Differential Privacy at Risk: Bridging Randomness and Privacy Budget

Ashish Dandekar, Debabrota Basu, Stéphane Bressan

Ashish Dandekar*, Debabrota Basu*, and Stéphane Bressan

Differential Privacy at Risk: Bridging Randomness and Privacy Budget

Abstract: The calibration of noise for a privacy-preserving mechanism depends on the sensitivity of the query and the prescribed privacy level. A data steward must make the non-trivial choice of a privacy level that balances the requirements of users and the monetary constraints of the business entity.

Firstly, we analyse roles of the sources of randomness, namely the explicit randomness induced by the noise distribution and the implicit randomness induced by the data-generation distribution, that are involved in the design of a privacy-preserving mechanism. The finer analysis enables us to provide stronger privacy guarantees with quantifiable risks. Thus, we propose *privacy at risk* that is a probabilistic calibration of privacy-preserving mechanisms. We provide a composition theorem that leverages privacy at risk. We instantiate the probabilistic calibration for the Laplace mechanism by providing analytical results.

Secondly, we propose a cost model that bridges the gap between the privacy level and the compensation budget estimated by a GDPR compliant business entity. The convexity of the proposed cost model leads to a unique fine-tuning of privacy level that minimises the compensation budget. We show its effectiveness by illustrating a realistic scenario that avoids overestimation of the compensation budget by using privacy at risk for the Laplace mechanism. We quantitatively show that composition using the cost optimal privacy at risk provides stronger privacy guarantee than the classical advanced composition. Although the illustration is specific to the chosen cost model, it naturally extends to any convex cost model. We also provide realistic illustrations of how a data steward uses privacy at risk to balance the trade-off between utility and privacy.

Keywords: Differential privacy, cost model, Laplace mechanism

DOI Editor to enter DOI

Received ..; revised ..; accepted ...

1 Introduction

Dwork et al. [12] quantify the privacy level ε in ε-*differential privacy* (or ε-DP) as an upper bound on the worst-case privacy loss incurred by a privacy-preserving mechanism. Generally, a privacy-preserving mechanism perturbs the results by adding the calibrated amount of random noise to them. The calibration of noise depends on the sensitivity of the query and the specified privacy level. In a real-world setting, a data steward must specify a privacy level that balances the requirements of the users and monetary constraints of the business entity. For example, Garfinkel et al. [14] report on issues encountered when deploying differential privacy as the privacy definition by the US census bureau. They highlight the lack of analytical methods to choose the privacy level. They also report empirical studies that show the loss in utility due to the application of privacy-preserving mechanisms.

We address the dilemma of a data steward in two ways. Firstly, we propose a probabilistic quantification of privacy levels. Probabilistic quantification of privacy levels provides a data steward with a way to take quantified risks under the desired utility of the data. We refer to the probabilistic quantification as *privacy at risk*. We also derive a composition theorem that leverages privacy at risk. Secondly, we propose a cost model that links the privacy level to a monetary budget. This cost model helps the data steward to choose the privacy level constrained on the estimated budget and vice versa. Convexity of the proposed cost model ensures the existence of a unique privacy at risk that would minimise the budget. We show that the composition with an optimal privacy at risk provides stronger privacy guarantees than the traditional advanced composition [12]. In the end, we illustrate a realistic scenario that exemplifies how the

***Corresponding Author: Ashish Dandekar:** DI ENS, ENS, CNRS, PSL University & Inria, Paris, France, E-mail: adandekar@ens.fr

***Corresponding Author: Debabrota Basu:** Dept. of Computer Sci. and Engg., Chalmers University of Technology, Göteborg, Sweden, E-mail: basud@chalmers.se

Stéphane Bressan: National University of Singapore, Singapore, E-mail: steph@nus.edu.sg

data steward can avoid overestimation of the budget by using the proposed cost model by using privacy at risk.

The probabilistic quantification of privacy levels depends on two sources of randomness: the *explicit randomness* induced by the noise distribution and the *implicit randomness* induced by the data-generation distribution. Often, these two sources are coupled with each other. We require analytical forms of both sources of randomness as well as an analytical representation of the query to derive a privacy guarantee. Computing the probabilistic quantification of different sources of randomness is generally a challenging task. Although we find multiple probabilistic privacy definitions in the literature [16, 27] [1], we miss an analytical quantification bridging the randomness and privacy level of a privacy-preserving mechanism. We propose a probabilistic quantification, namely *privacy at risk*, that further leads to analytical relation between privacy and randomness. We derive a composition theorem with privacy at risk for mechanisms with the same as well as varying privacy levels. It is an extension of the advanced composition theorem [12] that deals with a sequential and adaptive use of privacy-preserving mechanisms. We also prove that privacy at risk satisfies convexity over privacy levels and a weak relaxation of the post-processing property. To the best of our knowledge, we are the first to analytically derive the proposed probabilistic quantification for the widely used Laplace mechanism [10].

The privacy level proposed by the differential privacy framework is too abstract a quantity to be integrated in a business setting. We propose a cost model that maps the privacy level to a monetary budget. The proposed model is a convex function of the privacy level, which further leads to a convex cost model for privacy at risk. Hence, it has a unique probabilistic privacy level that minimises the cost. We illustrate this using a realistic scenario in a GDPR-compliant business entity that needs an estimation of the compensation budget that it needs to pay to stakeholders in the unfortunate event of a personal data breach. The illustration, which uses the proposed convex cost model, shows that the use of probabilistic privacy levels avoids overestimation of the compensation budget without sacrificing utility. The illustration naturally extends to any convex cost model.

In this work, we comparatively evaluate the privacy guarantees using privacy at risk of the Laplace mechanism. We quantitatively compare the composition under the optimal privacy at risk, which is estimated using the cost model, with traditional composition mechanisms – basic and advanced mechanisms [12]. We observe that it gives stronger privacy guarantees than the ones obtained by the advanced composition without sacrificing on the utility of the mechanism.

