Randomized Privacy Budget Differential Privacy

Meisam Mohammady

December 2018

1 Abstract

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 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. However, 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. In this report we present DP mechanisms with randomized parameters, i.e., randomized privacy budget, and formally analyze its privacy and utility and demonstrate that randomizing privacy budget in DP mechanisms will boost the accuracy in a humongous scale.

2 Backgrounds

Definition 2.1. Let $\epsilon, \delta \geq 0$. A mechanism $M : D \times \Omega \rightarrow \mathbb{R}$ is $(\epsilon, \delta)$-differentially private for Adj if for all $d, d' \in D$ such that $\text{Adj}(d, d')$, we have

$$\mathbb{P}(M(d) \in S) \leq e^{\epsilon} \mathbb{P}(M(d') \in S) + \delta, \quad \forall S \in \mathcal{M}. \quad (1)$$

If $\delta = 0$, the mechanism is said to be $\epsilon$-differentially private.

Definition 2.2. (Usefulness Definition). A database mechanism $M_q$ is $(\zeta, \gamma(\zeta, u))$-useful if with probability $1 - \gamma(\zeta, u)$, for every database $d \subseteq D$, $|M_q(d) - q(d)| \leq \zeta$.

3 Randomized Parameter DP

Let $M_q(d, u) = q(d) \oplus \omega(u)$ be a randomized $(\epsilon(u), \delta(u))$-differentially private mechanism where $\omega(u)$ is a random oracle with specified set of parameters $u$ and $\oplus$ stands for the corresponding operator. Also suppose $M_q(d, u)$ is $(\zeta, \gamma(\zeta, u))$-useful. Define by $M_q(d) = q(d) \oplus \omega(u)$, with $u \sim \mathcal{F}$, the distribution of all
possible randomized mechanism \( M_q(d, u) \) where \( F \) is a probability density function for all parameters in \( u \). The optimal utility achieved due to the application of optimal pdf \( F \) is shown in the following.

\[
U(\zeta) = \text{Max} \{ E[|M_q(d, u) - q(d)| < \zeta]\} \tag{2}
\]

Accordingly, we say that \( M_q(d) \) improves the privacy-utility trade-off if we have

- **Case I** (\( \delta = 0 \))
  \[
  E(e^u) = e^{\epsilon(u_0)} \Rightarrow U(\zeta) > 1 - \gamma(\zeta, u_0) \tag{3}
  \]
  over

- **Case II** (\( \delta > 0 \))
  \[
  E(\delta(u)) = \delta(u_0) \Rightarrow U(\zeta) > 1 - \gamma(\zeta, u_0) \tag{4}
  \]
  \[
  E(e^u) < e^{\epsilon(u_0)} \tag{5}
  \]

where, \( E(.) \) denotes the expected value over distribution \( F \). We now derive the corresponding conditions for two popular differentially private mechanisms. In particular, a Laplace Mechanism modifies an answer to a numerical query by adding independent and identically distributed (i.i.d.) zero-mean noise distributed \([1, 3, 2]\) according to a Laplace distribution. Recall that the Laplace distribution with mean zero and scale parameter \( b \), denoted \( \text{Lap}(b) \), has density

\[
p(x; b) = \frac{1}{2b} \exp(-\frac{|x|}{b}) \quad \text{and variance } 2b^2.
\]

Moreover, for \( \omega \in \mathbb{R}^k \) with \( \omega_i \) i.i.d. and \( \omega_i \sim \text{Lap}(b) \), denoted \( \omega \sim \text{Lap}(b)^k \), we have

\[
p(\omega; b) = \left(\frac{1}{2b}\right)^k \exp(-\frac{||\omega||_1}{b}),
\]

\[
E(||\omega||_1) = b, \quad \text{and } P(||\omega||_1 \geq tb) = e^{-t}.
\]

**Theorem 3.1.** Let \( q : D \rightarrow \mathbb{R}^k \) be a query, \( \epsilon > 0 \). Then the mechanism \( M_q : D \times \Omega \rightarrow \mathbb{R}^k \) defined by \( M_q(d) = q(d) + w \), with \( w \sim \text{Lap}(b)^k \), where \( b \geq \frac{\Delta}{\epsilon^2} \) is \( \epsilon \)-differentially private.

