R^{2}DP: A Universally Utility-driven Framework via Randomizing the Randomization Mechanism of Differential Privacy

***Abstract*—Differential privacy (DP) has emerged as a de facto rigorous privacy notion against adversaries with arbitrary background knowledge. To improve the utility of DP mechanisms, minimizing the distortion in randomization has attracted significant interests recently. However, existing studies only optimize specific utility metrics and lack a unified framework for a variety of applications. To address such deficiency, we propose a novel universal utility-driven framework that can integrate different metrics/applications and return the optimal utility while ensuring differential privacy. The proposed scheme *r*andomizes the *r*andomization mechanism of *d*ifferential *p*rivacy (R^{2}DP) by regarding the variance of the injected noise (or equivalently, the privacy budget ϵ) itself as a random variable, and optimizes the utility for any given metric. The universality of R^{2}DP is then proved by showing that R^{2}DP can generate the entire space of probability distribution functions (PDF) using an alternative specification of PDFs called *Moment Generating Functions*.**

Furthermore, we design and implement the R^{2}DP framework to enable both data owners and recipients to specify their preferences on privacy and utility (e.g., error bound), respectively. R^{2}DP then maximizes the probability of satisfying both privacy and utility. After formally benchmarking R^{2}DP under the Laplace mechanism, we show that the usefulness of our class of noise probability distributions asymptotically approaches to the optimal distribution (i.e., staircase distribution for large ϵ and Laplace distribution for small ϵ). Finally, we experimentally validate the performance of R^{2}DP using the Privacy Integrated Queries (PINQ) platform on both statistical queries and privacy preserving data analysis. The experimental results demonstrate a significantly increase in the utility (e.g., up to 230% of the utility of the baseline Laplace mechanism in PINQ).

I. Introduction

Significant amounts of individual information are being collected and analyzed today by a wide variety of applications across different industries [1]. While pursuing better utility by discovering knowledge from the data, individual's privacy may be compromised during an analysis. To that end, differential privacy has been widely recognized as the state-of-the-art [2], [3] privacy notion. By requiring the presence of any individual's data in the input to only marginally affect the distribution over the output, differential privacy provides strong protection against adversaries in possession of arbitrary background knowledge about the individuals. On the other hand, since the privacy constraints (e.g., the degree of randomization) imposed by differential privacy may render the released data less useful for analysis, the fundamental trade-off between privacy and utility (i.e., analysis accuracy) has attracted significant attention in various settings [3]–[8].

More often, studying the fundamental trade-off between privacy and utility in the context of differential privacy refers to deriving the optimal randomization mechanism, as the main focus of the existing solutions. For instance, Ghosh et al. [9], [10] studied the problem for a single counting query and under a Bayesian framework metric, and showed that a universally optimal mechanism (adding a specific class of *geometric noise*) can be used to preserve differential privacy for all negative expected loss utility metrics. Subsequently, Geng et al. [11] studied this critical problem for the class of noise adding ϵ-differentially private mechanism. In particular, they showed that under ℓ_1 and ℓ_2 norms the widely used standard Laplacian mechanism is optimal in cases where strong privacy guarantees are necessary, whereas the staircase mechanism (which can be viewed as a geometric mixture of uniform probability distributions) performs exponentially better than the Laplacian mechanism in case of weaker privacy guarantees (a comprehensive discussion will be given in Section VII).

However, to the best of our knowledge, there are very limited studies w.r.t. other utility metrics, e.g., usefulness (for machine learning applications [12]), entropy-based measures (for signal processing applications [13], [14]) and graph distance metrics (for social network applications [15]). Therefore, it would be highly beneficial to build a universal framework with the capability of leveraging any utility metric/application when seeking for the optimal distribution. Such a framework can result in a global view necessary to optimize the utility w.r.t. any metric such as *usefulness* [12]. To address the aforementioned issues, this paper proposes a novel privacy model by *r*andomizing the *r*andomization mechanism in *d*ifferential *p*rivacy (R^{2}DP). Precisely, R^{2}DP is a universal framework which derives the near-optimal utility w.r.t. any utility metric specified by an application. It is worth noting that R^{2}DP formulates both differential privacyguarantee and utility into one function, which can also facilitate any DP-related optimization task to identify the most appropriate utility metric for a specific application.

Specifically, our key observation is that, most existing works (on mechanisms that are oblivious to the database or queries beyond the global sensitivity) assume a rather inflexible randomization mechanism based on a specific class of noise probability distributions with a constant variance, e.g., a constant scale parameter in Laplace mechanism. Therefore, as illustrated in Figure 1, by having a two-dimensional search space (consisted of all possible queries and randomization mechanisms with constant-variance additive noises), those existing approaches are missing the opportunity of leveraging a much larger, three-dimensional search space in their optimization algorithms. In particular, we first define our model to randomize the variance of the injected noise (or equivalently the privacy budget ϵ) in a differentially private scheme, according to some parameterized probability distributions (we note that this type of two-folded distributions are generally called either the *mixture distributions* or *compound distributions*, depending on whether the set of component distributions is countable or not [16]). Next, R^{2}DP, under a specified privacy guarantee, computes the maximal utility under the desired utility metric/application by formulating and solving an optimization problem over the space of all probability distributions.

We formally benchmark R^{2}DP under Laplace mechanism and discuss other mechanisms in Sections V and Appendix 7 due to space limitations. Next, we show that R^{2}DP can the entire space of PDFs using an alternative specification of PDFs called *Moment Generating Functions*add citation. This is the key in leveraging a variety of utility metrics into one uni-variable optimization problem (universality). For example, our analysis demonstrates the following privacy and utility guarantees (refer to Section V for complete results). for l1 and l2? (reviewer may not understand)

$$e^{\epsilon} = \frac{\mathbb{M}_{\epsilon}(0)}{M'_{\epsilon}(-\Delta q)},\ \ell_1 \propto \int_0^{\infty} M_{\epsilon}(-x)dx,\ \ell_2 \propto \iint_0^{\infty} M_{\epsilon}(-u)dudx$$

Furthermore, we design and implement an R^{2}DP framework that enables both data owners and recipients to specify both a privacy bound ϵ and a query of interest and its utility requirement (e.g., tolerable error bound), respectively. With those inputs, the framework maximizes the probability of satisfying the utility requirement by formulating and solving an optimization problem over the space of all probability distributions (e.g., optimization for the *usefulness* metric). This addresses a practical limitation that, in most existing approaches, the randomization mostly depends on the specified privacy budget ϵ (e.g., the *scale parameter* b in Laplace mechanism is fully decided by the requested query and ϵ), and therefore the utility requirement specified by a data recipient cannot be flexibly accommodated.

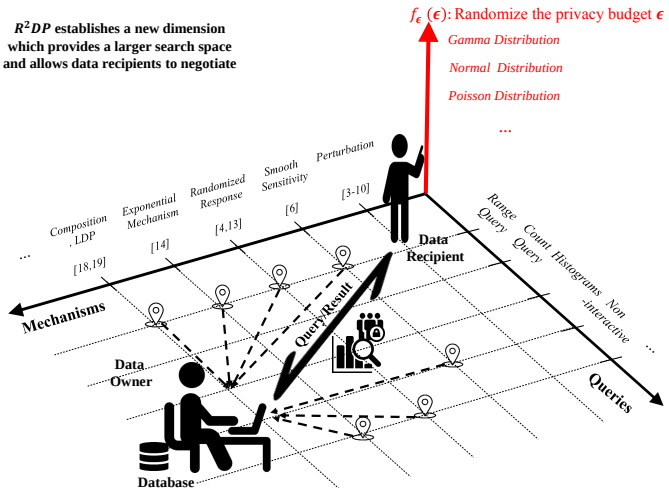


Fig. 1. R^{2}DP offers a larger search space for finding the utility-maximizing differential privacy mechanism.

A. Contributions

Therefore, R^{2}DP makes the following major contributions to improve the usability of differential privacy:

1) R^{2}DP significantly increases the search space of utility optimization for differentially private mechanisms (e.g., [6], [9], [17], [18]).
2) R^{2}DP can flexibly incorporate various utility metrics and the specified utility requirement (by the data recipient) into the optimization for differential privacy.
3) After tackling the challenges in the two-folded randomization in R^{2}DP, we formally prove that R^{2}DP under Laplace mechanism can unify the privacy and utility into an optimization problem in terms of the second-fold distribution's parameters. We also show that R^{2}DP generates near-optimal results (e.g., staircase-shape PDF for large values of ϵ and Laplace itself for small ϵ in Laplace mechanism).
4) We experimentally evaluate R^{2}DP on real data using both statistical queries (e.g., count, sum and average), and data analytics applications (e.g., machine learning and social network). The experimental results have demonstrated that R^{2}DP can significantly increase the utility of these data analyses (e.g., more than 230% of the utility of the baseline in PINQ).
5) The general concept of R^{2}DP can be adapted to improve a variety of other applications related to differential privacy, e.g., composition [18] and local differential privacy [17], as discussed in Section V.

The rest of the paper is organized as follows. Section II presents an overview of our R^{2}DP framework together with some background on differential privacy. Section III formally studies the differential privacyguarantee of the framework. Section IV analyzes the utility of the framework. Section V provides further discussions. Section VI

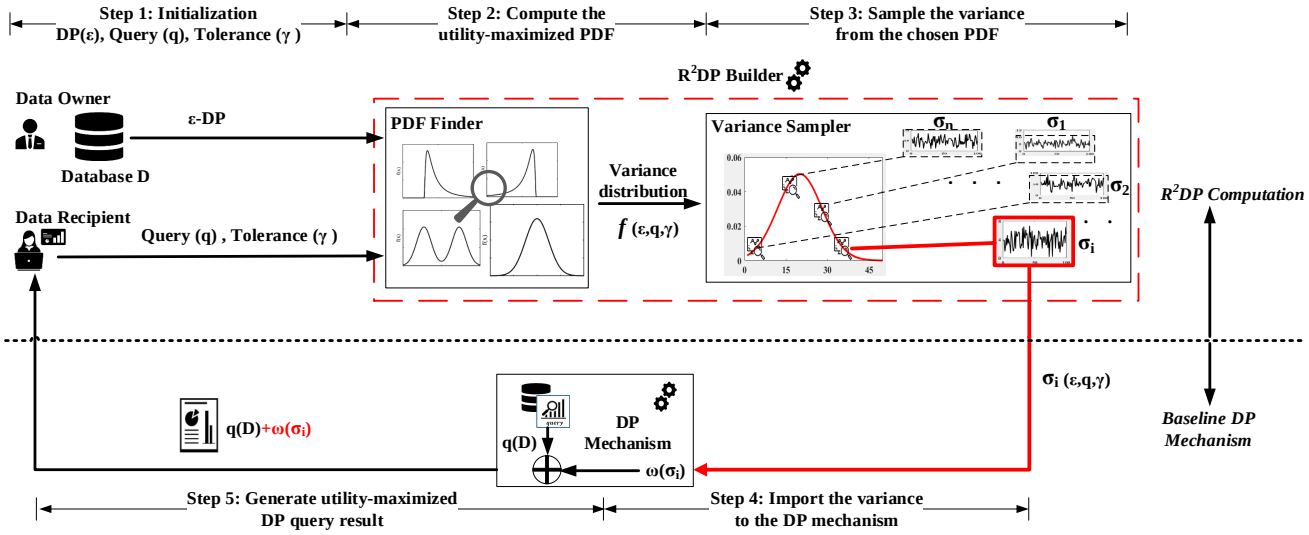


Fig. 2. An overview of the R^{2}DP approach

presents the experimental results. Section VII reviews the related work. Finally, Section VIII concludes the paper.

II. Preliminaries

In this section, we present an informal overview of the R^{2}DP framework and some required backgrounds on differential privacy.

A. Overview

As illustrated in Figure 2, our R^{2}DP framework consists of the following steps.

1) R^{2}DP Computation:

Step 1: The data owner inputs the differential privacyguarantee ϵ and the data recipient specifies his/her query of interest together with its error tolerance. We note that different queries may have their specific error tolerances. For instance, neural networks could have much higher error tolerance owing to their learning/training ability [19].

Step 2: Given the input triplets $(\epsilon, \Delta q, \gamma)$ where Δq is the sensitivity of the query (defined in Definition 2.1) the *PDF finder* module provably computes the optimal probability density function for the variance of the additive noise. For example, in Figure 2, the PDF finder returns a lower tail truncated Gaussian distribution for the specified inputs. This distribution maximizes the probability of generating an ϵ-DP result within γ distance from the original query result.

Step 3: The *variance sampler* randomly samples (w.r.t. to the PDF found in above step) a noise standard deviation σ_i.

2) Baseline DP Mechanism:

Step 4: Next, the computed standard deviation σ_i is used in generating a noise $\omega(\sigma_i)$ for the baseline DP mechanism.

Step 5: The computed noise $\omega(\sigma_i)$ is added to the query result $q(D)$ to provide utility-aware DP result.

We will more formally define the R^{2}DP framework and characterize its privacy and utility in Sections III and IV.

B. Backgrounds

In this section, we review some backgrounds on differential privacythat are required in sections III and IV, where we present the theoretical foundations of the PDF finder in our R^{2}DP framework. More importantly, we review the *usefulness* notion of utility which will be applied throughout our discussions later.