In conclusion, benefits of the probabilistic quantification i.e., of the privacy at risk are twofold. It not only quantifies the privacy level for a given privacy-preserving mechanism but also facilitates decision-making in problems that focus on the privacy-utility trade-off and the compensation budget minimisation.

2 Background

We consider a universe of datasets $\mathcal{D}$. We explicitly mention when we consider that the datasets are sampled from a data-generation distribution $\mathcal{G}$ with support $\mathcal{D}$. Two datasets of equal cardinality x and y are said to be *neighbouring datasets* if they differ in one data point. A pair of neighbouring datasets is denoted by $x \sim y$. In this work, we focus on a specific class of queries called *numeric queries*. A numeric query f is a function that maps a dataset into a real-valued vector, i.e. $f : \mathcal{D} \to \mathbb{R}^k$. For instance, a sum query returns the sum of the values in a dataset.

In order to achieve a privacy guarantee, researchers use a *privacy-preserving mechanism*, or *mechanism* in short, which is a randomised algorithm that adds noise to the query from a given family of distributions. Thus, a privacy-preserving mechanism of a given family, $\mathcal{M}(f, \Theta)$, for the query f and the set of parameters Θ of the given noise distribution, is a function i.e. $\mathcal{M}(f, \Theta) : \mathcal{D} \to \mathcal{R}$. In the case of numerical queries, $\mathcal{R}$ is $\mathbb{R}^k$. We denote a privacy-preserving mechanism as $\mathcal{M}$, when the query and the parameters are clear from the context.

Definition 1 (Differential Privacy [12])**.** *A privacy-preserving mechanism $\mathcal{M}$, equipped with a query f and with parameters Θ, is (ε, δ)-differentially private if for all $Z \subseteq Range(\mathcal{M})$ and $x, y \in \mathcal{D}$ such that $x \sim y$:*

$$\mathbb{P}(\mathcal{M}(f, \Theta)(x) \in Z) \leq e^{\varepsilon} \times \mathbb{P}(\mathcal{M}(f, \Theta)(y) \in Z) + \delta.$$

An $(\varepsilon, 0)$-differentially private mechanism is also simply said to be ε-differentially private. Often, ε-differential privacy is referred to as pure differential privacy whereas (ε, δ)-differential privacy is referred as approximate differential privacy.

1 A widely-used (ε, δ)-differential privacy is not a probabilistic relaxation of differential privacy [29].

A privacy-preserving mechanism provides perfect privacy if it yields indistinguishable outputs for all neighbouring input datasets. The privacy level ε quantifies the privacy guarantee provided by ε-differential privacy. For a given query, the smaller the value of the ε, the qualitatively higher the privacy. A randomised algorithm that is ε-differentially private is also ε'-differential private for any $\varepsilon' > \varepsilon$.

In order to satisfy ε-differential privacy, the parameters of a privacy-preserving mechanism requires a calculated calibration. The amount of noise required to achieve a specified privacy level depends on the query. If the output of the query does not change drastically for two neighbouring datasets, then a small amount of noise is required to achieve a given privacy level. The measure of such fluctuations is called the *sensitivity* of the query. The parameters of a privacy-preserving mechanism are calibrated using the sensitivity of the query that quantifies the smoothness of a numeric query.

Definition 2 (Sensitivity)**.** *The sensitivity of a query $f : \mathcal{D} \to \mathbb{R}^k$ is defined as*

$$\Delta_f \triangleq \max_{\substack{x,y \in \mathcal{D} \\ x \sim y}} \|f(x) - f(y)\|_1.$$

The Laplace mechanism is a privacy-preserving mechanism that adds scaled noise sampled from a calibrated Laplace distribution to the numeric query.

Definition 3 ([35])**.** *The Laplace distribution with mean zero and scale $b > 0$ is a probability distribution with probability density function*

$$\text{Lap}(b) \triangleq \frac{1}{2b} \exp\left(-\frac{|x|}{b}\right),$$

where $x \in \mathbb{R}$. We write Lap(b) *to denote a random variable $X \sim \text{Lap}(b)$*

Definition 4 (Laplace Mechanism [10])**.** *Given any function $f : \mathcal{D} \to \mathbb{R}^k$ and any $x \in \mathcal{D}$, the Laplace Mechanism is defined as*

$$\mathcal{L}_{\varepsilon}^{\Delta_f}(x) \triangleq \mathcal{M}\left(f,\ \frac{\Delta_f}{\varepsilon}\right)(x) = f(x) + (L_1, ..., L_k),$$

where L_i is drawn from Lap $\left(\frac{\Delta_f}{\varepsilon}\right)$ *and added to the i^{th} component of $f(x)$.*

Theorem 1 ([10])**.** *The Laplace mechanism, $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$, is ε_0-differentially private.*

3 Privacy at Risk: A Probabilistic Quantification of Randomness

The parameters of a privacy-preserving mechanism are calibrated using the privacy level and the sensitivity of the query. A data steward needs to choose an appropriate privacy level for practical implementation. Lee et al. [25] show that the choice of an actual privacy level by a data steward in regard to her business requirements is a non-trivial task. Recall that the privacy level in the definition of differential privacy corresponds to the worst case privacy loss. Business users are however used to taking and managing risks, if the risks can be quantified. For instance, Jorion [21] defines *Value at Risk* that is used by risk analysts to quantify the loss in investments for a given portfolio and an acceptable confidence bound. Motivated by the formulation of *Value at Risk*, we propose to use the use of probabilistic privacy level. It provides us with a finer tuning of an ε_0-differentially private privacy-preserving mechanism for a specified risk γ.