Hence, \( \omega(u) \) is a Laplace distribution where \( u = 1/b = \epsilon \). Also, \( \gamma(\zeta, u_0) = e^{-\zeta \epsilon} \). Thus, equations \( 2, 7 \) can be re-written as follows.

\[
U(\zeta) = \text{Max} \ E(1 - e^{-\zeta \epsilon}) \tag{6}
\]

\[
E(e^\epsilon) = e^{\epsilon_0} \Rightarrow U(\zeta) > 1 - e^{-\zeta \epsilon} \tag{7}
\]

Similarly, for a Gaussian mechanism, we have

\[
\text{Min} \ E(e^\epsilon) = \int_{-\infty}^{\infty} f(\epsilon)e^\epsilon d\epsilon \tag{8}
\]

over

\[
E(Q\left(\frac{\zeta}{\sqrt{2\epsilon + 1}}\right)) = \int_{-\infty}^{\infty} f(\epsilon)Q\left(\frac{\zeta}{\sqrt{2\epsilon + 1}}\right)d\epsilon \leq \delta \tag{9}
\]
4 Privacy and Utility Analysis

In this section, we formally characterize the privacy and the utility of the Randomized DP mechanism.

4.1 Deriving PDF of Randomized DP

we can write the CDF of the output of an Randomized DP Laplace mechanism in terms of the Moment Generating Function (MGF) for the probability distribution $f_{\frac{1}{\epsilon}}$. 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. In particular,

Definition 4.1. (Moment Generating Function) The moment-generating function of a random variable $x$ is $M_X(t) := E[e^{tX}], t \in \mathbb{R}$ wherever this expectation exists. The moment-generating function is the expectation of the random variable $e^{tX}$.

Accordingly, in the following, we give a general formula for the probability of any measurable event originated from an Randomized DP Laplace Mechanism.

Theorem 4.1. The search space of an Randomized DP Laplace mechanism is as large as the space of all PDFs with non-negative support and existing MGF. Moreover, generated PDFs are all log-convex.

Thus, for a PDF with non-negative support (scale parameter is always non-negative), the Randomized DP Laplace mechanism outputs another PDF using the MGF (CDF is the moment and PDF is its derivative) as shown in Equation 11 in Appendix [3]. However, a challenge is that not all random variables have moment generating functions (MGFs). Fortunately, MGFs possess an appealing composability property between independent probability distributions, which can be used to provide us with a search space of all linear combinations of a set of popular distributions with known MGFs (infinite number of RVs).

Theorem 4.2 (MGF of Linear Combination of RVs). If $x_1, x_2, \ldots, x_n$ are $n$ independent RVs with respective MGFs $M_{x_i}(t) = E(e^{tx_i})$ for $i = 1, 2, \ldots, n$, then the MGF of the linear combination $Y = \sum_{i=1}^{n} a_i x_i$ is $\prod_{i=1}^{n} M_{x_i}(a_i t)$.

Thus, our search space is given as all possible linear combinations of a set of independent RVs with existing MGF (Section 4.5 demonstrates on how to choose the set of independent RVs).

4.1.1 Determining the Optimal PDF

After giving the differential privacy guarantee and characterizing the utility of the Randomized DP Laplace mechanism (see Section 4), we will show that the
Randomized DP framework can unify two parallel concepts, i.e., privacy and utility, into one optimization problem defined over the defined search space of RVs.

4.2 Numerical Analysis

We now present numerical results to fine tune the Randomized DP parameters under more general settings. In particular, Figure 1 depicts the corresponding performance of Laplace mechanism, Randomized DP Laplace mechanism and Staircase mechanism. Figure 1 clearly demonstrates the fact that Randomized DP can achieve both objectives mentioned earlier, i.e., approaching the optimal mechanism and improving Laplace mechanisms for larger $\epsilon$. We now analyze the improvements provided by Randomized DP under two different settings. First, we discuss the performance of Randomized DP under a stronger privacy guarantee (e.g., $\epsilon < 2$). Next, we study the improvement for counting queries ($\Delta q = 1$) while varying the error bound $\gamma$.

4.3 Privacy Analysis

We now show the Randomized DP Laplace mechanism provides differential privacy guarantee. Using theorem 4.1, the DP bound is

$$
e^\epsilon = \max_{\forall S \in \mathbb{R}} \left\{ -M_1 \frac{(-|x-q(d)|)}{|S \geq q(d)} + M_1 \frac{(-|x-q(d')|)}{|S \leq q(d')} - M_1 \frac{(-|x-q(d)|)}{|s \geq q(d')} + M_1 \frac{(-|x-q(d')|)}{|s \leq q(d')} \right\}
$$
Figure 2: Validating the effectiveness of Randomized DP for small $\epsilon$

Hence, the value of $e^\epsilon$ only depends on the distribution of reciprocal of the scale parameter $b$, i.e., $f_\frac{1}{b}$. Moreover, an MGF is positive and log-convex 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 $P(\{q(d) + \text{Lap}(b)\} \in S)$ in Equation 9 can be easily applied to calculate the differential privacy guarantee.