1) Differential Privacy: We follow the standard definitions of ϵ-differential privacy [6], [8]. Let D be a dataset of interest and d, d' be two adjacent subsets of D meaning that we can obtain d' from d simply by adding or subtracting the data of one individual. A randomization mechanism $\mathcal{M} : \mathsf{D} \times \Omega \rightarrow \mathsf{R}$ which is ϵ-differentially private, necessarily randomizes its output in such a way that for all $S \subset \mathsf{R}$ the following property is satisfied.

$$\mathbb{P}(\mathcal{M}(d) \in S) \leq e^{\epsilon}\mathbb{P}(\mathcal{M}(d') \in S), \quad \forall S \subset \mathsf{R}. \tag{1}$$

If the inequality fails, then a leakage (ϵ breach) takes place. This simply means the difference between the prior distribution and posterior one is tangible. We recall below a basic mechanism that can be used to answer queries in a ϵ-differentially private way. Note that we are only concerned in this paper with queries that return numerical

answers, i.e., here a query is a map $q : \mathsf{D} \to \mathsf{R}$, where the output space R equals the set of real numbers $\mathbb{R}$. The following quantity plays an important role in the design of differentially private mechanism [3].

Definition 2.1: [6], [8] The sensitivity of a query $q : \mathsf{D} \to \mathsf{R}$ is defined as $\Delta q = \max_{d,d':\mathrm{Adj}(d,d')} |q(d) - q(d')|$.

2) Laplace Mechanism: The Laplace mechanism [3] modifies an answer to a numerical query by adding zero-mean noise distributed according to a Laplace distribution. Recall that the Laplace distribution with mean zero and scale parameter b, denoted $Lap(b)$, has density $p(x;b) = \frac{1}{2b} exp(-\frac{|x|}{b})$ and variance $2b^2$.

Theorem 2.1: Let $q : \mathsf{D} \to \mathbb{R}$ be a query , $\epsilon > 0$. Then the mechanism $\mathcal{M}_q : \mathsf{D} \times \Omega \to \mathbb{R}$ defined by $\mathcal{M}_q(d) = q(d) + w$, with $w \sim Lap(b)$, where $b \geq \frac{\Delta q}{\epsilon}$ is ϵ-differentially private.

3) Utility Measure: Usefulness: Following Blum et al. [12], we employ the following notion of utility.

Definition 2.2: (Usefulness Definition). A database mechanism $\mathcal{M}_q$ is (γ, ζ)-useful if with probability $1 - \zeta$, for every database $d \subseteq \mathsf{D}$, $|\mathcal{M}_q(d) - q(d)| \leq \gamma$.

Theorem 2.2: The Laplace Mechanism 2.1 is $(\frac{\Delta q}{\epsilon} \ln \frac{1}{\zeta}, \zeta)$-useful [20]. Equivalently, one could say that the Laplace Mechanism 2.1 is $(\gamma, e^{\frac{-\gamma}{b(\epsilon)}})$-useful.

4) Randomizing the Randomization (R^2 Distributions): In probability and statistics, a random variable (RV) that is distributed according to some parameterized PDF, with (some of) the parameters of that PDF themselves being random variables, is known as a *mixture* distribution [16] (when the underlying RV is discrete) or a *compound* distribution when the RV is continuous. Compound (or mixture) distributions arise in many contexts in the literature and arise naturally where a statistical population contains two or more sub-populations (see Appendix A for a formal definition).

In general, we shall call any differentially private query answering mechanisms that leverage R^2 probability distribution functions in their randomization, the *R^2DP mechanisms*, as formally defined in the following.

Definition 2.3: (*R^2DP Mechanisms*) Let $\mathcal{M}_q(d,u) = q(d) \bigoplus \omega(u)$ be a mechanism which randomizes the answer of a query q using a random oracle $\omega(u)$, where $\bigoplus$ stands for the corresponding operator and u is the set of parameters (mean, variance, etc) of the PDF of ω. We call $\mathcal{M}_q(d,u) = q(d) \bigoplus \omega(u)$, an *$R^2$DP mechanism* if at least one of the parameters $u_i \in u$, $i \leq |u|$ is/are chosen randomly w.r.t. a specified probability distribution $f_{u_i} \in \mathcal{F}$.

To make our discussions more concrete, Sections III and IV will focus on the R^2DP Laplace mechanisms.

III. R^2DP Laplace Mechanism

In this section, we define the R^2DP Laplace mechanism and formally characterize its privacy.

A. Mechanism Definition

Basically, the R^2DP Laplace mechanism will modify an answer to a numerical query by adding zero-mean noise distributed according to a mixture/compound Laplace distribution with scale parameter b itself distributed according to some distribution f_b.

Example 3.1: Suppose that the scale parameter b in a Laplace mechanism is randomized as follows:

$$b = \begin{cases} b_1 & \text{w.p.} \quad p, \\ b_2 & \text{w.p.} \quad q = 1 - p. \end{cases}$$

Then, the perturbed query result $q(D) + Lap(b)$ is an example of an R^2DP Laplace mechanism using a Bernoulli distribution.

Definition 3.1: Let $q : \mathsf{D} \to \mathbb{R}$ be a query and suppose $f_b \in \mathcal{F}$ is a probability density function of the scale parameter b. Then, the mechanism $\mathcal{M}_q : \mathsf{D} \times \Omega \to \mathbb{R}$, defined by $\mathcal{M}_q(d,b) = q(d) + Lap(b)$ is an R^2DP Laplace mechanism that utilizes PDF f_b.

B. Differential Privacy of the Mechanism

We now show the R^2DP Laplace Mechanisms provide differential privacy guarantee. The ϵ-differential privacy $\mathcal{M}_q(d,b)$, for RV b, can be derived as follows.

$$e^{\epsilon} = \max_{\forall S \in \mathsf{R}} \left\{ \frac{\mathbb{P}(\mathcal{M}_q(d,b) \in S)}{\mathbb{P}(\mathcal{M}_q(d',b) \in S)} \right\}, \tag{2}$$

For the two probabilities $\mathbb{P}(\mathcal{M}_q(d',b) \in S)$ and $\mathbb{P}(\mathcal{M}_q(d,b) \in S)$ (which can be obtained using Equation 12 in Appendix A), since a Laplace distribution is of a ($\propto x.e^{x \cdot t}$) order, and since $x.e^{x \cdot t} = \dfrac{\mathrm{d}e^{x \cdot t}}{\mathrm{d}t}$, hence, the cumulative distribution function (CDF) resulted from randomizing x (the inverse of the scale parameter) can be expressed in terms of the expectation $\mathbb{E}(e^{x \cdot t})$.

Example 3.2: Following example 3.1, for a Bernoulli distributed scale parameter b, we have

$$\begin{aligned} &\mathbb{P}(\mathcal{M}_q(d,b) \in S) \\ &= \int_{\mathbb{R}} \frac{p}{2b_1} \cdot \mathbb{1}_S\{q(d) + w\} e^{\frac{-|w|}{b_1}} + \frac{q}{2b_2} \cdot \mathbb{1}_S\{q(d) + w\} e^{\frac{-|w|}{b_2}} dw \\ &= \int_{\mathbb{R}} \big(\frac{p}{2b_1} \cdot e^{\frac{-|w|}{b_1}} + \frac{q}{2b_2} \cdot e^{\frac{-|w|}{b_2}} \big) \mathbb{1}_S\{q(d) + w\} dw \end{aligned}$$

where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function. It can be verified that the term in the parenthesis is the derivative of $\mathbb{E}(e^{\frac{1}{b} \cdot -|w|})$ w.r.t $-|w|$, and hence the above probability can be expressed in terms of the expectation.

Therefore, we can re-write the probabilities in Equation 2 in terms of the *Moment Generating Function (MGF)* for the probability distribution $f_{\frac{1}{b}}$. Recall that MGF of a random variable is an alternative specification of its probability distribution, and hence provides the basis of an alternative route to analytical results compared with

working directly with probability density functions or cumulative distribution functions [21]. In particular,

Definition 3.2: (*Moment Generating Function* [21]) The moment-generating function of a random variable x is

$$M_X(t) := \mathbb{E}\left[e^{tX}\right], \quad t \in \mathbb{R} \tag{3}$$

wherever this expectation exists. In other words, the moment-generating function is the expectation of the random variable e^{tX}.

Next, we give a general formula for the probabilities in Equation 2 as follows (see Appendix A for Proofs).

Theorem 3.1: For all $S \subset \mathbb{R}$ measurable and dataset d in D,

$$\mathbb{P}(\{q(d) + Lap(b)\} \in S) =$$
$$\tfrac{1}{2} \cdot \Big[- M_{\frac{1}{b}}(-|x - q(d)|)|_{S_{\geq q(d)}} + M_{\frac{1}{b}}(-|x - q(d)|)|_{S_{<q(d)}} \Big]$$

According to Equation 13, R^{2}DP Laplace mechanism can in fact generate the entire space of PDFs using an alternative specification (CDF is the moment and PDF is its derivative). Respectively, the DP bound is

$$e^{\epsilon} = \max_{\forall S \in \mathsf{R}} \left\{ \frac{-M_{\frac{1}{b}}(-|x-q(d)|)|_{S_{\geq q(d)}} + M_{\frac{1}{b}}(-|x-q(d)|)|_{S_{<q(d)}}}{-M_{\frac{1}{b}}(-|x-q(d')|)|_{S_{\geq q(d')}} + M_{\frac{1}{b}}(-|x-q(d')|)|_{S_{<q(d')}}} \right\}$$

Hence, the value of e^{ϵ} only depends on the distribution of reciprocal of the scale parameter b, i.e., $f_{\frac{1}{b}}$. Moreover, a MGF is positive and log-convex [21] where the latter property is desirable in defining various natural logarithm upper-bounds, e.g., DP bound. In the following theorem, we demonstrate the fact that our MGF-based formula for the probability $\mathbb{P}(\{q(d) + Lap(b)\} \in S)$ in Equation 4 can be easily applied to calculate the differential privacyguarantee.

Theorem 3.2: The R^{2}DP mechanism $\mathcal{M}_q(d, b)$, is

$$\ln \left[\max_{\forall x \in \mathbb{R}} \left\{ \frac{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d)|}}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d')|}} \right\} \right] \text{-differentially private} \tag{4}$$

Proof. According to Equation 13,

$$\mathbb{P}(\mathcal{M}_q(d, b) \in S) = \tfrac{1}{2} \int_S \frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d)|} dx$$
$$= \tfrac{1}{2} \int_S \frac{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d)|}}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d')|}} \cdot \frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d')|} dx$$

Denote by

$$e^{\epsilon} = \sup \left\{ \frac{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d)|}}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-|x-q(d')|}}, \forall x \in S \right\},$$
$$\Rightarrow \mathbb{P}(\mathcal{M}_q(d, b) \in S) \leq e^{\epsilon} \cdot \mathbb{P}(\mathcal{M}_q(d', b) \in S)$$

and the choice of $S = \mathbb{R}$ concludes the proof. □

The result in Theorem 3.2 can be further simplified (see Appendix A for the proof).

Corollary 3.3: The R^{2}DP mechanism $\mathcal{M}_q(d, b)$ is $\ln \left[\frac{\mathbb{E}(\frac{1}{b})}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-\Delta q}} \right]$-differentially private.

However, a challenge here is that not all random variables have moment-generating functions [22]. Therefore, we will study popular distributions (or combinations of these distributions) with known moment generating functions in order to identify those that can improve the privacy utility trade-off. Another criteria we consider in choosing the appropriate probability distribution is to pick the ones that are general enough to cover a large family of other probability distributions. For instance, since Exponential distribution, Erlang distribution, and Chi-squared distribution are special cases of Gamma distribution, we will only consider Gamma distribution.

IV. Utility of R^{2}DP Laplace Mechanism

In this section, we formally analyze the utility provided by R^{2}DP mechanisms. We will show that R^{2}DP model can unify two parallel concepts, i.e., privacy and utility, into one optimization problem under the differential privacy-constraint and defined over the space of the probability distributions with positive support. As a result, we will show that R^{2}DP model can provide a strictly dominating payoff in comparison to the baseline Laplace mechanisms (with a fixed parameter). In particular, we derive the accurate differential privacyand usefulness guarantee for *five* well-known probability distributions with non-negative supports (as $b \geq 0$). These distributions consist of *two* discrete and *three* continuous distributions; whose moment generating functions are computable.

A. Characterizing the Utility

We now characterize the utility of the generic R^{2}DP Laplace mechanism, using the notion of usefulness (see Section II-B) and some other well-known metrics.