Definition 5 (Privacy at Risk)**.** *For a given data generating distribution $\mathcal{G}$, a privacy-preserving mechanism $\mathcal{M}$, equipped with a query f and with parameters Θ, satisfies ε-differential privacy with a* privacy at risk $0 \leq \gamma \leq 1$ *if, for all $Z \subseteq Range(\mathcal{M})$ and x, y sampled from $\mathcal{G}$ such that $x \sim y$:*

$$\mathbb{P}\left[\left|\ln \frac{\mathbb{P}(\mathcal{M}(f,\Theta)(x) \in Z)}{\mathbb{P}(\mathcal{M}(f,\Theta)(y) \in Z)}\right| > \varepsilon\right] \leq \gamma, \tag{1}$$

where the outer probability is calculated with respect to the probability space Range$(\mathcal{M} \circ \mathcal{G})$ *obtained by applying the privacy-preserving mechanism $\mathcal{M}$ on the data-generation distribution $\mathcal{G}$.*

If a privacy-preserving mechanism is ε_0-differentially private for a given query f and parameters Θ, for any privacy level $\varepsilon \geq \varepsilon_0$, the privacy at risk is 0. We are interested in quantifying the risk γ with which an ε_0-differentially private privacy-preserving mechanism also satisfies a stronger ε-differential privacy, i.e., with $\varepsilon < \varepsilon_0$.

Unifying Probabilistic and Random DP

Interestingly, Equation (1) unifies the notions of probabilistic differential privacy and random differential privacy by accounting for both sources of randomness in a privacy-preserving mechanism. Machanavajjhala et

al. [27] define probabilistic differential privacy that incorporates the explicit randomness of the noise distribution of the privacy-preserving mechanism, whereas Hall et al. [16] define random differential privacy that incorporates the implicit randomness of the data-generation distribution. In probabilistic differential privacy, the outer probability is computed over the sample space of Range($\mathcal{M}$) and all datasets are equally probable.

Connection with Approximate DP

Despite a resemblance with probabilistic relaxations of differential privacy [13, 16, 27] due to the added parameter δ, (ε, δ)-differential privacy (Definition 1) is a non-probabilistic variant [29] of regular ε-differential privacy. Indeed, unlike the auxiliary parameters in probabilistic relaxations, such as γ in privacy at risk (ref. Definition 5), the parameter δ of approximate differential privacy is an absolute slack that is independent of the sources of randomness. For a specified choice of ε and δ, one can analytically compute a matching value of δ for a new value of ε[2]. Therefore, as other probabilistic relaxations, privacy at risk cannot be directly related to approximate differential privacy. An alternative is to find out a privacy at risk level γ for a given privacy level (ε, δ) while the original noise satisfies (ε_0, δ).

Theorem 2. *If a privacy preserving mechanism satisfies (ε, γ) privacy at risk, it also satisfies (ε, γ) approximate differential privacy.*

We obtain this reduction as the probability measure induced by the privacy preserving mechanism and data generating distribution on any output set $Z \subseteq Range(\mathcal{M})$ is additive. [3] The proof of the theorem is in Appendix A.

3.1 Composition theorem

The application of ε-differential privacy to many real-world problem suffers from the degradation of privacy guarantee, i.e., privacy level, over the composition. The basic composition theorem [12] dictates that the privacy guarantee degrades linearly in the number of evaluations of the mechanism. The advanced composition theorem [12] provides a finer analysis of the privacy loss over multiple evaluations with a square root dependence on the number of evaluations. In this section, we provide the composition theorem for privacy at risk.

Definition 6 (Privacy loss random variable)**.** *For a privacy-preserving mechanism $\mathcal{M} : \mathcal{D} \to R$, any two neighbouring datasets $x, y \in \mathcal{D}$ and an output $r \in R$, the value of the privacy loss random variable $\mathcal{C}$ is defined as:*

$$\mathcal{C}(r) \triangleq \ln \frac{\mathbb{P}(\mathcal{M}(x) = r)}{\mathbb{P}(\mathcal{M}(y) = r)}.$$

Lemma 1. *If a privacy-preserving mechanism $\mathcal{M}$ satisfies ε_0-differential privacy, then*

$$\mathbb{P}[|\mathcal{C}| \leq \varepsilon_0] = 1.$$

Theorem 3. *For all $\varepsilon_0, \varepsilon, \gamma, \delta > 0$, the class of ε_0-differentially private mechanisms, which satisfy (ε, γ)-privacy at risk under a uniform data-generation distribution, are (ε', δ)-differential privacy under n-fold composition where*

$$\varepsilon' = \varepsilon_0 \sqrt{2n \ln \frac{1}{\delta}} + n\mu,$$

where $\mu = \frac{1}{2}[\gamma\varepsilon^2 + (1-\gamma)\varepsilon_0^2]$.

Proof. Let, $\mathcal{M}^{1\ldots n} : \mathcal{D} \to R^1 \times R^2 \times \ldots \times R^n$ denote the n-fold composition of privacy-preserving mechanisms $\{\mathcal{M}^i : \mathcal{D} \to R^i\}_{i=1}^n$. Each ε_0-differentially private $\mathcal{M}^i$ also satisfies (ε, γ)-privacy at risk for some $\varepsilon \leq \varepsilon_0$ and appropriately computed γ. Consider any two neighbouring datasets $x, y \in \mathcal{D}$. Let,

$$B = \left\{ (r_1, \ldots, r_n) \middle| \bigwedge_{i=1}^{n} \frac{\mathbb{P}(\mathcal{M}^i(x) = r_i)}{\mathbb{P}(\mathcal{M}^i(y) = r_i)} > \mathrm{e}^{\varepsilon} \right\}$$

Using the technique in [12, Theorem 3.20], it suffices to show that $\mathbb{P}(\mathcal{M}^{1\ldots n}(x) \in B) \leq \delta$.