**Theorem 4.3.** The Randomized DP mechanism $M_q(d, b)$ is

$$\ln \left[ \frac{\mathbb{E}(\frac{1}{b})}{\frac{dM_t}{dt}|_{t=-\Delta q}} \right] - \text{differentially private.} \quad (10)$$

Finally, Theorem 4.2 can be directly applied to calculate the differential privacy guarantee of any RV from our defined search space (all linear combinations of a set of independent RVs with known MGFs).

**Corollary 4.4** (differential privacy of combined PDFs). If $x_1, x_2, \cdots, x_n$ are $n$ independent random variables with respective MGFs $M_{x_i}(t) = \mathbb{E}(e^{t x_i})$ for $i = 1, 2, \cdots, n$, then the Randomized DP mechanism $M_q(d, b)$ where $\frac{1}{b}$ is defined as the linear combination $\frac{1}{b} = \sum_{i=1}^{n} a_i x_i$ is
Figure 3: Comparing the performance of (a) the baseline Laplace mechanisms, for count queries and (b) Randomized DP, when varying $\gamma$ and $\epsilon$

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

- differentially private.

Therefore, we have established a search space of probability distributions with a universal formulation for their differential privacy guarantees, which is the key enabler for the universality of Randomized DP. Next, we characterize the utility of Randomized DP Laplace mechanisms.

### 4.4 Characterizing the Utility

We now characterize the utility of the Randomized DP Laplace mechanism. To make concrete discussions, we first focus our discussion on the usefulness metric (see Section ??), then discuss how a similar logic applies to other metrics. Denote by $U(\epsilon, \Delta q, \gamma)$ the usefulness of a Randomized DP Laplace mechanism for all $\epsilon > 0$, sensitivity $\Delta q$ and error bound $\gamma$. The optimal usefulness is then given as the answer of the following optimization problem over the search space of PDFs.

\[
\max_{f \in F} \left\{ U(\epsilon, \Delta q, \gamma) \right\} = \max_{f \in F} \left\{ \frac{1}{2} \left[ - M_{x}( - |x - q(d)| )^{q(d)+\gamma} + M_{x}( - |x - q(d)| )^{q(d)-\gamma} \right] \right\},
\]

subject to $\epsilon = \ln \left[ \frac{E(\frac{1}{b})}{dM_{x}(t)} \right]_{t=-\Delta q}$
Note that $\epsilon$ and $\Delta q$ do not directly impact the usefulness but they do so indirectly through the differential privacy constraint. Furthermore, as shown in Theorem 4.3, the differential privacy guarantee $\epsilon$ over the established search space $\mathcal{F}$ is a unique function of the parameters of the second fold distribution.

**Corollary 4.5.** Denote by $u$, the set of parameters for a probability distribution $f_{1b}$, and by $M_{f(u)}$ its MGF. Then, the optimal usefulness of an Randomized DP mechanism utilizing $f_{1b}$, at each triplet $(\epsilon, \Delta q, \gamma)$ is

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

subject to $\epsilon = \ln \left[ \frac{\mathbb{E}(\frac{1}{t})}{\frac{dM_{1b}(t)}{dt}}|_{t=-\Delta q} \right]$.

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

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

Next, under the DP guarantee of several probability distributions, we will apply Equation 12 to find the optimal trade-off.

**Utility under Other Metrics.** We derive the utility of the Randomized DP Laplace mechanism under some well-known utility metrics. Due to space limitation, we present only the final results in Table 1.

```
| Metric | Usefulness |
|--------|------------|
| $\ell_1$ | $\int_{-\infty}^{\infty} M_{1b}(-x) \cdot |x - q(d)| \cdot \ln M_{1b}(-x) dx$ |
| $\ell_2$ | $\int_{-\infty}^{\infty} M_{1b}(-x) \cdot (x - q(d))^2 dx$ |
| Entropy | $\int_{-\infty}^{\infty} M_{1b}(-x) \cdot \ln M_{1b}(-x) dx$ |
```

The results in Table 1 can be easily applied to optimize each measure in different applications.