1) Usefulness Metric: First, we have shown in Section III that

$$\forall S \in \mathsf{R} \Rightarrow \mathbb{P}(\{q(d) + Lap(b)\} \in S)$$
$$= \tfrac{1}{2} \cdot \Big[- M_{\frac{1}{b}}(-|x - q(d)|)|_{S_{\geq q(d)}} + M_{\frac{1}{b}}(-|x - q(d)|)|_{S_{<q(d)}} \Big]$$

Therefore, denote by $U(\epsilon, \Delta q, \gamma)$, the usefulness of an R^{2}DP Laplace mechanism for all $\epsilon > 0$, sensitivity Δq and error bound γ. The optimal usefulness, is then given as the answer (if there exists one) of the following optimization problem.

$$\max_{f_{\frac{1}{b}} \in F} \{U(\epsilon, \Delta q, \gamma)\} = \max_{f_{\frac{1}{b}} \in F} \Bigg\{ \frac{1}{2} \cdot \Big[- M_{\frac{1}{b}}(-|x - q(d)|)|_{q(d)}^{q(d)+\gamma} + M_{\frac{1}{b}}(-|x - q(d)|)|_{q(d)-\gamma}^{q(d)} \Big] \Bigg\},$$

$$\text{subject to} \quad \epsilon = \ln \left[\frac{\mathbb{E}(\frac{1}{b})}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-\Delta q}} \right]$$

Note that ϵ and Δq do not directly impact the usefulness but they do so indirectly through the differential privacy-constraint. Optimizing the objective function over all possible $f_{1/b \geq 0} \in \mathcal{F}$, is out of the scope of this paper, and in general, not all random variables have moment-generating functions [22]. Alternatively, for a predefined probability distribution, e.g., Gamma distribution, we maximize the usefulness over its parameters, e.g., mean and variance. Later, in Section IV-C, using the composability property of MGFs (see Theorem 4.11), the search space will be further extended to all linear combinations of predefined distributions (a space of infinite RVs) which results in a significant improvement in utility as it will be illustrated in Figure 3.

Corollary 4.1: Denote by u, the set of parameters for a probability distribution $f_{\frac{1}{b}}$, and by $M_{f(u)}$ its moment generating function. Then, the optimal usefulness of an R^{2}DP mechanism utilizing $f_{\frac{1}{b}}$, at each triplet $(\epsilon, \Delta q, \gamma)$ is

$$U_f(\epsilon, \Delta q, \gamma) = \max_{u \in \mathbb{R}^{|u|}} \Bigg\{ \frac{1}{2} \cdot \Big[- M_{f(u)}(-|x - q(d)|)|_{q(d)}^{q(d)+\gamma} + M_{f(u)}(-|x - q(d)|)|_{q(d)-\gamma}^{q(d)} \Big] \Bigg\},$$

$$\text{subject to} \quad \epsilon = \ln \left[\frac{\mathbb{E}(\frac{1}{b})}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-\Delta q}} \right]$$

However, moment generating functions are positive and log-convex, with $M(0) = 1$ and hence, $U_f(\epsilon, \Delta q, \gamma) = 1 - \min_{u \in \mathbb{R}^{|u|}} M_{f(u)}(-\gamma)$. Therefore, for usefulness metric, the best distribution for ϵ is the one with minimum MGF evaluated at γ. In particular, for a set of privacy and utility parameter, one can find the optimal point using the *Lagrange multiplier* method [23]. i.e.,

$$\mathcal{L}(u, \lambda) = M_{f(u)}(-\gamma) + \lambda \cdot (\ln \left[\frac{\mathbb{E}(\frac{1}{b})}{\frac{\mathrm{d}M_{\frac{1}{b}}(t)}{\mathrm{d}t}|_{t=-\Delta q}} \right] - \epsilon) \quad (5)$$

Next, under the differential privacyguarantee of several probability distributions, we will apply Equation 5 to find the optimal trade-off.

However, not all probability distributions can boost the utility of the baseline Laplace mechanism. Accordingly, in the following theorem, we derive a necessary condition on the differential privacyguarantee of an R^{2}DP Laplace mechanism to boost the utility of the baseline Laplace mechanism (refer to Appendix A for the proof). Using this necessary condition, we can easily filter out those probability distributions that cannot deliver any utility improvement.

Theorem 4.2: The usefulness of an R^{2}DP Laplace mechanisms with $\epsilon \geq \ln \left[\mathbb{E}_{\frac{1}{b}} \left(e^{\epsilon(b)} \right) \right]$, is always upper bounded by the usefulness of the ϵ-differentially private baseline Laplace mechanism. Equivalently, for an R^{2}DP Laplace mechanism to boost the utility, the following relation is necessarily true.

$$e^{\epsilon} = \frac{\mathbb{E}(\frac{1}{b})}{M'_{\frac{1}{b}}(-\Delta q)} < M_{\frac{1}{b}}(\Delta q) \quad (6)$$

We note that $\epsilon = \ln \left[\mathbb{E}_{\frac{1}{b}} \left(e^{\epsilon(b)} \right) \right]$ provides a tight upper bound as it gives the overall e^{ϵ} of an R^{2}DP Laplace mechanism as the average of differential privacyleakages.

2) Other Metrics: Similarly, we can derive the utility of R^{2}DP Laplace mechanism under some well-known utility metrics. Due to space limitation, we present only the final results in Table I.

TABLE I
UTILITY OF R^{2}DP LAPLACE MECHANISM UNDER DIFFERENT METRICS

ℓ_1	ℓ_2	Entropy	Usefulness
$\int_0^{\infty} M_{\frac{1}{b}}(-x)dx$	$\sqrt{2\int_0^{\infty}\int M_{\frac{1}{b}}(-u)dudx}$	$\int_0^{\infty} \frac{M_{\frac{1}{b}}(-x) \cdot M"_{\frac{1}{b}}(-x)}{M'_{\frac{1}{b}}(-x)}dx$	$1 - M_{\frac{1}{b}}(-\gamma)$

Where by entropy measure here, we follow Wang et al. [24]. In particular, given a mechanism $Y = \mathcal{M}(X)$, we use the output entropy $H(\mathcal{M}) = \sup \left(\int_{\mathbb{R}} -\mathbb{P}(x, y) \ln(\mathbb{P}(x, y)) dy, x \in \mathbb{R} \right)$. The precise set of results in Table I can be easily used to test the appropriateness of each measure in different applications. Note that Laplace mechanism has already been studied under the first three measure. Therefore, the rest of the paper will be focused on usefulness.

B. Finding Utility-Maximizing Probability Distributions

We now examine a set of well-known probability distributions. Promisingly, our analytic evaluations for *three* of these distributions, i.e., Gamma, Uniform and truncated Gaussian distribution offer a significantly improved ϵ compared with the bound given in Theorem 4.2.

1) Discrete Probability Distributions: First, we consider two different mixture Laplace distributions that can be applied for constructing R^{2}DP Laplace mechanisms with discrete probability distribution f_b.

(1) **Degenerate distribution.** A degenerate distribution is a probability distribution in a space (discrete or continuous) with support only on a space of lower dimension [25]. If the degenerate distribution is uni-variate

(involving only a single random variable) it is a deterministic distribution and takes only a single value. Therefore, the degenerate distribution is identical to the baseline Laplace mechanism as it also assigns the mechanism one single scale parameter b_0. Specifically, the probability mass function of the uni-variate degenerate distribution is:

$$f_{\delta,k_0}(x) = \begin{cases} 1 & x = k_0 \\ 0 & x \neq k_0 \end{cases}$$

The MGF for the degenerate distribution δ_{k_0} is given by $M_k(t) = e^{t \cdot k_0}$ [22]. Using Equation 4, Theorem 4.3 gives the same differential privacyguarantee as the baseline Laplace mechanism.

Theorem 4.3: The R^{2}DP Laplace mechanism $M_q(d, \epsilon)$, $\epsilon \sim f_{\delta,\frac{1}{b_0}}(\epsilon)$, is $\frac{\Delta q}{b_0}$-differentially private.

Obviously, this distribution does not improve the bound in Theorem 3.2 and is only presented to show the soundness of our findings.

(2) **Bernoulli distribution.** The probability mass function of this distribution, over possible outcomes k, is

$$f_B(k; p) = \begin{cases} p & \text{if } k = 1, \\ q = 1 - p & \text{if } k = 0. \end{cases}$$

Note that the binary outcomes $k = 0$ and $k = 1$ can be mapped to any two outcomes X_0 and X_1, respectively. Therefore, we consider the following Bernoulli outcomes

$$f_{B,X_0,X_1}(X; p) = \begin{cases} p & \text{if } X = X_1, \\ q = 1 - p & \text{if } X = X_0. \end{cases}$$

The MGF for Bernoulli distribution $f_{B,X_0,X_1}(X; p)$ is $M_X(t) = p \cdot e^{t \cdot X_0} + (1 - p) \cdot e^{t \cdot X_1}$ [22]. We now derive the precise differential privacyguarantee of a R^{2}DP Laplace mechanism with its scale parameter randomized according to a Bernoulli distribution.

Theorem 4.4: The R^{2}DP Laplace mechanism $M_q(d, \epsilon)$, $\epsilon \sim f_{B,\frac{1}{b_0},\frac{1}{b_1}}(\epsilon; p)$, is $\ln[p \cdot e^{\frac{\Delta q}{b_0}} + (1-p) \cdot e^{\frac{\Delta q}{b_1}}]$-differentially private.

However, this bound is exactly the mean value of $e^{\epsilon(b)}$ and therefore, this distribution does not improve the bound given in Theorem 3.2, either.

2) Continuous Probability Distributions: We now investigate three compound Laplace distributions.

(1) **Gamma distribution.** The gamma distribution is a two-parameter family of continuous probability distributions with a shape parameter $k > 0$ and a scale parameter θ. Besides the generality, the gamma distribution is the maximum entropy probability distribution (both w.r.t. a uniform base measure and w.r.t. a $1/x$ base measure) for a random variable X for which $\mathbb{E}(X) = k\theta = \alpha/\beta$ is fixed and greater than zero, and $\mathbb{E}[\ln(X)] = \psi(k) + \ln(\theta) = \psi(\alpha) - \ln(\beta)$ is fixed (ψ is the digamma function). Therefore, it may provide a relatively higher privacy-utility trade-off in comparison to the other candidates [26], [27]. A random variable X that is gamma-distributed with shape α and rate β is denoted by $X \sim \Gamma(k, \theta)$ and the corresponding probability density function is

$$f_\Gamma(X; k, \theta) = \frac{x^{k-1} e^{-\frac{x}{\theta}}}{\Gamma(k) \cdot \theta^k} \quad \text{for } X > 0 \text{ and } k, \theta > 0,$$

where $\Gamma(\alpha)$ is the gamma function. We now investigate the differential privacyguarantee provided by assuming that the the reciprocal of the scale parameter b in Laplace mechanism is distributed according to the gamma distribution (see Appendix A for the proof).

Theorem 4.5: The R^{2}DP Laplace mechanism $M_q(d, \epsilon)$, $\epsilon \sim f_\Gamma(\epsilon; k, \theta)$, is $\big((k+1) \cdot \ln(1 + \Delta q \cdot \theta)\big)$-differentially private.

We now test the necessary condition given in Equation 6.

Lemma 4.6: R^{2}DP using Gamma distribution may satisfy the necessary condition in Equation 6.

Proof. We need to show that there exists k, θ such that $(k+1) \cdot \ln(1 + \Delta q \cdot \theta) < -k \cdot \ln(1 - \Delta q \cdot \theta)$, $\theta < \frac{1}{\Delta q}$. Suppose $\theta = \frac{1}{2\Delta q}$, then we need to show that there exist k such that

$$k \cdot \ln(2) > (k+1) \cdot \ln(1.5)$$

which always holds for all $k > 1.4094$. □

Therefore, Gamma distribution may improve over the baseline, and this can be computed by optimizing the privacy-utility trade-off using Equation 5. In particular, our optimization shows that, this distribution is more effective in large values of ϵ (weaker privacy guarantees).

(2) **Uniform distribution.** In probability theory and statistics, the continuous uniform distribution or rectangular distribution is a family of symmetric probability distributions such that for each member of the family, all intervals of the same length on the distribution's support are equally probable. The support is defined by the two parameters, a and b, which are its minimum and maximum values. The distribution is often abbreviated $U(a, b)$. It is the maximum entropy probability distribution for a random variable X under no constraint other than that it is contained in the distribution's support [26], [27]. The MGF for $U(a, b)$ is

$$M_X(t) = \begin{cases} \frac{e^{tb} - e^{ta}}{t(b-a)} & \text{for } t \neq 0, \\ 1 & \text{for for } t = 0. \end{cases}$$

. We now , using Corollary 3.3, derive the precise differential privacyguarantee of a R^{2}DP Laplace mechanism for uniform distribution $U(a, b)$.

Theorem 4.7: The R^{2}DP Laplace mechanism $M_q(d, \epsilon)$, $\epsilon \sim f_{U(a,b)}(\epsilon)$, is $\ln\big[\frac{\alpha^2 - \beta^2}{2((1+\beta)e^{-\beta} - (1+\alpha)e^{-\alpha})}\big]$-differentially private, where $\alpha = a \cdot \Delta q$ and $\beta = b \cdot \Delta q$.

We now test the necessary condition given in Equation 6. One can easily verify that the inequality 6 holds for infinite number of settings, e.g., $a = 0.5$, $b = 9$ and $\Delta q = 1.2$.

Lemma 4.8: R^{2}DP using Uniform distribution may satisfy the necessary condition in Equation 6.

Therefore, uniform distribution may improve over the baseline, and this can be computed by optimizing the privacy-utility trade-off using Equation 5. In particular, our optimization shows that, this distribution can also be effective in both small and large values of ϵ (stronger privacy guarantees).

