mathrm{d}_j})^{-1} \mathbf{D}_{-\mathrm{d}_j} = \big[\mathbf{A}^{1-j}; \ldots; \mathbf{A}^{-1}; \mathbf{A}; \ldots; \mathbf{A}^{T-j}\big]$ and the definition of $\mathbf{D}^\perp$ in (19). This further implies $\bar{\Lambda}^\dagger \mathbf{F} \Psi_j \mathbf{E}_{-\mathrm{d}_j} \mathbf{E}_{-\mathrm{d}_j}^\top \mathbf{e}_i = \bar{\Lambda}^\dagger \mathbf{O}_T \mathbf{A}^{1-j} \mathbf{E}_{-\mathrm{d}_j}^\top \mathbf{e}_i$ for $i \in \mathrm{V}$ and $j \in \mathcal{J}$.

It is also noted that for any $i \in \mathrm{d}_j$, $\mathbf{E}_{-\mathrm{d}_j}^\top \mathbf{e}_i = \mathbf{v}_k$ with $k := i - (j-1)n_x \in [1, n_x]$, while for any $i \in -\mathrm{d}_j$, $\mathbf{E}_{-\mathrm{d}_j}^\top \mathbf{e}_i = 0$. Therefore, we have

$$\begin{array}{rcl} \bar{\Delta}_{\mathcal{N}} & := & \max_{(i,j) \in \mathrm{V} \times \mathcal{J}} \|\bar{\Lambda}^\dagger \mathbf{F} \Psi_j \mathbf{E}_{-\mathrm{d}_j} \mathbf{E}_{-\mathrm{d}_j}^\top \mathbf{e}_i\| \\ & = & \max_{(j,k) \in \mathcal{J} \times [1, n_x]} \|\bar{\Lambda}^\dagger \mathbf{O}_T \mathbf{A}^{1-j} \mathbf{v}_k\|\,, \end{array}$$

completing the proof.