Consider

$$\begin{aligned} &\ln \frac{\mathbb{P}(\mathcal{M}^{1\ldots n}(x) = (r_1, \ldots, r_n))}{\mathbb{P}(\mathcal{M}^{1\ldots n}(y) = (r_1, \ldots, r_n))} \\ &= \ln \prod_{i=1}^{n} \frac{\mathbb{P}(\mathcal{M}^i(x) = r_i)}{\mathbb{P}(\mathcal{M}^i(y) = r_i)} \\ &= \sum_{i=1}^{n} \ln \frac{\mathbb{P}(\mathcal{M}^i(x) = r_i)}{\mathbb{P}(\mathcal{M}^i(y) = r_i)} \triangleq \sum_{i=1}^{n} \mathcal{C}^i \end{aligned} \tag{2}$$

where $\mathcal{C}^i$ in the last line denotes the privacy loss random variable related to $\mathcal{M}^i$.

Consider an ε-differentially private mechanism $\mathcal{M}_\varepsilon$ and ε_0-differentially private mechanism $\mathcal{M}_{\varepsilon_0}$. Let $\mathcal{M}_{\varepsilon_0}$ satisfy (ε, γ)-privacy at risk for $\varepsilon \leq \varepsilon_0$ and appropriately

2 For any $0 < \varepsilon' \leq \varepsilon$, any (ε, δ)-differentially private mechanism also satisfies $(\varepsilon', (\mathrm{e}^{\varepsilon} - \mathrm{e}^{\varepsilon'} + \delta))$-differential privacy.

3 The converse is not true as explained before.

computed γ. Each $\mathcal{M}^i$ can be simulated as the mechanism $\mathcal{M}_\varepsilon$ with probability γ and the mechanism $\mathcal{M}_{\varepsilon_0}$ otherwise. Therefore, the privacy loss random variable for each mechanism $\mathcal{M}^i$ can be written as

$$\mathcal{C}^i = \gamma\mathcal{C}^i_\varepsilon + (1-\gamma)\mathcal{C}^i_{\varepsilon_0}$$

where $\mathcal{C}^i_\varepsilon$ denotes the privacy loss random variable associated with the mechanism $\mathcal{M}_\varepsilon$ and $\mathcal{C}^i_{\varepsilon_0}$ denotes the privacy loss random variable associated with the mechanism $\mathcal{M}_{\varepsilon_0}$. Using [5, Remark 3.4], we can bound the mean of every privacy loss random variable as:

$$\mu \triangleq \mathbb{E}[\mathcal{C}^i] \leq \frac{1}{2}[\gamma\varepsilon^2 + (1-\gamma)\varepsilon_0^2].$$

We have a collection of n independent privacy random variables $\mathcal{C}^i$'s such that $\mathbb{P}\left[|\mathcal{C}^i| \leq \varepsilon_0\right] = 1$. Using Hoeffding's bound [18] on the sample mean for any $\beta > 0$,

$$\mathbb{P}\left[\frac{1}{n}\sum_i C^i \geq \mathbb{E}[\mathcal{C}^i] + \beta\right] \leq \exp\left(-\frac{n\beta^2}{2\varepsilon_0^2}\right).$$

Rearranging the inequality by renaming the upper bound on the probability as δ, we get:

$$\mathbb{P}\left[\sum_i C^i \geq n\mu + \varepsilon_0\sqrt{2n\ln\frac{1}{\delta}}\right] \leq \delta.$$

□

Theorem 3 is an analogue, in the privacy at risk setting, of the advanced composition of differential privacy [12, Theorem 3.20] under a constraint of independent evaluations. Note that if one takes $\gamma = 0$, then we obtain the exact same formula as in [12, Theorem 3.20]. It provides a sanity check for the consistency of composition using privacy at risk.

Corollary 1 (Heterogeneous Composition)**.** *For all $\varepsilon_l, \varepsilon, \gamma_l, \delta > 0$ and $l \in \{1, \ldots, n\}$, the composition of $\{\varepsilon_l\}_{l=1}^n$-differentially private mechanisms, which satisfy (ε, γ_l)-privacy at risk under a uniform data-generation distribution, also satisfies (ε', δ)-differential privacy where*

$$\varepsilon' = \sqrt{2\left(\sum_{l=1}^n \varepsilon_l^2\right)\ln\frac{1}{\delta}} + \mu,$$

where $\mu = \frac{1}{2}[\varepsilon^2(\sum_{l=1}^n \gamma_l) + \sum_{l=1}^n (1-\gamma_l)\varepsilon_l^2]$.

Proof. The proof follows from the same argument as that of Theorem 3 of bounding the loss random variable at step l using $\gamma_l\mathcal{C}^l_\varepsilon + (1-\gamma_l)\mathcal{C}^l_{\varepsilon_l}$ and then applying the concentration inequality. □

A detailed discussion and analysis of proving such heterogeneous composition theorems is available in [22, Section 3.3].

In fact, if we consider both sources of randomness, the expected value of the loss function must be computed by using the law of total expectation.

$$\mathbb{E}[\mathcal{C}] = \mathbb{E}_{x,y\sim\mathcal{G}}[\mathbb{E}[\mathcal{C}]|x, y]$$

Therefore, the exact computation of privacy guarantees after the composition requires access to the data-generation distribution. We assume a uniform data-generation distribution while proving Theorem 3. We can obtain better and finer privacy guarantees accounting for data-generation distribution, which we keep as a future work.

3.2 Convexity and Post-Processing

We show that privacy at risk satisfies the convexity property and does not satisfy the post-processing property.

Lemma 2 (Convexity)**.** *For a given ε_0-differentially private privacy-preserving mechanism, privacy at risk satisfies the convexity property.*

Proof. Let $\mathcal{M}$ be a mechanism that satisfies ε_0-differential privacy. By the definition of the privacy at risk, it also satisfies $(\varepsilon_1, \gamma_1)$-privacy at risk as well as $(\varepsilon_2, \gamma_2)$-privacy at risk for some $\varepsilon_1, \varepsilon_2 \leq \varepsilon_0$ and appropriately computed values of γ_1 and γ_2. Let $\mathcal{M}^1$ and $\mathcal{M}^2$ denote the hypothetical mechanisms that satisfy $(\varepsilon_1, \gamma_1)$-privacy at risk and $(\varepsilon_2, \gamma_2)$-privacy at risk respectively. We can write privacy loss random variables as follows:

$$\mathcal{C}^1 \leq \gamma_1\varepsilon_1 + (1-\gamma_1)\varepsilon_0$$
$$\mathcal{C}^2 \leq \gamma_2\varepsilon_2 + (1-\gamma_2)\varepsilon_0$$

where $\mathcal{C}^1$ and $\mathcal{C}^2$ denote privacy loss random variables for $\mathcal{M}^1$ and $\mathcal{M}^2$.