**Necessary Condition on Selected Distributions.** Not all second fold probability distributions can boost the utility of the baseline Laplace mechanism. Accordingly, in the following theorem, we derive a necessary condition on the differential privacy guarantee of an Randomized DP Laplace mechanism to boost...
the utility of the baseline Laplace mechanism (refer to Appendix in [3] for the proof). Using this necessary condition, we can easily filter out those probability distributions that cannot deliver any utility improvement.

**Theorem 4.6.** The utility of an Randomized DP Laplace mechanism with $\epsilon \geq \ln \left[ \mathbb{E}_\frac{1}{b}(e^{\epsilon(b)}) \right]$ is always upper bounded by the utility of the $\epsilon$-differentially private baseline Laplace mechanism. Equivalently, for an Randomized DP Laplace mechanism to boost the utility, the following relation is necessarily true.

$$e^\epsilon = \frac{\mathbb{E}(\frac{1}{b})}{M'_1(-\Delta q)} < M_2(\Delta q)$$

(13)

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

### 4.5 Finding Utility-Maximizing Probability Distributions

We now examine a set of well-known probability distributions to establish the required search space by selecting those offer a significantly improved $\epsilon$ compared with the bound given in Theorem 4.6. Promisingly, our analytic evaluations for three of these distributions, i.e., Gamma, Uniform and truncated Gaussian distributions demonstrates such a payoff. Finally, we note that those chosen distributions 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.

#### 4.5.1 Discrete Probability Distributions

First, we consider two different mixture Laplace distributions that can be applied for constructing Randomized DP Laplace mechanisms with discrete probability distribution $f_b$.

(1) **Degenerate distribution.** A degenerate distribution is a probability distribution in a (discrete or continuous) space with support only in a space of lower dimension. If the degenerate distribution is uni-variate (involving only a single random variable) it will be 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 $\delta_{k_0}$ is given by $M_k(t) = e^{t k_0}$. Using Equation 10, Theorem 4.7 gives the same DP guarantee as the baseline Laplace mechanism.
Theorem 4.7. The Randomized DP Laplace mechanism $M_q(d, \epsilon)$, $\epsilon \sim f_{\frac{1}{b_0}}(\epsilon)$, is $\Delta_{\frac{b_0}{b}}$-differentially private.

Obviously, this distribution does not improve the bound in Theorem 4.6 but shows 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 $M_X(t) = p \cdot e^{t \cdot X_0} + (1 - p) \cdot e^{t \cdot X_1}$. We now derive the precise differential privacy guarantee of an Randomized DP Laplace mechanism with its scale parameter randomized according to a Bernoulli distribution.

Theorem 4.8. The Randomized DP Laplace mechanism $M_q(d, \epsilon)$, $\epsilon \sim f_{\frac{1}{b_0}}(\epsilon; p)$, satisfies $\ln[p \cdot e^{\frac{\Delta q \cdot b_0}{b}} + (1 - p) \cdot e^{\frac{\Delta q \cdot b_1}{b}}]$-DP.

However, this bound is exactly the mean value of $e^{\epsilon(b)}$ and therefore, this distribution does not improve the bound given in Theorem 4.6 either.

4.5.2 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 $\theta$. 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 $E(X) = k\theta = \alpha/\beta$ is fixed and greater than zero, and $E[\ln(X)] = \psi(k) + \ln(\theta) = \psi(\alpha) - \ln(\beta)$ is fixed ($\psi$ is the digamma function). Therefore, it may provide a relatively higher privacy-utility trade-off in comparison to the other candidates. A random variable $X$ that is gamma-distributed with shape $\alpha$ and rate $\beta$ is denoted by $X \sim \Gamma(k, \theta)$ and the corresponding PDF 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 privacy guarantee provided by assuming that the reciprocal of the scale parameter $b$
in Laplace mechanism is distributed according to the gamma distribution (see Appendix [3] for the proof).

**Theorem 4.9.** The Randomized DP Laplace mechanism $M_q(d, \epsilon), \epsilon \sim f_\Gamma(\epsilon; k, \theta)$, satisfies $((k + 1) \cdot \ln(1 + \Delta q \cdot \theta))$-DP.