(3) **Truncated Gaussian distribution.** The last distribution we consider is the Truncated Gaussian distribution. This distribution is derived from that of a normally distributed random variable by bounding the random variable from either below or above (or both). Therefore, we can benefit from the numerous useful properties of Gaussian distribution, by truncating the negative region of the Gaussian distribution. Suppose $X \sim \mathcal{N}(\mu, \sigma^2)$ has a Gaussian distribution and lies within the interval $X \in (a, b),\ -\infty \leq a < b \leq \infty$. Then, X conditional on $a < X < b$ has a truncated Gaussian distribution with the following probability density function.

$$f_{\mathcal{N}^T}(X; \mu, \sigma, a, b) = \frac{\phi(\frac{X-\mu}{\sigma})}{\sigma \cdot \left(\Phi(\frac{b-\mu}{\sigma}) - \Phi(\frac{a-\mu}{\sigma})\right)} \quad \text{for } a \leq x \leq b$$

and by $f_{\mathcal{N}^T} = 0$ otherwise. Here, $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ and $\Phi(x) = 1 - Q(x)$ are PDF and CDF of the standard Gaussian distribution, respectively. Next, using Corollary 3.3, the differential privacyguarantee provided by the mechanism assuming that the reciprocal of b is distributed according to the truncated Gaussian distribution is given.

Theorem 4.9: The R^{2}DP Laplace mechanism $\mathcal{M}_q(d, \epsilon)$, and $\epsilon \sim f_{\mathcal{N}^T}(\epsilon; \mu, \sigma, a, b)$, is ϵ_{N^T}-differentially private, where

$$\epsilon_{N^T} = \ln \left[\frac{\mu + \dfrac{\sigma \cdot (\phi(\alpha) - \phi(\beta))}{(\Phi(\beta) - \Phi(\alpha))}}{\dfrac{\mathrm{d}M_{N^T}(t)}{\mathrm{d}t}|_t = -\Delta q} \right] \tag{7}$$

where $\phi(\cdot)$ is the probability density function of the standard normal distribution, $\phi(\cdot)$ is its cumulative distribution function and $\alpha = \frac{a-\mu}{\sigma}$ and $\beta = \frac{b-\mu}{\sigma}$.

Lemma 4.10 (see Appendix A for the proof): R^{2}DP using truncation of Gaussian distribution may satisfy the necessary condition in Equation 6.

Therefore, truncated Gaussian distribution may improve over the baseline, and this can be computed by optimizing the privacy-utility trade-off using Equation 5. In particular, our optimization shows that, this distribution can also be effective in smaller values of ϵ (stronger privacy guarantees).

C. Maximizing Utility with Combined PDFs

Fortunately, MGFs possess an appealing composability property between independent probability distributions, which allows us to further enlarge the search space in finding the maximal utility through combining random variables.

Theorem 4.11 (MGF of Linear Combination of RVs): If $x_1, x_2, \cdots, x_n$ are n independent random variables with respective moment-generating functions $M_{x_i}(t) = \mathbb{E}(e^{tx_i})$ for $i = 1, 2, \cdots, n$, then the moment-generating function of the linear combination $Y = \sum_{i=1}^{n} a_i x_i$ is $\prod_{i=1}^{n} M_{x_i}(a_i t)$.

Theorem 4.11 can be applied in designing a utility-maximizing R^{2}DP Laplace mechanism to achieve a sufficiently large search space (infinite number of different random variables).

Corollary 4.12 (Optimal Utility for Linear Combination of RVs): If $x_1, x_2, \cdots, x_n$ are n independent random variables with respective moment-generating functions $M_{x_i}(t) = \mathbb{E}(e^{tx_i})$ for $i = 1, 2, \cdots, n$, then for the linear combination $Y = \sum_{i=1}^{n} a_i x_i$, the optimal utility under ϵ-differential privacyconstraint is given as

$$U_Y(\epsilon, \Delta q, \gamma) = 1 - \min_{\mathcal{A}, \mathcal{U}} \left\{ \prod_{i=1}^{n} M_{x_i}(-a_i \gamma) \right\} \tag{8}$$

subject to

$$\epsilon = ln \left[\frac{\sum_{j=1}^{n} a_j \cdot E_{x_j}(\frac{1}{b})}{\sum_{j=1}^{n} a_j \cdot M'_{x_j}(a_j \cdot -\Delta q) \cdot \prod_{\substack{i=1 \\ i \neq j}}^{n} M_{x_i}(-a_i \cdot \Delta q)} \right]$$

where $\mathcal{A} = \{a_1, a_2, \cdots, a_n\}$ is the set of the coefficients and $\mathcal{U} = \{u_1, u_2, \cdots, u_n\}$ is the set of parameters of the probability distributions of RVs $x_i,\ \forall i \leq n$.

Similar to the single RV, we can compute the optimal solution for this optimization problem using the Lagrange multiplier function in Equation 5. We will focus on all RVs that are produced using linear combinations of the Gamma, Uniform and truncated Gaussian distributions (compose both weak and strong privacy preserving pdfs). Therefore, the corresponding Lagrange multiplier function is

$$\begin{aligned} &\mathcal{L}(a_1, a_2, a_3, k, \theta, a_u, b_u, \mu, \sigma, a_{\mathcal{N}^T}, b_{\mathcal{N}^T}, \Lambda) \\ &\quad = M_{\Gamma(k,\theta)}(-a_1\gamma) \cdot M_{U(a_u,b_u)}(-a_2\gamma) \\ &\qquad \cdot M_{\mathcal{N}^T(\mu,\sigma,a_{\mathcal{N}^T},b_{\mathcal{N}^T})}(-a_3\gamma) + \Lambda \cdot (\ln \left[\frac{\mathsf{N}}{\mathsf{D}} \right] - \epsilon) \end{aligned} \tag{9}$$

where the numerator and the denominator N, D are

$$\mathsf{N} = (a_1 \cdot k \cdot \theta) + (a_2 \cdot \tfrac{a+b}{2}) + (a_3 \cdot (\mu + (\frac{\sigma \cdot \phi(\alpha) - \phi(\beta))}{(\Phi(\beta) - \Phi(\alpha))}))$$

$$\begin{aligned} \mathsf{D} = {} & a_1 \cdot M'_{\Gamma(k,\theta)}(-a_1 \cdot \Delta q) \cdot M_{U(a_u,b_u)}(-a_2 \cdot \Delta q) \\ & \cdot M_{\mathcal{N}^T(\mu,\sigma,a_{\mathcal{N}^T},b_{\mathcal{N}^T})}(-a_3 \cdot \Delta q) \\ & + a_2 \cdot M_{\Gamma(k,\theta)}(-a_1 \cdot \Delta q) \cdot M'_{U(a_u,b_u)}(-a_2 \cdot \Delta q) \\ & \cdot M_{\mathcal{N}^T(\mu,\sigma,a_{\mathcal{N}^T},b_{\mathcal{N}^T})}(-a_3 \cdot \Delta q) \\ & + a_3 \cdot M_{\Gamma(k,\theta)}(-a_1 \cdot \Delta q) \cdot M_{U(a_u,b_u)}(-a_2 \cdot \Delta q) \\ & \cdot M'_{\mathcal{N}^T(\mu,\sigma,a_{\mathcal{N}^T},b_{\mathcal{N}^T})}(-a_3 \cdot \Delta q) \end{aligned}$$

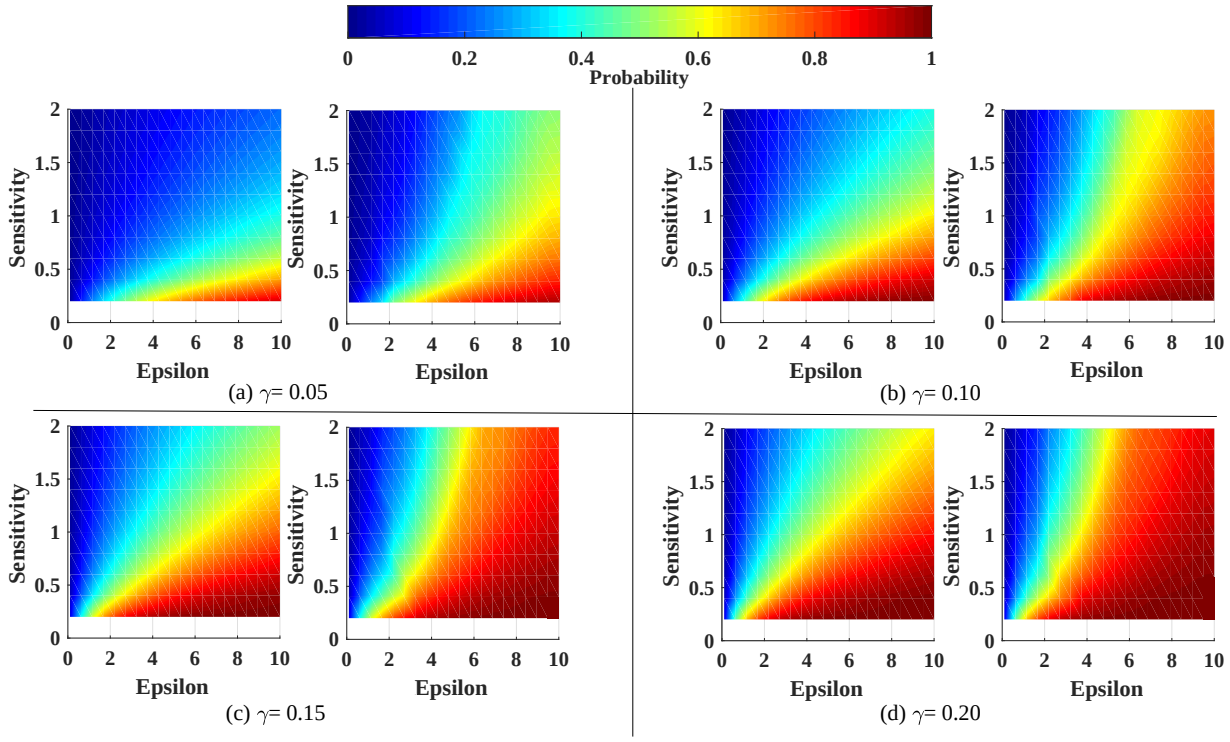


Fig. 3. Numerical results to illustrate the superiority of R^{2}DP Laplace mechanism. The left figure in each of the configurations is the usefulness for the baseline (Laplace mechanism) and the right figure is the R^{2}DP Laplace mechanism's performance.

Algorithm 1 (in Appendix A) summarizes our utility-aware R^{2}DP Laplace mechanism using linear combination of these three PDFs, and Figure 3 depicts the corresponding performance for four different values of γ-error. We observe that R^{2}DP can in fact deliver its promises as illustrated in Figure 3. In particular, R^{2}DP using Gamma, Uniform and truncated Gaussian Distributions produces a dominating payoff under most of the settings where the rate of this payoff depends on the privacy and utility parameters. Specifically, as we show in Figure 10 in Appendix, the results generated by R^{2}DP is quite similar to the optimal class of noise, i.e., Laplace for small ϵ and Staircase-shape for large ϵ [28]. We further verify this fact through our Experiments in Section VI.

V. Discussion

In this section, we first present numerical results on the utility improvements and parameter tuning of R^{2}DP, and highlight open issues and possible improvements. We then discuss the applications of R^{2}DP in other differential privacy-related research areas.

A. Private Data Analysis with R^2DP

We now analyze the improvements provided by R^{2}DP through numerical results under different settings. We first discuss the performance of R^{2}DP under stronger privacy guarantees ($\epsilon < 2$). Next, we study the improvement for counting queries ($\Delta q = 1$) while varying γ error. Moreover, we discuss how to tune the R^{2}DP γ parameter to keep the query result at the minimized distance from the original result in mean-square sense. Finally, we discuss open issues and possible improvements.

R^{2}DP Under Stronger Privacy Guarantees. As discussed earlier, Gamma distribution is more effective in the case of weaker privacy (bigger ϵ) whereas truncated Gaussian distribution is more effective in the case of stronger privacy (smaller ϵ). Therefore, using truncated Gaussian distribution with appropriate parameters, R^{2}DP can generate improved results compared with the Laplace mechanism in fairly strong values of ϵ. In fact, since Laplace is asymptotically optimal for when $\epsilon \to 0$, the optimal staircase shape distribution cannot be easily tuned due to catastrophic cancellation (parameters and corresponding figure discussed in Geng et al. [11]). In contrast, R^{2}DP can successfully be implemented in those values of ϵ. Figure 4 confirms this through numerical results in four different values of γ. In particular, Figure 4 shows that as γ decreases truncated Gaussian distribution generates a further improved results (up to 1.4 times). However, our analysis shows that trend doesn not hold for further decreasing γ. In fact, the performance ratio depends on all the three parameters.

Performance of R^{2}DP for Counting Queries. As shown in Figure 5, R^{2}DP significantly improves the usefulness of the counting queries. The numerical results, for

some pair (ϵ, γ), predicts up to 125% improvement. We will further validate the numerical results shown in Figure 5 through experiments in Section VI.

Tuning Parameter γ for Data Recipients. The R^{2}DP model formulates the privacy-utility trade-off under the usefulness definition of utility, which means a bigger γ error will result in a higher probability. On the other hand, results with a bigger γ may also have a larger distance from the original query result. Therefore, tuning γ based on specific needs of the application and the analysis might lead to better results in practice. We will further study this trade-off about the choice of γ through experiments in next section by considering outputs outside the specified error-bound as false positives. Since each γ is tuned for one specific application and requirement, we will not consider the comparison between different values of γ. Alternatively, to maximize the utility, the data recipients may want to minimize the mean square error (MSE) $\mathbb{E}(|\omega(\epsilon, \gamma)|^2)$ whose utility function is shown in Table I. Designing R^{2}DP models for minimizing the MSE (variance of the compound Laplace distribution) is an interesting future direction.