Let us consider a privacy-preserving mechanism $\mathcal{M}$ that uses $\mathcal{M}^1$ with a probability p and $\mathcal{M}^2$ with a probability $(1-p)$ for some $p \in [0, 1]$. By using the techniques in the proof of Theorem 3, the privacy loss random variable $\mathcal{C}$ for $\mathcal{M}$ can be written as:

$$\begin{aligned}\mathcal{C} &= p\mathcal{C}^1 + (1-p)\mathcal{C}^2 \\ &\leq \gamma'\varepsilon' + (1-\gamma')\varepsilon_0\end{aligned}$$

where

$$\varepsilon' = \frac{p\gamma_1\varepsilon_1 + (1-p)\gamma_2\varepsilon_2}{p\gamma_1 + (1-p)\gamma_2}$$

$$\gamma' = (1 - p\gamma_1 - (1-p)\gamma_2)$$

Thus, $\mathcal{M}$ satisfies (ε', γ')-privacy at risk. This proves that privacy at risk satisfies convexity [23, Axiom 2.1.2]. □

Meiser [29] proved that a relaxation of differential privacy that provides probabilistic bounds on the privacy loss random variable does not satisfy post-processing property of differential privacy. Privacy at risk is indeed such a probabilistic relaxation.

Corollary 2 (Post-processing)**.** *Privacy at risk does not satisfy the post-processing property for every possible mapping of the output.*

Though privacy at risk is not preserved after post-processing, it yields a weaker guarantee in terms of approximate differential privacy after post-processing. The proof involves reduction of privacy at risk to approximate differential privacy and preservation of approximate differential privacy under post-processing.

Lemma 3 (Weak Post-processing)**.** *Let $\mathcal{M} : \mathcal{D} \to \mathcal{R} \subseteq \mathbb{R}^k$ be a mechanism that satisfy (ε, γ)-privacy at risk and $f : \mathcal{R} \to \mathcal{R}'$ be any arbitrary data independent mapping. Then, $f \circ \mathcal{M} : \mathcal{D} \to \mathcal{R}'$ would also satisfy (ε, γ)-approximate differential privacy.*

Proof. Let us fix a pair of neighbouring datasets x and y, and also an event $Z' \subseteq \mathcal{R}'$. Let us define pre-image of Z' as $Z \triangleq \{r \in \mathcal{R} : f(r) \in Z\}$. Now, we get

$$\begin{aligned}\mathbb{P}(f \circ \mathcal{M}(x) \in Z') &= \mathbb{P}(\mathcal{M}(x) \in Z) \\ &\underset{(a)}{\leq} e^{\varepsilon}\mathbb{P}(\mathcal{M}(y) \in Z) + \gamma \\ &= e^{\varepsilon}\mathbb{P}(f \circ \mathcal{M}(y) \in Z') + \delta\end{aligned}$$

(a) is a direct consequence of Theorem 2. □

4 Privacy at Risk for Laplace Mechanism

The Laplace and Gaussian mechanisms are widely used privacy-preserving mechanisms in the literature. The Laplace mechanism satisfies pure ε-differential privacy whereas the Gaussian mechanism satisfies approximate (ε, δ)-differential privacy. As previously discussed, it is not straightforward to establish a connection between the non-probabilistic parameter δ of approximate differential privacy and the probabilistic bound γ of privacy at risk. Therefore, we keep privacy at risk for Gaussian mechanism as the future work.

In this section, we instantiate privacy at risk for the Laplace mechanism in three cases: two cases involving two sources of randomness and a third case involving the coupled effect. These three different cases correspond to three different interpretations of the confidence level, represented by the parameter γ, corresponding to three interpretations of the support of the outer probability in Definition 5. In order to highlight this nuance, we denote the confidence levels corresponding to the three cases and their three sources of randomness as γ_1, γ_2, and γ_3, respectively.

4.1 The Case of Explicit Randomness

In this section, we study the effect of the explicit randomness induced by the noise sampled from Laplace distribution. We provide a probabilistic quantification for fine tuning for the Laplace mechanism. We fine-tune the privacy level for a specified risk under by assuming that the sensitivity of the query is known a priori.

For a Laplace mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$ calibrated with sensitivity Δ_f and privacy level ε_0, we present the analytical formula relating privacy level ε and the risk γ_1 in Theorem 4. The proof is available in Appendix B.

Theorem 4. *The risk $\gamma_1 \in [0, 1]$ with which a Laplace Mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$, for a numeric query $f : \mathcal{D} \to \mathbb{R}^k$ satisfies a privacy level $\varepsilon \geq 0$ is given by*

$$\gamma_1 = \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \varepsilon_0)}, \tag{3}$$

where T is a random variable that follows a distribution with the following density function.

$$P_T(t) = \frac{2^{1-k} t^{k-\frac{1}{2}} K_{k-\frac{1}{2}}(t)\varepsilon_0}{\sqrt{2\pi}\Gamma(k)\Delta_f}$$

where $K_{n-\frac{1}{2}}$ is the Bessel function of second kind.

Figure 1a shows the plot of the privacy level against risk for different values of k and for a Laplace mechanism $\mathcal{L}_{1.0}^{1.0}$. As the value of k increases, the amount of noise added in the output of numeric query increases. Therefore, for a specified privacy level, the privacy at risk level increases with the value of k.