We now apply the necessary condition given in Equation [13]

**Lemma 4.10.** Randomized DP using Gamma distribution can satisfy the necessary condition in Equation [13].

Proof. We need to show that there exist $k$ and $\theta$ such that $(k + 1) \cdot \ln(1 + \Delta q \cdot \theta) < -k \cdot \ln(1 - \Delta q \cdot \theta)$, $\theta < \frac{1}{\Delta q}$. Given $\theta = \frac{1}{\Delta q}$, we need to show that $\exists k, k \cdot \ln(2)$, which always holds for all $k > 1.4094$. \[\square\]

Therefore, Gamma distribution may improve over the baseline, and this can be computed by optimizing the privacy-utility trade-off using the Lagrange multiplier function in Equation [12]. Also, our optimization shows that, this distribution is more effective for large $\epsilon$ (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 support of the distribution are equally probable. The support is defined by the two parameters, $a$ and $b$, which are the minimum and maximum values. The distribution is often abbreviated as $U(a, b)$, which 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. The MGF for $U(a, b)$ is

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

Using Theorem [4.3], we now drive the precise differential privacy guarantee of an Randomized DP Laplace mechanism for uniform distribution $U(a, b)$.

**Theorem 4.11.** The Randomized DP Laplace mechanism $M_q(d, \epsilon), \epsilon \sim f_U(a, b)(\epsilon)$, is $\ln(\alpha + \beta)$-differentially private, where $\alpha = a - \Delta q$ and $\beta = b - \Delta q$.

We now apply the necessary condition given in Equation [13]. One can easily verify that the inequality holds for infinite number of settings, e.g., $a = 0.5, b = 9$ and $\Delta q = 1.2$.

**Lemma 4.12.** Randomized DP using uniform distribution can satisfy the necessary condition in Equation [13].

Therefore, Randomized DP using uniform distribution may improve over the baseline, and this can be computed by optimizing the privacy-utility trade-off.
using the Lagrange multiplier function in Equation 12. Also, our simulation shows that, this distribution can also be effective for both small and large $\epsilon$.

(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 $X \in (a, b)$, $-\infty < a < b \leq \infty$. Then, $X$ conditional on $a < X < b$ has a truncated Gaussian distribution with the following probability density function.

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

and by $f_{\mathcal{N}^T} = 0$ otherwise. Here, $\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$ and $\Phi(x) = 1 - Q(x)$ are PDF and CDF of the standard Gaussian distribution, respectively. Next, using Theorem 4.13, we give the differential privacy guarantee provided by the mechanism assuming that the reciprocal of $b$ is distributed according to the truncated Gaussian distribution.

**Theorem 4.13.** The Randomized DP Laplace mechanism $\mathcal{M}_q(d, \epsilon)$, and $\epsilon \sim f_{\mathcal{N}^T}(\epsilon; \mu, \sigma, a, b)$, satisfies $\epsilon_{\mathcal{N}^T}$-DP, where

$$\epsilon_{\mathcal{N}^T} = \ln \left[ \frac{\mu + \sigma \cdot (\phi(\alpha) - \phi(\beta))}{\Phi(\beta) - \Phi(\alpha)} \right]$$

$$\frac{dM_{\mathcal{N}^T}(t)}{dt} \big|_{t = -\Delta q} \quad \text{(14)}$$

in which $\phi(\cdot)$ 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 4.14.** Randomized DP using truncated Gaussian distribution can satisfy the necessary condition in Equation 13.

Therefore, truncated Gaussian distribution may improve over the baseline, and this can be computed by optimizing the privacy-utility trade-off using the Lagrange multiplier function in Equation 12. In particular, our simulation shows that, this distribution can also be effective for smaller $\epsilon$ (stronger privacy guarantees).

### 4.6 Expanding the Search Space with Combined PDFs

Theorem 4.2 can be directly applied to design a utility-maximizing Randomized DP Laplace mechanism with a sufficiently large search space (infinite number of different random variables). Since the Laplace mechanism has already been studied under $\ell_1, \ell_2$ and entropy, we will focus on the usefulness metric.
Corollary 4.15 (Optimal Utility for Combined RVs). If $x_1, x_2, \cdots, x_n$ are $n$ independent random variables with respective MGFs $M_{x_i}(t) = \mathbb{E}(e^{tx_i})$ for $i = 1, 2, \cdots, n$, then for the linear combination $Y = \sum_{i=1}^n a_i x_i$, the optimal usefulness (similar relation holds for other metrics) under $\epsilon$-differential privacy constraint is given as

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

subject to

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

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 case of single RVs, we can compute the optimal solution for this optimization problem using the Lagrange multiplier function in Equation 12.