Open Issues and Improvements. Despite the effectiveness of our R^{2}DP, one possible problem is to maximize the utility over the global space of PDFs (those with infinite MGF). Address this issue might lead to a more complete view of the privacy-utility trade-off under R^{2}DP. One future direction is that, since the characteristic function of a continuous random variable X is the Fourier transform of its probability density function $f_X(x)$, we may define a linear programming optimization problem in Fourier domain for the probability distribution function which will help to construct the optimal PDF. Second, we have only applied R^{2}DP to a few popular probability distributions, and a future direction is to study R^{2}DP under other differential privacymechanisms, e.g., Gaussian mechanism and exponential mechanism. Third, we consider tuning the parameter of R^{2}DP as an off-line process, i.e., the optimal distribution of the scale parameter according to all input triplets is computed one time and the users will tune the Laplace noise parameter using a pre-computed b_r. A future direction is to bring this process online by dynamically tuning them based on streams of data and incoming queries.

B. Other Applications of R^{2}DP

In this paper, we have focused on the application of R^{2}DP to privacy preserving query answering. However, R^{2}DP represents a very general concept which could potentially be applied in a broader range of contexts. In general, applying R^{2}DP to design more application-aware distortions may further improve the utility of many existing solutions, e.g., smooth sensitivity (sample and aggregate) [6] (Section VII surveys other existing utility-maximizing solutions). In particular, the following outlines some of the possible applications.

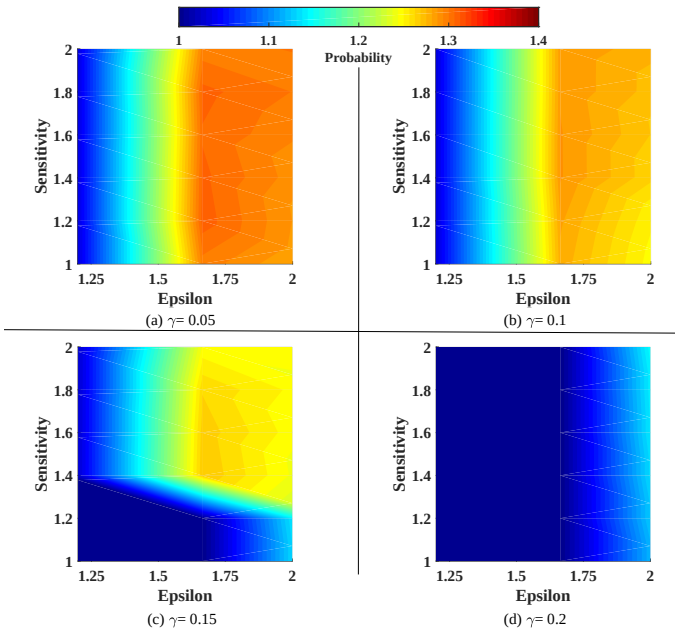


Fig. 4. Numerical results to demonstrate the effectiveness of R^{2}DP in stronger values of ϵ.

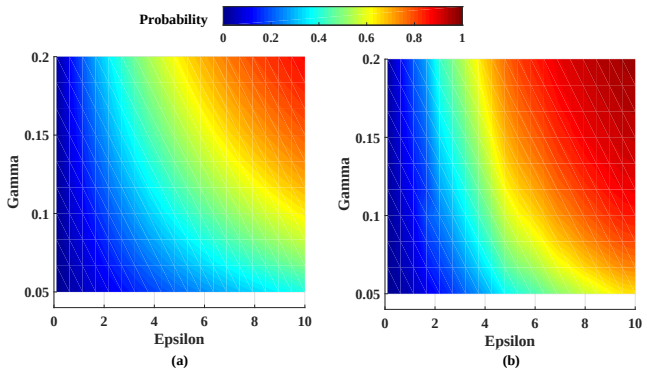


Fig. 5. Numerical results to compare the performance of (a) the baseline Laplace mechanisms, for count queries and (b) R^{2}DP, when varying γ error and ϵ.

R^{2}DP and Composition. R^{2}DP may be applied for reducing the differential privacyleakage due to sequential querying of a dataset. In those scenarios, the objective will be to maximize the number of compositions under a specified ϵ-differential privacyconstraint.

R^{2}DP and Local Differential Privacy. In this context, R^{2}DP can be regarded as a new randomized response model. In particular, the randomized response scheme presented in [29] can be produced using R^{2}DP for the Bernoulli distribution when $b_0 \to 0$ and $b_1 \to \infty$. Therefore, designing more efficient local differential privacy schemes using R^{2}DP is an interesting future direction.

R^{2}DP for Continual Observation Applications. Providing differential privacy guarantees on data streams

represents another important future direction for R^{2}DP. As an example, the multi-input multi-output (MIMO) systems process streams of signals originated from many sensors capturing privacy-sensitive events about individuals, and statistics of interest need to be continuously published in real time [30], [31], e.g., privacy-preserving traffic monitoring over multi-lane roads [32]. In this context, R^{2}DP can leverage the constraint related to the number of inputs and the number of outputs (e.g., the sensitivity of the output of MIMO filter G with m inputs and p outputs is proportional to the $\mathcal{H}_2$ norm of G which itself is an increasing function of m and p [33]) into its model to build more efficient differentially private mechanisms for the MIMO scenarios.

VI. Experimental Evaluations

In this section, we experimentally evaluate the performance of our R^{2}DP mechanism, and compare it with the theoretically proven optimal results (i.e., Laplace and the optimal staircase PDFs) via real world applications and datasets.

A. Experimental Setting

We perform all the experiments and comparisons on the Privacy Integrated Queries (PINQ) platform [34]. Besides two basic statistical queries, two applications in the current suite (*machine learning* and *social network analysis*) are employed to evaluate the accuracy of R^{2}DP, Laplace and staircase PDFs. Specifically, we conduct the following experiments.

1) **Statistical Queries**: In the first set of our experiments, we examine the benefits of the proposed technique using two basic statistical queries, i.e., count and average. The dataset used here comes from a sensor network experiment carried out in the Mitsubishi Electric Research Laboratories (MERL) and described in [35]. MERL has been collecting motion sensor data from a network of over 200 sensors for a year that contains over 30 million raw motion records.
2) **Machine Learning**: Naive Bayes classification (injected with R^{2}DP and Laplace noise, respectively) is performed on two datasets: Adult dataset (in the UCI ML Repository) [36] and KDD Cup 99 dataset (available at http://kdd.ics.uci.edu/databases/kddcup99/kddcup99). First, the adult dataset includes the demographic information of 48,842 different adults in the US (14 features in total). It can be utilized to train a Naive bayes claissifier to predict whether any adult's annual salary is greater than 50k or not. Second, the second dataset in a KDD competition was utilized to build a network intrusion detector (based on 24 training attack types) by classifying "bad" connections and "good" connections.
3) **Social Network**: Social network degree distribution and recommendation application (injected with R^{2}DP and Laplace noise, respectively) are performed on the DBLP dataset [37] in which the co-authorship network of computer science academic articles is connected by representing different authors as nodes (317,080 in total) and co-author relationship as edges (1,049,866 in total).

All the experiments are performed on an HP PC with Inter Core i7-7700 CPU 3.60GHz and 32G RAM.

B. Statistical Queries Results

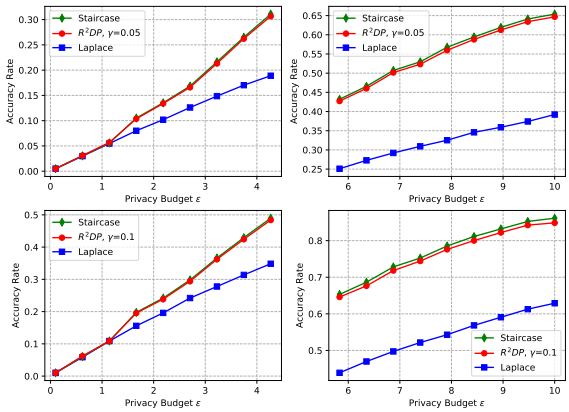
We first verify the effectiveness our R^{2}DP framework using count and average queries over the MERL dataset. As shown in Figure VI, our evaluation by varying the privacy budget and for two error bounds $\gamma = 5\%, 10\%$ shows that compared with the baseline Laplace mechanism, our framework can provide a dominating trade-off for privacy budgets larger than a lower-bound depending on the sensitivity of the query. In particular, for the count query (sensitivity=1) and the average query (sensitivity=0.2), the lower-bounds are $\epsilon = 1.2$ and $\epsilon = 2.2$, respectively. Moreover, we observe the multiplicative ratio gain shown in Geng et al. [28] for relatively larger values ϵ (in usefulness sense). Finally, we observe that R^{2}DP generates a very close to optimal results.

C. Machine Learning Results

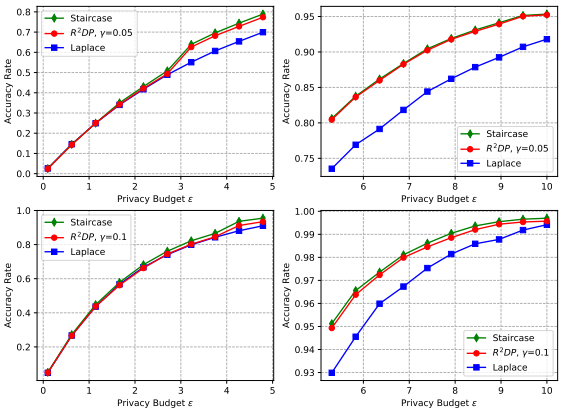
After executing the naive bayes classifier on the Adult dataset (45K training records and 5K testing records), the precision and recall of the classification are derived as 0.814 and 0.825, respectively. We consider such results as the baseline for our comparison. Then, we evaluate the precision and recall of R^{2}DP and Laplace based naive classification [38] by varying the privacy budget for each PINQ query $\epsilon \in [0.1, 10]$ (sensitivity=1) where two different error bounds $\gamma = 0.05, 0.1$ are specified for our R^{2}DP. Then, we have the following observations:

- As shown in Figure 7(a), 7(b) our R^{2}DP based classification is more accurate than the Laplace mechanism for the same privacy budget of the PINQ queries ϵ. As the privacy budget ϵ increases, following our statistical query experiments, R^{2}DP offers a far better precision/recall (close to the baseline without privacy consideration) because the noise adding mechanism approaches to the optimal PDF.
- Among the precision/recall results derived with two different γ in R^{2}DP, for each ϵ, one out of the two specified error bounds (e.g., $\gamma = 5\%$) may reach the highest accuracy (not necessarily the result with the smallest γ) (refer to discussion on tuning parameter γ in section V-A)
- As shown in Figure 8(a), 8(b), we can draw similar observations from the KDD Cup 99 dataset.

The above experimental results have validated the effectiveness of integrating our R^{2}DP based queries to improve the output utility for classification while ensuring ϵ-differential privacy.

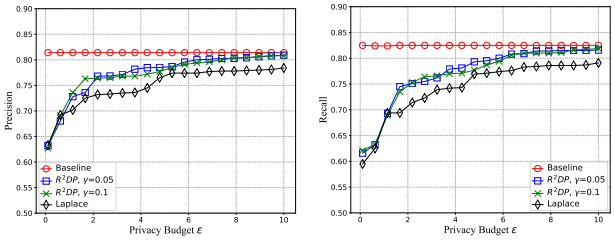


(a) Accuracy comparison for count query.



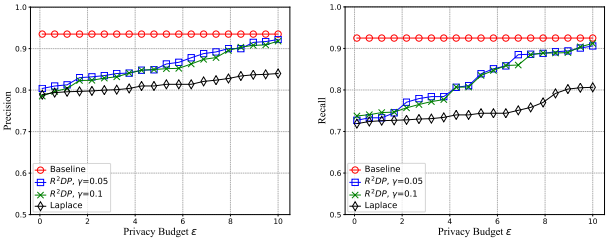
(b) Accuracy comparison for average query.

Fig. 6. Accuracy Evaluation of R^{2}DP for Statistical Query in comparison to the baseline Laplace and the Optimal Staircase PDFs.



(a) Precision vs. ϵ (b) Recall vs. ϵ

Fig. 7. Accuracy Evaluation for Classification (UCI Adult Dataset)



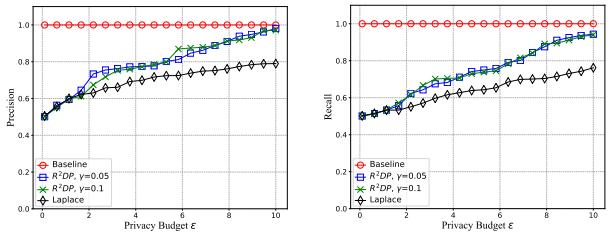
(a) Precision vs. ϵ (b) Recall vs. ϵ

Fig. 8. Accuracy Evaluation for Classification (KDDCUP99 Dataset)

D. Social Network Analysis Results

Besides classification, we also conduct experiments to compare the performance of R^{2}DP and Laplace based on PINQ queries in social network analysis. We implement the social recommendation (e.g., follow, connection) [39] by using Jaccard Coefficient to measure the similarity between any two unlinked nodes in the social graph:

$$Sim(x,y) = \frac{|\Psi(x) \cap \Psi(y)|}{|\Psi(x) \cup \Psi(y)|} \tag{10}$$



(a) Precision vs. ϵ (b) Recall vs. ϵ

Fig. 9. Accuracy Evaluation for Social Recommendation

where $\Psi(x)$ and $\Psi(y)$ represent the set of linked nodes for node x and y, respectively, and $|\cdot|$ returns the cardinality of the set. Thus, a threshold τ for the similarity between two unlinked nodes can be specified to derive a set of recommendations.