The analytical formula representing γ_1 as a function of ε is bijective. We need to invert it to obtain the privacy level ε for a privacy at risk γ_1. However the analytical closed form for such an inverse function is not

explicit. We use a numerical approach to compute privacy level for a given privacy at risk from the analytical formula of Theorem 4.

Result for a Real-valued Query. For the case $k = 1$, the analytical derivation is fairly straightforward. In this case, we obtain an invertible closed-form of a privacy level for a specified risk. It is presented in Equation 4.

$$\varepsilon = \ln\left(\frac{1}{1-\gamma_1(1-e^{-\varepsilon_0})}\right) \tag{4}$$

Remarks on ε_0. For $k = 1$, Figure 1b shows the plot of privacy at risk level ε versus privacy at risk γ_1 for the Laplace mechanism $\mathcal{L}^{1.0}_{\varepsilon_0}$. As the value of ε_0 increases, the probability of Laplace mechanism generating higher value of noise reduces. Therefore, for a fixed privacy level, privacy at risk increases with the value of ε_0. The same observation is made for $k > 1$.

4.2 The Case of Implicit Randomness

In this section, we study the effect of the implicit randomness induced by the data-generation distribution to provide a fine tuning for the Laplace mechanism. We fine-tune the risk for a specified privacy level without assuming that the sensitivity of the query.

If one takes into account randomness induced by the data-generation distribution, all pairs of neighbouring datasets are not equally probable. This leads to estimation of sensitivity of a query for a specified data-generation distribution. If we have access to an analytical form of the data-generation distribution and to the query, we could analytically derive the sensitivity distribution for the query. In general, we have access to the datasets, but not the data-generation distribution that generates them. We, therefore, statistically estimate sensitivity by constructing an empirical distribution. We call the sensitivity value obtained for a specified risk from the empirical cumulative distribution of sensitivity the *sampled sensitivity* (Definition 7). However, the value of sampled sensitivity is simply an estimate of the sensitivity for a specified risk. In order to capture this additional uncertainty introduced by the estimation from the empirical sensitivity distribution rather than the true unknown distribution, we compute a lower bound on the accuracy of this estimation. This lower bound yields a probabilistic lower bound on the specified risk. We refer to it as *empirical risk*. For a specified absolute risk γ_2, we denote by $\hat{\gamma_2}$ corresponding empirical risk.

For the Laplace mechanism $\mathcal{L}_{\varepsilon}^{\Delta_{S_f}}$ calibrated with sampled sensitivity Δ_{S_f} and privacy level ε, we evaluate the empirical risk $\hat{\gamma_2}$. We present the result in Theorem 5. The proof is available in Appendix C.

Theorem 5. *Analytical bound on the empirical risk, $\hat{\gamma_2}$, for Laplace mechanism $\mathcal{L}_{\varepsilon}^{\Delta_{S_f}}$ with privacy level ε and sampled sensitivity Δ_{S_f} for a query $f : \mathcal{D} \to \mathbb{R}^k$ is*

$$\hat{\gamma_2} \geq \gamma_2(1 - 2e^{-2\rho^2 n}) \tag{5}$$

where n is the number of samples used for estimation of the sampled sensitivity and ρ is the accuracy parameter. γ_2 denotes the specified absolute risk.

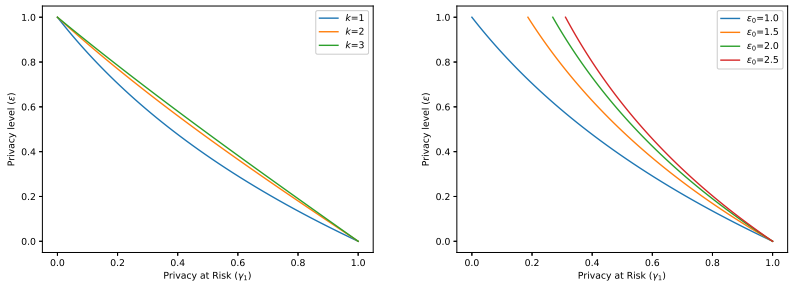
The error parameter ρ controls the closeness between the empirical cumulative distribution of the sensitivity to the true cumulative distribution of the sensitivity. Lower the value of the error, closer is the empirical cumulative distribution to the true cumulative distribution. Mathematically,

$$\rho \geq \sup_{\Delta} |F^n_S(\Delta) - F_S(\Delta)|,$$

where F^n_S is the empirical cumulative distribution of sensitivity after n samples and F_S is the actual cumulative distribution of sensitivity.

Figure 2 shows the plot of number of samples as a function of the privacy at risk and the error parameter. Naturally, we require higher number of samples in order to have lower error rate. The number of samples reduces as the privacy at risk increases. The lower risk demands precision in the estimated sampled sensitivity, which in turn requires larger number of samples.

If the analytical form of the data-generation distribution is not known a priori, the empirical distribution of sensitivity can be estimated in two ways. The first way is to fit a known distribution on the available data and later use it to build an empirical distribution of the sensitivities. The second way is to sub-sample from a large dataset in order to build an empirical distribution of the sensitivities. In both of these ways, the empirical distribution of sensitivities captures the inherent randomness in the data-generation distribution. The first way suffers from the goodness of the fit of the known distribution to the available data. An ill-fit distribution does not reflect the true data-generation distribution and hence introduces errors in the sensitivity estimation. Since the second way involves subsampling, it is immune to this problem. The quality of sensitivity estimates obtained by sub-sampling the datasets depend on the availability of large population.



(a) (b)

Fig. 1. Privacy level ε for varying privacy at risk γ_1 for Laplace mechanism $\mathcal{L}^{1.0}_{\varepsilon_0}$. In Figure 1a, we use $\varepsilon_0 = 1.0$ and different values of k. In Figure 1b, for $k = 1$ and different values of ε_0.

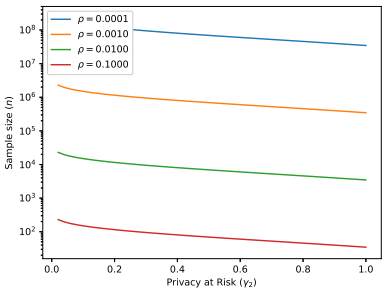


Fig. 2. Number of samples n for varying privacy at risk γ_2 for different error parameter ρ.