We will focus on all RVs that are produced using linear combinations of the Gamma, uniform and truncated Gaussian distributions (which include both weak and strong privacy-preserving PDFs). Therefore, the corresponding Lagrange multiplier function is

$$\mathcal{L}(a_1, a_2, k, \theta, a_u, b_u, \mu, \sigma, a_{N_T}, b_{N_T}, \Lambda) = M_{T(k, \theta)}(-a_1 \gamma) \cdot M_{U(a_u, b_u)}(-a_2 \gamma) \cdot M_{N_T}(\mu, \sigma, a_{N_T}, b_{N_T})(-a_3 \gamma) + \Lambda \cdot (\ln \left[ N \right] - \epsilon)$$

where the numerator and the denominator $N$, $D$ are

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

$$D = a_1 \cdot M'_{T(k, \theta)}(-a_1 \cdot \Delta q) \cdot M_{U(a_u, b_u)}(-a_2 \cdot \Delta q) \cdot M_{N_T}(\mu, \sigma, a_{N_T}, b_{N_T})(-a_3 \cdot \Delta q)$$

$$+ a_2 \cdot M_{T(k, \theta)}(-a_1 \cdot \Delta q) \cdot M'_{U(a_u, b_u)}(-a_2 \cdot \Delta q) \cdot M_{N_T}(\mu, \sigma, a_{N_T}, b_{N_T})(-a_3 \cdot \Delta q)$$

$$+ a_3 \cdot M_{T(k, \theta)}(-a_1 \cdot \Delta q) \cdot M'_{U(a_u, b_u)}(-a_2 \cdot \Delta q) \cdot M_{N_T}(\mu, \sigma, a_{N_T}, b_{N_T})(-a_3 \cdot \Delta q)$$
Finally, Algorithm 1 details our Randomized DP Laplace mechanism using linear combinations of these three PDFs. In Section ??, using experiments and simulation results, we show that Algorithm 1 can indeed outputs near-optimal results.

**Input**: Dataset $D$, Privacy budget $\epsilon$, Query $q(\cdot)$, Metric and its parameters (from data recipient)

**Output**: Query result $q(D) + \text{Lap}(b_r)$ which is $\epsilon$-DP and is near-optimal w.r.t. the utility requirement

$\Delta q \leftarrow$ Sensitivity ($q(\cdot)$)

Find optimal parameters from Lagrange Multiplier $\mathcal{L}(\epsilon, \Delta q, \text{metric}) = \{a_1^{opt}, a_2^{opt}, a_3^{opt}, k^{opt}, \theta^{opt}, a_u^{opt}, k_u^{opt}, \mu^{opt}, \sigma^{opt}, a_{NT}^{opt}, b_{NT}^{opt}\}$

$X_1 \sim \Gamma(k^{opt}, \theta^{opt})$

$X_2 \sim U(a_u^{opt}, b_u^{opt})$

$X_3 \sim \mathcal{N}(\mu^{opt}, \sigma^{opt}, a_{NT}^{opt}, b_{NT}^{opt})$

$\sigma^2_n = a_1^{opt} \cdot X_1 + a_2^{opt} \cdot X_2 + a_3^{opt} \cdot X_3$

**Algorithm 1**: Randomized DP with 3 PDFs

**References**

[1] J. Le Ny and M. Mohammady, "Differentially Private MIMO Filtering for Event Streams," in IEEE Transactions on Automatic Control, vol. 63, no. 1, pp. 145-157, Jan. 2018, doi: 10.1109/TAC.2017.2713643.

[2] Meisam Mohammady, Lingyu Wang, Yuan Hong, Habib Louafi, Makan Pourzandi, and Mourad Debbabi. 2018. Preserving Both Privacy and Utility in Network Trace Anonymization. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (CCS ’18). Association for Computing Machinery, New York, NY, USA, 459–474. https://doi.org/10.1145/3243734.3243809.

[3] Meisam Mohammady, Shangyu Xie, Yuan Hong, Mengyuan Zhang, Lingyu Wang, Makan Pourzandi, and Mourad Debbabi. 2020. R2DP: A Universal and Automated Approach to Optimizing the Randomization Mechanisms of Differential Privacy for Utility Metrics with No Known Optimal Distributions. In Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security (CCS ’20). Association for Computing Machinery, New York, NY, USA, 677–696. https://doi.org/10.1145/3372297.3417259