Again, for each group of experiments (e.g., given a τ), we derive the recommendation set as the baseline, and calculate the precision and recall for PINQ queries based recommendation sets for R^{2}DP and Laplace (via benchmarking the non-noisy results). More specifically, by fixing the similarity threshold $Sim(x,y) > \tau = 0.5$, we evaluate the performance for different privacy budgets of the PINQ queries ϵ in $[0.1, 10]$. Then we have the following observations for the social network analysis:

- As the privacy budget ϵ for PINQ queries declines, we can draw similar observations from the recommendation function as the machine learning results (Figure 9(a) and 9(b)) (Laplace is optimal).
- As shown in Figure 9(a) and 9(b), as the similarity threshold increases from 50% to 90%, both precision and recall converge to the baseline (no privacy consideration). This reflects that higher threshold similarity

would generate recommendation for unlinked nodes with more common edges and then the results would be more tolerable to noise.
In the meanwhile, the accuracy of R^{2}DP (all two γ) still outperforms the Laplace noise for any similarity threshold where smaller γ results in higher recommendation accuracy.

VII. RELATED WORK

Differential Privacy [3] is a model for preserving privacy while releasing the results of various useful functions, such as contingency tables, histograms, means, etc. [2]. Many existing works focus on improving the utility based on different differential privacy mechanisms. In the following, we briefly review several categories of utility-maximizing techniques.

Noise Perturbation. Based on the general utility maximization framework from Ghosh et al. [9], Gupte and Sundararajan [10] further study the optimal noise probability distributions for single count queries. Later, Geng el al. [11], [40] demonstrate the optimal noise distribution has a staircase-shaped probability density function for Laplace mechanism. Similar to Laplace mechanism, Balle and Wang [41] develop an optimal Gaussian mechanism in high privacy regime to minimize the noise and increase the utility for queries. Geng et al. [42] further show the optimal noise distribution is a uniform distribution over Gaussian mechanism. Moreover, Hardt et al. [43] study the privacy/utility trade-off for answering a set of linear queries over a histogram, where the error of the analysis is defined as the worst expectation of the ℓ^2-norm (identical to variance) of the noise among all possible outputs. Subsequently, Brenner et al. [44] show that, for general query functions, no *universally optimal* differential privacy mechanisms exist.

Sampling and Aggregation. Sampling and aggregation frameworks normally split the database into chunks, and aggregate the result using a differential private algorithm after executing the query on each chunk [6]. To expand the applicability of output perturbation, Nissim et al. [6] propose a framework to formally analyze the effect of instance-based noise. Observing the highly compressible nature of many real-life data, researchers propose lossy compression techniques to add noise calibrated to the compressed data. Acs et al. [45] propose an optimization of Foruier perturbation algorithm that cluster and exploit the redundancy between bins. Instead of directly adding noise to histogram counts, this approach first lossily compresses the data, then adds noise calibrated to the data. Li et al. [46] propose an algorithm that first partitions a data domain into uniform regions then adapts the strategy to fit the specific set of range queries to achieve a lower error rate. Zhang et al. [47] improve the clustering mechanism by sorting histogram bins based on the noisy counts.

Data Composition. Barak et al. [48] propose transforming the data into the Fouier domain, which could avoid the violation of consistency for low-order marginals in database tables. As efficiency is the main bottleneck for this approach when the number of attributes is large, Hay et al. [49] ensure that the error rate does not grow with the size of a database. The proposed hierarchical histogram method also achieves a lower error for a fixed domain. Different than one-dimensional datasets solution proposed by Hay et al. [49], Xiao et al. [50] propose *Privelet* that improves accuracy on datasets with arbitrary dimensions, which could reduce error to 25% compared to 70% as baseline error rate. Cormode et al. [51] apply *quadtrees* and *kd-trees* as new techniques for parameter setting to improve the accuracy on spatial data. Ding et al. [52] introduce a general noise-control framework on data cubes. Li et al. [53] unify the two range queries over histograms into one framework. Other techniques, such as principal component analysis (PCA), linear discriminant analysis (LDA) [54], and random projection [55], [56] are also used to lower the data dimension for reducing the perturbation errors.

Adaptive Queries. In this technique, the improvement of utilities takes advantage of a known set of queries, for example, Dwork et al. [57] propose *Boosting for Queries* algorithm to obtain a better accuracy of learning algorithms. Hardt et al. [58], [59] present multiplicative weights mechanism to improve the efficiency of interactive queries. Instead of polynomial running time [8], this work achieves a nearly linear running time with a relaxed utility requirement. Yuan et al. [60], [61] propose low-rank mechanism (LRM) to further improve the adaptive queries. Other techniques such as, correlated noise [62] and sparse vector technique (SVT) [63] are also used in adaptive queries.

Applications. Many researchers also work on improving the utility for different types of data, such as, the Fourier Perturbation Algorithm (FPA$_k$) [7] in time-series data (e.g., location traces, web history, and personal health data), *kd-trees* on spatial data [51], and matrix-valued query [64].

a) Summary: Our R^{2}DP framework provides a complementary approach to those existing works by providing the opportunity of searching for the maximal utility along an extra dimension. This framework also enables data recipients to specify their utility requirements and the computed parameter could be incorporated into existing solutions to further improve utility.

VIII. Conclusion

This paper has proposed a universally utility-driven framework called R^{2}DP which can leverage variety of utility metrics/applications. For this purpose, R^{2}DP randomizes the randomization mechanism of differentially private solutions that leads to remodeling the entire space of PDFs. Specifically, we have shown that a differentially

private mechanism could be defined based on a random variable which was itself distributed according to some parameterized distribution. We have also shown that such a mechanism could explicitly take into account both the privacy requirement and the utility requirement specified by the data owner and data recipient, respectively. We have presented our model based on the well-known Laplace mechanism and formally proved the improvement of the computed statistics over the baseline Laplace mechanism. Finally, our experimental results for statistical queries, machine learning and social network applications have demonstrated that our proposed framework could indeed significantly improve the usefulness of differential privacysolutions.

References

[1] Gregory A Aarons, Amy E Green, Lawrence A Palinkas, Shannon Self-Brown, Daniel J Whitaker, John R Lutzker, Jane F Silovsky, Debra B Hecht, and Mark J Chaffin. Dynamic adaptation process to implement an evidence-based child maltreatment intervention. *Implementation Science*, 7(1):32, 2012.

[2] Cynthia Dwork. Differential privacy: A survey of results. In *Theory and Applications of Models of Computation*, volume 4978, pages 1–19, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[3] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography*, pages 265–284, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[4] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the ACM SIGSAC conference on computer and communications security*, pages 1054–1067, Scottsdale, AZ, USA, 2014. ACM.

[5] Sangmin Lee, Edmund L Wong, Deepak Goel, Mike Dahlin, and Vitaly Shmatikov. πbox: A platform for privacy-preserving apps. In *Proceedings of the 10th {USENIX} Symposium on Networked Systems Design and Implementation*, {NSDI} '13, pages 501–514, Lombard, IL, USA, 2013.

[6] Kobbi Nissim, Sofya Raskhodnikova, and Adam D. Smith. Smooth sensitivity and sampling in private data analysis. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, STOC '07, pages 75–84, San Diego, California, USA, 2007.

[7] Vibhor Rastogi and Suman Nath. Differentially private aggregation of distributed time-series with transformation and encryption. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, SIGMOD '10, pages 735–746, Indianapolis, Indiana, USA, 2010.

[8] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil P. Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, STOC '09, pages 381–390, Bethesda, MD, USA, 2009.

[9] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, STOC '09, pages 351–360, New York, NY, USA, 2009. ACM.

[10] Mangesh Gupte and Mukund Sundararajan. Universally optimal privacy mechanisms for minimax agents. In *Proceedings of the 29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '10, pages 135–146, New York, NY, USA, 2010. ACM.

[11] Quan Geng and Pramod Viswanath. The optimal mechanism in differential privacy. In *2014 IEEE International Symposium on Information Theory*, pages 2371–2375, Honolulu, HI, USA, June 2014.

[12] Avrim Blum, Katrina Ligett, and Aaron Roth. A learning theory approach to non-interactive database privacy. In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, STOC '08, pages 609–618, New York, NY, USA, 2008. ACM.

[13] Joel E Cohen, YVES Derriennic, and GH Zbaganu. Majorization, monotonicity of relative entropy, and stochastic matrices. *Contemporary Mathematics*, 149:251–251, 1993.

[14] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud. Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems. In *53rd IEEE Conference on Decision and Control*, pages 2130–2135, Dec 2014.

[15] Shiva Prasad Kasiviswanathan, Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Analyzing graphs with node differential privacy. In Amit Sahai, editor, *Theory of Cryptography*, pages 457–476, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[16] Charalambos A. Charalambides. *Combinatorial Methods in Discrete Distributions (Wiley Series in Probability and Statistics)*, volume 600. Wiley-Interscience, New York, NY, USA, 2005.

[17] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. Extremal mechanisms for local differential privacy. In *Advances in Neural Information Processing Systems 27*, pages 2879–2887. Curran Associates, Inc., 2014.

[18] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. *IEEE Transactions on Information Theory*, 63(6):4037–4049, June 2017.

[19] Zidong Du, Krishna V. Palem, Lingamneni Avinash, Olivier Temam, Yunji Chen, and Chengyong Wu. Leveraging the error resilience of machine-learning applications for designing highly energy efficient accelerators. In *2014 19th Asia and South Pacific Design Automation Conference*, ASP-DAC '14, pages 201–206, 2014.

[20] T.-H. Hubert Chan, Elaine Shi, and Dawn Song. Private and continual release of statistics. *ACM Transactions Information System Security*, 14(3):26:1–26:24, 2011.

[21] Marek Fisz. *Probability theory and mathematical statistics*, volume 3. 2018.

[22] Michael George Bulmer. *Principles of statistics*. Courier Corporation, 1979.

[23] Dimitri P Bertsekas. *Constrained optimization and Lagrange multiplier methods*. Academic press, 2014.

[24] Y. Wang, Z. Huang, S. Mitra, and G. E. Dullerud. Entropy-minimizing mechanism for differential privacy of discrete-time linear feedback systems. In *53rd IEEE Conference on Decision and Control*, pages 2130–2135, Dec 2014.

[25] Hans Bremermann. Distributions, complex variables, and fourier transforms. 1965.

[26] R. Bruce Kellogg. CRC standard mathematical tables and formulae. *SIAM Review*, 38(4):691–692, 1996.

[27] MV Jambunathan. Some properties of beta and gamma distributions. *The annals of mathematical statistics*, 25(2):401–405, 1954.

[28] Quan Geng and Pramod Viswanath. Optimal noise adding mechanisms for approximate differential privacy. *IEEE Transactions on Information Theory*, 62(2):952–969, Feb 2016.

[29] Yue Wang, Xintao Wu, and Donghui Hu. Using randomized response for differential privacy preserving data collection. In *Proceedings of EDBT/ICDT Workshops Joint Conference*, EDBT/ICDT '16, Bordeaux, France, 2016.

[30] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N. Rothblum. Differential privacy under continual observation. In *Proceedings of the 42rd ACM Symposium on Theory of Computing*, STOC '10, pages 715–724, New York, NY, USA, 2010. ACM.

[31] Jerome Le Ny and George J Pappas. Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2):341–354, 2014.

[32] Joshua W. S. Brown, Olga Ohrimenko, and Roberto Tamassia. Haze: Privacy-preserving real-time traffic statistics. In *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '13, pages 540–543, New York, NY, USA, 2013. ACM.

[33] Jerome Le Ny and Meisam Mohammady. Differentially private MIMO filtering for event streams. *IEEE Transactions on Automatic Control*, 63(1):145–157, Jan 2018.

[34] Frank McSherry. Privacy integrated queries: an extensible platform for privacy-preserving data analysis. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, SIGMOD '09, pages 19–30, Rhode Island, USA, 2009. ACM.

[35] Christopher R. Wren, Yuri A. Ivanov, Darren Leigh, and Jonathan Westhues. The merl motion detector dataset. In *Proceedings of the 2007 Workshop on Massive Datasets*, MD '07, pages 10–14, New York, NY, USA, 2007. ACM.

[36] Ron Kohavi. Scaling up the accuracy of naive-bayes classifiers: A decision-tree hybrid. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD' 96, pages 202–207. AAAI Press, 1996.

[37] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection, Jun 2014. http://snap.stanford.edu/data.

[38] Jaideep Vaidya, Basit Shafiq, Anirban Basu, and Yuan Hong. Differentially private naive bayes classification. In *2013 IEEE/WIC/ACM International Conferences on Web Intelligence*, WI '13, pages 571–576, Atlanta, GA, USA, 2013.