Let, $\mathcal{G}$ denotes the data-generation distribution, either known apriori or constructed by subsampling the available data. We adopt the procedure of [38] to sample two neighbouring datasets with p data points each. We sample $p-1$ data points from $\mathcal{G}$ that are common to both of these datasets and later two more data points, *independently*. From those two points, we allot one data point to each of the two datasets.

Let, $S_f = \|f(x) - f(y)\|_1$ denotes the sensitivity random variable for a given query f, where x and y are two neighbouring datasets sampled from $\mathcal{G}$. Using n pairs of neighbouring datasets sampled from $\mathcal{G}$, we construct the empirical cumulative distribution, F_n, for the sensitivity random variable.

Definition 7. *For a given query f and for a specified risk γ_2, sampled sensitivity, Δ_{S_f}, is defined as the value of sensitivity random variable that is estimated using its empirical cumulative distribution function, F_n, constructed using n pairs of neighbouring datasets sampled from the data-generation distribution $\mathcal{G}$.*

$$\Delta_{S_f} \triangleq F_n^{-1}(\gamma_2)$$

If we knew analytical form of the data generation distribution, we could analytically derive the cumulative distribution function of the sensitivity, F, and find the sensitivity of the query as $\Delta_f = F^{-1}(1)$. Therefore, in order to have the sampled sensitivity close to the sensitivity of the query, we require the empirical cumulative distributions to be close to the cumulative distribution of the sensitivity. We use this insight to derive the analytical bound in the Theorem 5.

4.3 The Case of Explicit and Implicit Randomness

In this section, we study the combined effect of both explicit randomness induced by the noise distribution and implicit randomness in the data-generation distribution respectively. We do not assume the knowledge of the sensitivity of the query.

We estimate sensitivity using the empirical cumulative distribution of sensitivity. We construct the empirical distribution over the sensitivities using the sampling technique presented in the earlier case. Since we use the sampled sensitivity (Definition 7) to calibrate the Laplace mechanism, we estimate the *empirical risk* $\hat{\gamma_3}$.

For Laplace mechanism $\mathcal{L}^{\Delta_{S_f}}_{\varepsilon_0}$ calibrated with sampled sensitivity Δ_{S_f} and privacy level ε_0, we present the analytical bound on the empirical sensitivity $\hat{\gamma_3}$ in Theorem 6 with proof in the Appendix D.

Theorem 6. *Analytical bound on the empirical risk $\hat{\gamma_3} \in [0, 1]$ to achieve a privacy level $\varepsilon > 0$ for Laplace mechanism $\mathcal{L}^{\Delta_{S_f}}_{\varepsilon_0}$ with sampled sensitivity Δ_{S_f} of a query $f : \mathcal{D} \to \mathbb{R}^k$ is*

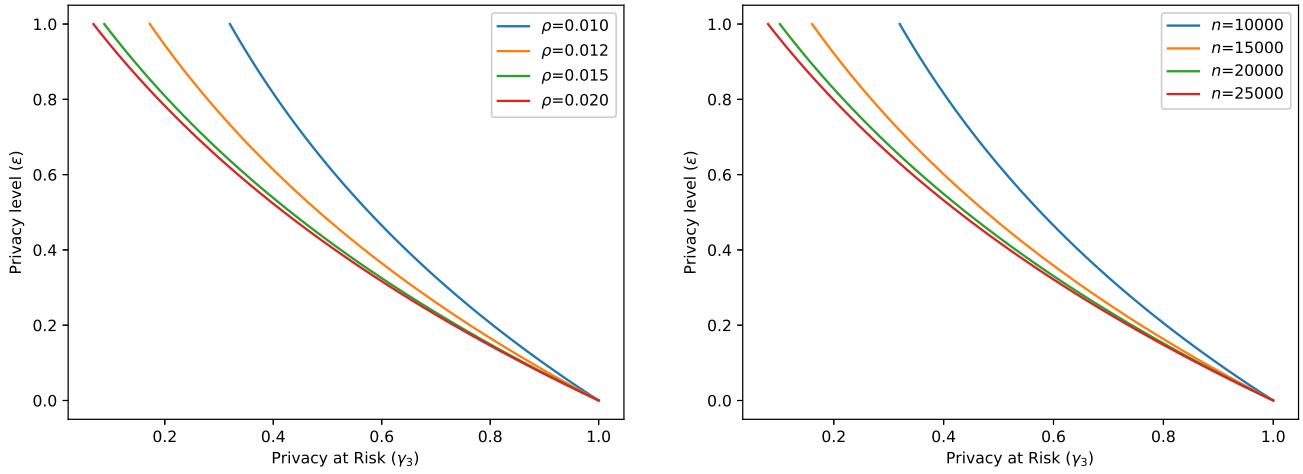
$$\hat{\gamma_3} \geq \gamma_3(1 - 2e^{-2\rho^2 n}) \tag{6}$$

where n is the number of samples used for estimating the sensitivity, ρ is the accuracy parameter. γ_3 denotes the specified absolute risk defined as:

$$\gamma_3 = \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)} \cdot \gamma_2$$

Here, η is of the order of the ratio of the true sensitivity of the query to its sampled sensitivity.