[39] Balazs Horanyi. Follow recommendations in social networks, 2017. https://getstream.io/blog/follow-recommendations-in-social-networks/.

[40] Quan Geng, Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The staircase mechanism in differential privacy. *IEEE Journal of Selected Topics Signal Processing*, 9(7):1176–1184, 2015.

[41] Borja Balle and Yu-Xiang Wang. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *Proceedings of the 35th International Conference on Machine Learning*, ICML '18, pages 403–412, Stockholm, Sweden, 2018.

[42] Quan Geng, Wei Ding, Ruiqi Guo, and Sanjiv Kumar. Optimal noise-adding mechanism in additive differential privacy. *CoRR*, abs/1809.10224, 2018.

[43] Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the 42rd ACM Symposium on Theory of Computing*, STOC '10, pages 705–714, New York, NY, USA, 2010. ACM.

[44] Hai Brenner and Kobbi Nissim. Impossibility of differentially private universally optimal mechanisms. In *IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 71–80, Las Vegas, Nevada, USA, 2010.

[45] Gergely Acs, Claude Castelluccia, and Rui Chen. Differentially private histogram publishing through lossy compression. In *12th IEEE International Conference on Data Mining*, ICDM '12, pages 1–10, Brussels, Belgium, 2012.

[46] Chao Li, Michael Hay, Gerome Miklau, and Yue Wang. A data- and workload-aware algorithm for range queries under differential privacy. *VLDB*, 7(5):341–352, January 2014.

[47] Xiaojian Zhang, Rui Chen, Jianliang Xu, Xiaofeng Meng, and Yingtao Xie. Towards accurate histogram publication under differential privacy. In *Proceedings of the SIAM International Conference on Data Mining*, SDM '14, pages 587–595, Philadelphia, Pennsylvania, USA, 2014.

[48] Boaz Barak, Kamalika Chaudhuri, Cynthia Dwork, Satyen Kale, Frank McSherry, and Kunal Talwar. Privacy, accuracy, and consistency too: A holistic solution to contingency table release. In *Proceedings of the 26th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '07, pages 273–282, Beijing, China, 2007. ACM.

[49] Michael Hay, Vibhor Rastogi, Gerome Miklau, and Dan Suciu. Boosting the accuracy of differentially private histograms through consistency. *VLDB*, 3(1-2):1021–1032, September 2010.

[50] Xiaokui Xiao, Guozhang Wang, and Johannes Gehrke. Differential privacy via wavelet transforms. *IEEE Transactions on Knowledge and Data Engineering*, 23(8):1200–1214, August 2011.

[51] Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava, Entong Shen, and Ting Yu. Differentially private spatial decompositions. In *IEEE 28th International Conference on Data Engineering*, ICDE '12, pages 20–31, Washington, DC, USA, April 2012. IEEE Computer Society.

[52] Bolin Ding, Marianne Winslett, Jiawei Han, and Zhenhui Li. Differentially private data cubes: Optimizing noise sources and consistency. In *Proceedings of the ACM SIGMOD International Conference on Management of Data*, SIGMOD '11, pages 217–228, Athens, Greece, 2011.

[53] Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. Optimizing linear counting queries under differential privacy. In *Proceedings of the 29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, PODS '10, pages 123–134, Indianapolis, Indiana, USA, 2010. ACM.

[54] Xiaoqian Jiang, Zhanglong Ji, Shuang Wang, Noman Mohammed, Samuel Cheng, and Lucila Ohno-Machado. Differential-private data publishing through component analysis. *Transactions Data Privacy*, 6(1):19–34, 2013.

[55] Thee Chanyaswad, Changchang Liu, and Prateek Mittal. Rongauss: Enhancing utility in non-interactive private data release. *PoPETs*, 2019(1):26–46, 2019.

[56] Chugui Xu, Ju Ren, Yaoxue Zhang, Zhan Qin, and Kui Ren. Dppro: Differentially private high-dimensional data release via random projection. *IEEE Transactions Information Forensics and Security*, 12(12):3081–3093, 2017.

[57] Cynthia Dwork, Guy N. Rothblum, and Salil P. Vadhan. Boosting and differential privacy. In *Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '10, pages 51–60, Las Vegas, Nevada, USA, Oct 2010.

[58] Moritz Hardt and Guy N. Rothblum. A multiplicative weights mechanism for privacy-preserving data analysis. In *Proceedings of the 51th Annual IEEE Symposium on Foundations of Computer Science*, FOCS '10, pages 61–70, Las Vegas, Nevada, USA, 2010.

[59] Moritz Hardt, Katrina Ligett, and Frank McSherry. A simple and practical algorithm for differentially private data release. In *Proceedings of the 26th Annual Conference on Neural Information Processing Systems*, NIPS '12, pages 2348–2356, Lake Tahoe,Nevada, USA, 2012.

[60] Ganzhao Yuan, Zhenjie Zhang, Marianne Winslett, Xiaokui Xiao, Yin Yang, and Zhifeng Hao. Low-rank mechanism: Optimizing batch queries under differential privacy. *PVLDB*, 5(11):1352–1363, 2012.

[61] Ganzhao Yuan, Zhenjie Zhang, Marianne Winslett, Xiaokui Xiao, Yin Yang, and Zhifeng Hao. Optimizing batch linear queries under exact and approximate differential privacy. *ACM Transactions Database Systems*, 40(2):11:1–11:47, 2015.

[62] Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Symposium on Theory of Computing Conference*, STOC '13, pages 351–360, Palo Alto, CA, USA, 2013.

[63] Min Lyu, Dong Su, and Ninghui Li. Understanding the sparse vector technique for differential privacy. *PVLDB*, 10(6):637–648, 2017.

[64] 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*, CCS '18, pages 230–246, Toronto, ON, Canada, 2018.

[65] Edward Neuman. Inequalities involving a logarithmically convex function and their applications to special functions. *JIPAM. Journal of Inequalities in Pure & Applied Mathematics [electronic only]*, 7, 01 2000.

[66] Johan Ludwig William Valdemar Jensen. Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta mathematica*, 30(1):175–193, 1906.

[67] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *25th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, pages 486–503, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

Appendix

For a compound PDF, we have

Theorem A.1: Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X be a RV that is distributed according to some parameterized distribution $f(\theta) \in \mathcal{F}$ with an unknown parameter θ that is again distributed according to some other distribution g. The resulting distribution h is said to be the distribution that results from compounding f with g,

$$h(X) = \int_{\mathbb{R}} f(X|\theta) g(\theta) \,\mathrm{d}\,\theta \tag{11}$$

Then for any Borel subset B of $\mathbb{R}$,

$$\mathbb{P}(X \in B) = \int_B \int_{\mathbb{R}} f(X|\theta) g(\theta) \,\mathrm{d}\,\theta dX \tag{12}$$

Theorem 3.1. For an R^{2}DP Laplace mechanism, we have

$$\begin{aligned}
&\mathbb{P}(\mathcal{M}_q(d,b) \in S) \\
&= \int_{\mathbb{R}_{\geq 0}} f(b) \frac{1}{2b} \int_{\mathbb{R}} \mathbb{1}_S\{q(d)+w\} e^{\frac{-|w|}{b}} \; dw \; db \\
&= \int_{\mathbb{R}_{\geq 0}} g(u) \frac{u}{2} \int_{\mathbb{R}} \mathbb{1}_S\{q(d)+w\} e^{-|w|\cdot u} \; dw \; du \\
&= \int_{\mathbb{R}} \mathbb{1}_S\{q(d)+w\} \int_{\mathbb{R}_{\geq 0}} g(u) \frac{u}{2} e^{-|w|\cdot u} \; du \; dw \\
&= \int_{\mathbb{R}} \mathbb{1}_S\{q(d)+w\} \frac{1}{2} \frac{dM_u(t)}{dt}|_{t=-|w|} \; dw \\
&= \tfrac{1}{2} \int_S \frac{dM_u(t)}{dt}|_{t=-|x-q(d)|} dx
\end{aligned} \tag{13}$$

$$= \tfrac{1}{2} \cdot \Big[- M_u(-|x-q(d)|)|_{S_{\geq q(d)}} + M_u(-|x-q(d)|)|_{S_{<q(d)}} \Big]$$
$$= \tfrac{1}{2} \cdot \Big[- M_{\frac{1}{b}}(-|x-q(d)|)|_{S_{\geq q(d)}} + M_{\frac{1}{b}}(-|x-q(d)|)|_{S_{<q(d)}} \Big]$$

where $u = b^{-1}$, is reciprocal of random variable b and $g(u) = \frac{1}{u^2} \cdot f(\frac{1}{u})$. Moreover, $M_u(t)$ is the moment-generating function of random variable u which is identical with $M_{\frac{1}{b}}(t)$. □

Corollary 3.3. Following the DP guarantee in Theorem 3.2, and next, using triangle inequality, we have

$$e^{\epsilon} = \max_{\forall x \in \mathbb{R}} \left\{ \frac{\mathbb{E}(\epsilon \cdot e^{(-|x-q(d)|\cdot\epsilon)})}{\mathbb{E}(\epsilon \cdot e^{(-|x-q(d')|\cdot\epsilon)})} \right\} \leq \max_{\forall t \in \mathbb{R}_{\leq 0}} \left\{ \frac{\mathbb{E}(\epsilon \cdot e^{(t\cdot\epsilon)})}{\mathbb{E}(\epsilon \cdot e^{((t-\Delta q)\cdot\epsilon)})} \right\}$$

Next, we show that the latter is non-decreasing w.r.t. t. First, using Theorem 2.1 in [65], we know that for $f : \mathbb{R}_- \to \mathbb{R}_+$ a differentiable log-convex function, $g(x) = \frac{f'(x)}{f(x)}$ is an decreasing function. Therefore, since a moment generating function is log-convex and differentiable, we have for all t_1, t_2 if $t_1 < t_2 \leq 0$, $\frac{M(t_1)}{M'(t_1)} \geq \frac{M(t_2)}{M'(t_2)}$. This means if $t_1 < t_2 \leq 0$, $\frac{M(t_2)}{M(t_1)} \leq \frac{M'(t_2)}{M'(t_1)}$. However, we know that if $t_1 < t_2 \leq 0$, $M(t_1) < M(t_2)$, and hence, we can easily conclude that $M'(t)$ is a strictly increasing function. Finally, evaluating $e^{\epsilon(t)}$ at $t = 0$, concludes our proof. □

Theorem 4.2. Following Theorem 2.2, an ϵ-differentially private Laplace mechanism is $(\gamma, e^{\frac{-\gamma}{b(\epsilon)}})$-useful for all $\gamma \geq 0$, where $b(\epsilon) = \frac{\Delta q}{\epsilon}$. Therefore, for the usefulness of the baseline Laplace mechanism at $\epsilon = \ln[\mathbb{E}_{\frac{1}{b}}(e^{\epsilon(b)})]$, we have

$$e^{\frac{-\gamma \cdot \ln[\mathbb{E}_{\frac{1}{b}}(e^{\epsilon(b)})]}{\Delta q}} = \big(\mathbb{E}_{\frac{1}{b}}(e^{\epsilon(b)})\big)^{\frac{-\gamma}{\Delta q}} = \big(\mathbb{E}_{\frac{1}{b}}(e^{\frac{\Delta q}{b}})\big)^{\frac{-\gamma}{\Delta q}} \leq \mathbb{E}_{\frac{1}{b}}\big(e^{\frac{-\gamma}{b}}\big)$$

where the last inequality relation is verified by Jensen inequality [66] as $g(x) = x^{\frac{-\gamma}{b}}$ is a convex function. Recall the following Jensen inequality: Let $(\Omega, \mathfrak{F}, \mathrm{P})$ be a probability space, X an integrable real-valued random variable and g a convex function. Then

$$g(\mathbb{E}(X)) \leq \mathbb{E}(g(X))$$

Therefore,

$$1 - e^{\frac{-\gamma \cdot \ln[\mathbb{E}_{\frac{1}{b}}(e^{\epsilon(b)})]}{\Delta q}} \geq 1 - \mathbb{E}_{\frac{1}{b}}\big(e^{\frac{-\gamma}{b}}\big) = U(\ln[\mathbb{E}_{\frac{1}{b}}(e^{\epsilon(b)})], \Delta q, \gamma)$$

□

Theorem 4.3. For $\frac{1}{b} \sim f_{\delta, \frac{1}{b_0}}(\frac{1}{b})$, the MGF is given by $M_{\frac{1}{b}}(t) = e^{\frac{t}{b_0}}$. Following Theorem 3.2, one can write

$$\begin{aligned}
e^{\epsilon} &= \max_{\forall x \in \mathbb{R}} \left\{ \frac{\frac{1}{b_0} \cdot e^{\frac{-|x-q(d)|}{b_0}}}{\frac{1}{b_0} \cdot e^{\frac{-|x-q(d')|}{b_0}}} \right\} = \max_{\forall x \in \mathbb{R}} \left\{ e^{\frac{|x-q(d')|-|x-q(d)|}{b_0}} \right\} \\
&\leq \max_{\forall x \in \mathbb{R}} \left\{ e^{\frac{|q(d)-q(d')|}{b_0}} \right\} = e^{\frac{\Delta q}{b_0}}
\end{aligned}$$

where the last inequality is resulted form triangle inequality. □

Theorem 4.4. The R^{2}DP Laplace mechanism $\mathcal{M}_q(d,b)$, $\frac{1}{b} \sim f_{B, \frac{1}{b_0}, \frac{1}{b_1}}(\frac{1}{b}; p)$ returns with probability p, a Laplace mechanism with scale parameter b_1, and with probability $1-p$ another Laplace mechanism with scale parameter b_2. To this end, we are looking for

$$e^{\epsilon} = \max_{\forall x \in \mathbb{R}} \left\{ \frac{\frac{p}{b_0} \cdot e^{\frac{-|x-q(d)|}{b_0}} + \frac{1-p}{b_1} \cdot e^{\frac{-|x-q(d)|}{b_1}}}{\frac{p}{b_0} \cdot e^{\frac{-|x-q(d')|}{b_0}} + \frac{1-p}{b_1} \cdot e^{\frac{-|x-q(d')|}{b_1}}} \right\}$$

Therefore, using triangle inequality, we have

$$\begin{aligned}
e^{\epsilon_1} &= \max_{\forall S \in \mathbb{R}} \left\{ \frac{p \cdot e^{\frac{-|x-q(d)|}{b_0}} + (1-p) \cdot e^{\frac{-|x-q(d)|}{b_1}}}{p \cdot e^{\frac{-|x-q(d')|}{b_0}} + (1-p) \cdot e^{\frac{-|x-q(d')|}{b_1}}} \right\} \\
&\leq \max_{\forall x \geq q(d)} \left\{ \frac{p \cdot e^{\frac{\Delta q - |x-q(d')|}{b_0}} + (1-p) \cdot e^{\frac{\Delta q + -|x-q(d')|}{b_1}}}{p \cdot e^{\frac{-|x-q(d')|}{b_0}} + (1-p) \cdot e^{\frac{-|x-q(d')|}{b_1}}} \right\}
\end{aligned}$$

Let us make the substitutions $X = e^{\frac{-|x-q(d')|}{b_0}}$, $a = e^{\frac{\Delta q}{b_0}}$ and k=$\frac{b_0}{b_1} > 1$. Hence, we have

$$e^{\epsilon} \leq \max_{\forall X \in (0,1)} \left\{ \frac{p \cdot a \cdot X + (1-p) \cdot (a \cdot X)^k}{p \cdot X + (1-p) \cdot X^k} \right\}$$

To obtain e^{ϵ}, we need to find all the critical points of $e^{\epsilon_1}(X) = \frac{p \cdot a \cdot X + (1-p) \cdot (a \cdot X)^k}{p \cdot X + (1-p) \cdot X^k}$. However, the critical points

of a fractional function are the roots of the numerator of its derivative. Hence, suppose

$$\frac{\mathrm{d}e^{\epsilon}(X)}{\mathrm{d}X} = \frac{N(X)}{D(X)}$$

then

$$\begin{aligned}\Rightarrow N(X) &= \big(p\cdot a + (1-p)\cdot k\cdot a\cdot (a\cdot X)^{k-1}\big)\\ \cdot\big(p\cdot X + (1-p)\cdot X^k\big) &- \big(p + (1-p)\cdot k\cdot X^{k-1}\big)\\ \cdot\big(p\cdot a\cdot X + (1-p)\cdot (a\cdot X)^k\big)&\\ = p\cdot(1-p)\cdot(k-1)\cdot(a^{k-1}-1)\cdot X^k&\end{aligned}$$

However, all the terms in the last expression are strictly positive. Therefore, the only critical points are $X = 0$ and $X = 1$ and as the function is strictly increasing,

$$\begin{aligned}e^{\epsilon} \le e^{\epsilon}(1) &= p\cdot a + (1-p)\cdot (a)^k\\ &= p\cdot e^{\frac{\Delta q}{b_0}} + (1-p)\cdot e^{\frac{\Delta q}{b_1}}\end{aligned}$$

which is the bound in the Theorem.

□

Theorem 4.5. For a Gamma distribution with shape parameters k and scale parameters θ, the MGF at point t is given as $(1-\theta\cdot t)^{-k}$. Since $\frac{1}{b} \sim f_{\Gamma}(\frac{1}{b}; k, \theta)$, following Theorem 3.2, one can write

$$\begin{aligned}e^{\epsilon} &= \max_{\forall x\in\mathbb{R}} \left\{ \frac{k\cdot\theta\cdot(1+\theta\cdot|x-q(d)|)^{-k-1}}{k\cdot\theta\cdot(1+\theta\cdot|x-q(d')|)^{-k-1}} \right\}\\ \Rightarrow \epsilon &= \max_{\forall x\in\mathbb{R}} \left\{ (k+1)\cdot \ln\left[\frac{(1+\theta\cdot|x-q(d')|)}{(1+\theta\cdot|x-q(d)|)} \right] \right\}\end{aligned}$$

to find the maximum of the ln term, denote by $X = 1 + \theta\cdot|x-q(d)|$. Moreover, since $|x-q(d')| \le |x-q(d)| + \Delta q$, we have

$$\Rightarrow \epsilon \le \max_{\forall X\ge 1} \left\{ \frac{X + \Delta q\cdot\theta}{X} \right\}$$

However, since

$$\forall X \ge 1,\ \frac{X + \Delta q\cdot\theta}{X}$$

is strictly decreasing, we have

$$\Rightarrow \epsilon = (k+1)\cdot\ln\big[1 + \theta\cdot\Delta q\big].$$

□

Lemma 4.10. Using some exhaustive search, suppose $\mu = 0.5223$,$\sigma = 1.5454$, $a = 0.5223$ and for $\epsilon = 1.1703$ and $\Delta q = 0.6$, we will get $\ln(M_{\mathcal{N}^T}(\Delta q)) = 1.2417$. □

Figures 10 and 10, respectively, illustrate the RoI given by the Gamma distribution and the truncated Gaussian distribution. The focus of each distribution has already been shown in section V.

R^{2}DP Gaussian Mechanism. A differentially private mechanism proposed in [67] modifies an answer to a numerical query by adding iid zero-mean Gaussian noise.

Algorithm 1: Usefulness Optimization

Input : dataset D, privacy budget ϵ (Data Owner) query $q(\cdot)$, usefulness γ (Data Recipient)
Output: query result $q(D) + Lap(b_r)$ which satisfies ϵ-DP and is maximally γ-useful

Function: $\mathcal{F}(q(D), \epsilon, \gamma, a_1^{opt}, a_2^{opt}, \lambda^{opt}, \theta^{opt}, \mu^{opt}, \sigma^{opt}, a^{opt})$
$\Delta q \leftarrow$ Sensitivity $(q(\cdot))$
Data recipient generates the optimal parameters from Lagrange Multiplier $\{a_1^{opt}, a_2^{opt}, \lambda^{opt}, \theta^{opt}, \mu^{opt}, \sigma^{opt}, a^{opt}\}$
$X_1 \sim \Gamma(\lambda^{opt}, \theta^{opt})$
$X_2 \sim \mathcal{N}^T(\mu^{opt}, \sigma^{opt}, a^{opt})$
$\frac{1}{b_r} = a_1^{opt}\cdot X_1 + a_2^{opt}\cdot X_2$: Reciprocal of randomized scale parameter b_r
return $q(D) + Lap(b_r)$

Recall the definition of the $\mathcal{Q}$-function $\mathcal{Q}(x) := \frac{1}{\sqrt{2\pi}}\int_x^{\infty} e^{-\frac{u^2}{2}}du$, We have the following theorem [31], [67].

Theorem A.2: Let $q : \mathsf{D} \to \mathbb{R}$ be a query and $\epsilon > 0$. Then the Laplace mechanism $\mathcal{M}_q : \mathsf{D}\times\Omega \to \mathbb{R}$ defined by $\mathcal{M}_q(d) = q(d) + w$, with $w \sim \mathcal{N}\big(0, \sigma^2\big)$, where $\sigma \ge \frac{\Delta q}{2\epsilon}(K + \sqrt{K^2 + 2\epsilon})$ and $K = \mathcal{Q}^{-1}(\delta)$, is (ϵ, δ)-differentially private.

For the rest of the paper, we define $\kappa_{\delta,\epsilon} = \frac{1}{2\epsilon}(K + \sqrt{K^2+2\epsilon})$, so that the standard deviation σ in Theorem A.2 can be written as $\sigma(\delta,\epsilon) = \kappa_{\delta,\epsilon}\Delta q$. It can be shown that $\kappa_{\delta,\epsilon}$ behaves roughly as $O(\ln(1/\delta))^{1/2}/\epsilon$. For example, to guarantee (ϵ,δ)-differential privacy with $\epsilon = \ln(2)$ and $\delta = 0.05$, the standard deviation of the Gaussian noise introduced should be about 2.65 times the ℓ_1-sensitivity of q.

Theorem A.3: The Gaussian Mechanism A.2 is $(\gamma, 2\cdot \mathcal{Q}\big(\frac{\gamma}{\sigma(\delta,\epsilon)}\big))$-useful.

Similar to our R^{2}DP Laplace mechanism, we can formulate an optimization problem for the R^{2}DP model using Gaussian mechanism. Therefore, using Theorems A.2 and A.3, we have the following.

Corollary A.4: Denote by u, the set of parameters for a probability distribution f_σ. Then, the optimal usefulness of an R^{2}DP Gaussian mechanism utilizing f_σ, at each quadruplet $(\epsilon, \delta, \Delta q, \gamma)$ is

$$\begin{aligned}U_f(\epsilon,\delta,\Delta q,\gamma) = \max_{u\in\mathbb{R}^{|u|}} &\big(1 - 2\cdot\mathbb{E}_\sigma(Q\big(\tfrac{\gamma}{\sigma(\delta,\epsilon)}\big))\big)\\ &\text{subject to}\\ \max_{\forall S\in\mathsf{R}} &\left\{\frac{\mathbb{P}(\mathcal{M}_q(d,\sigma)\in S)}{\mathbb{P}(\mathcal{M}_q(d',\sigma)\in S)}\right\} = \epsilon,\\ &\mathbb{E}_\sigma(Q(\epsilon\sigma - \tfrac{1}{2\sigma})) = \delta\end{aligned} \tag{14}$$

We leave further discussion for R^{2}DP Gaussian mechanism to future works.

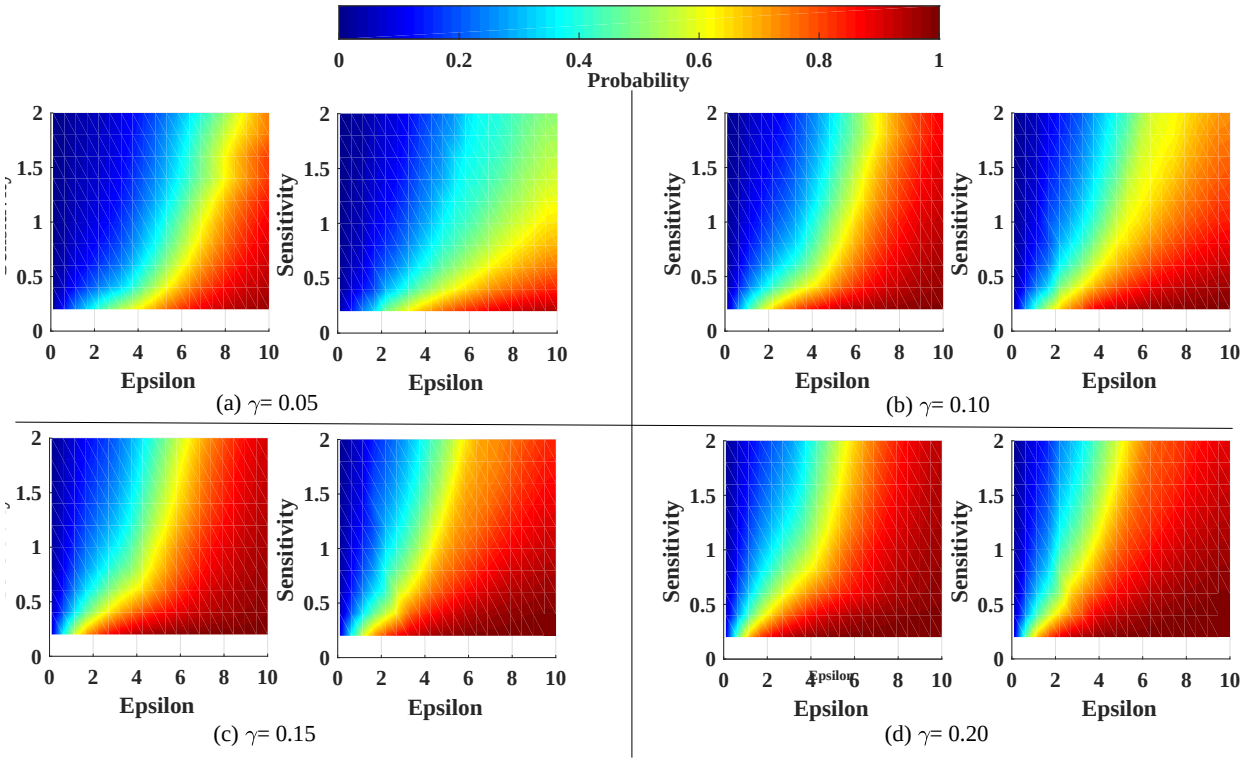


Fig. 10. Comparison between the performances of R^2DP and optimal noise, i.e., Laplace distribution in high privacy regime and Staircase shape distribution in low privacy regime.