The error parameter ρ controls the closeness between the empirical cumulative distribution of the sensitivity to the true cumulative distribution of the sensitivity. Figure 3 shows the dependence of the error parameter



(a) (b)

Fig. 3. Dependence of error and number of samples on the privacy at risk for Laplace mechanism $\mathcal{L}_{1.0}^{\Delta_{S_f}}$. For the figure on the left hand side, we fix the number of samples to 10000. For the Figure 3b we fix the error parameter to 0.01.

on the number of samples. In Figure 3a, we observe that for a fixed number of samples and a privacy level, the privacy at risk decreases with the value of error parameter. For a fixed number of samples, smaller values of the error parameter reduce the probability of similarity between the empirical cumulative distribution of sensitivity and the true cumulative distribution. Therefore, we observe the reduction in the risk for a fixed privacy level. In Figure 3b, we observe that for a fixed value of error parameter and a fixed level of privacy level, the risk increases with the number of samples. For a fixed value of the error parameter, larger values of the sample size increase the probability of similarity between the empirical cumulative distribution of sensitivity and the true cumulative distribution. Therefore, we observe the increase in the risk for a fixed privacy level.

Effect of the consideration of implicit and explicit randomness is evident in the analytical expression for γ_3 in Equation 7. Proof is available in Appendix D. The privacy at risk is composed of two factors whereas the second term is a privacy at risk that accounts for inherent randomness. The first term takes into account the implicit randomness of the Laplace distribution along with a coupling coefficient η. We define η as the ratio of the true sensitivity of the query to its sampled sensitivity. We provide an approximation to estimate η in the absence of knowledge of the true sensitivity. It can be found in Appendix D.

$$\gamma_3 \triangleq \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)} \cdot \gamma_2 \tag{7}$$

5 Minimising Compensation Budget for Privacy at Risk

Many service providers collect users' data to enhance user experience. In order to avoid misuse of this data, we require a legal framework that not only limits the use of the collected data but also proposes reparative measures in case of a data leak. General Data Protection Regulation (GDPR)[4] is such a legal framework.

Section 82 in GDPR states that any person who suffers from material or non-material damage as a result of a personal data breach has the right to demand compensation from the data processor. Therefore, every GDPR compliant business entity that either holds or processes personal data needs to secure a certain budget in the scenario of the personal data breach. In order to reduce the risk of such an unfortunate event, the business entity may use privacy-preserving mechanisms that provide provable privacy guarantees while publishing their results. In order to calculate the compensation budget for a business entity, we devise a cost model that maps the privacy guarantees provided by differential privacy and privacy at risk to monetary costs. The discussions demonstrate the usefulness of probabilistic quantification of differential privacy in a business setting.

4 https://gdpr-info.eu/

5.1 Cost Model for Differential Privacy

Let E be the compensation budget that a business entity has to pay to every stakeholder in case of a personal data breach when the data is processed without any provable privacy guarantees. Let E^{dp}_{ε} be the compensation budget that a business entity has to pay to every stakeholder in case of a personal data breach when the data is processed with privacy guarantees in terms of ε-differential privacy.

Privacy level, ε, in ε-differential privacy is the quantifier of indistinguishability of the outputs of a privacy-preserving mechanism when two neighbouring datasets are provided as inputs. When the privacy level is zero, the privacy-preserving mechanism outputs all results with equal probability. The indistinguishability reduces with increase in the privacy level. Thus, privacy level of zero bears the lowest risk of personal data breach and the risk increases with the privacy level. E^{dp}_{ε} needs to be commensurate to such a risk and, therefore, it needs to satisfy the following constraints.

1. For all $\varepsilon \in \mathbb{R}^{\geq 0}$, $E^{dp}_{\varepsilon} \leq E$.
2. E^{dp}_{ε} is a monotonically increasing function of ε.
3. As $\varepsilon \to 0$, $E^{dp}_{\varepsilon} \to E_{min}$ where E_{min} is the unavoidable cost that business entity might need to pay in case of personal data breach even after the privacy measures are employed.
4. As $\varepsilon \to \infty$, $E^{dp}_{\varepsilon} \to E$.

There are various functions that satisfy these constraints. In absence of any further constraints, we model E^{dp}_{ε} as defined in Equation (8).

$$E^{dp}_{\varepsilon} \triangleq E_{min} + Ee^{-\frac{c}{\varepsilon}}. \tag{8}$$

E^{dp}_{ε} has two parameters, namely $c > 0$ and $E_{min} \geq 0$. c controls the rate of change in the cost as the privacy level changes and E_{min} is a privacy level independent bias. For this study, we use a simplified model with $c = 1$ and $E_{min} = 0$.

5.2 Cost Model for Privacy at Risk

Let, $E^{par}_{\varepsilon_0}(\varepsilon, \gamma)$ be the compensation that a business entity has to pay to every stakeholder in case of a personal data breach when the data is processed with an ε_0-differentially private privacy-preserving mechanism along with a probabilistic quantification of privacy level. Use of such a quantification allows us to provide a stronger privacy guarantee *viz.* $\varepsilon < \varepsilon_0$ for a specified privacy at risk at most γ. Thus, we calculate $E^{par}_{\varepsilon_0}$ using Equation 9.

$$E^{par}_{\varepsilon_0}(\varepsilon, \gamma) \triangleq \gamma E^{dp}_{\varepsilon} + (1-\gamma)E^{dp}_{\varepsilon_0} \tag{9}$$

Note that the analysis in this section is specific to the cost model in Equation 8. It naturally extends to any choice of convex cost model.

5.2.1 Existence of Minimum Compensation Budget

We want to find the privacy level, say ε_{min}, that yields the lowest compensation budget. We do that by minimising Equation 9 with respect to ε.

Lemma 4. *For the choice of cost model in Equation 8, $E^{par}_{\varepsilon_0}(\varepsilon, \gamma)$ is a convex function of ε.*

By Lemma 4, there exists a unique ε_{min} that minimises the compensation budget for a specified parametrisation, say ε_0. Since the risk γ in Equation 9 is itself a function of privacy level ε, analytical calculation of ε_{min} is not possible in the most general case. When the output of the query is a real number, i. e. $k = 1$, we derive the analytic form (Equation 4) to compute the risk under the consideration of explicit randomness. In such a case, ε_{min} is calculated by differentiating Equation 9 with respect to ε and equating it to zero. It gives us Equation 10 that we solve using any root finding technique such as Newton-Raphson method [37] to compute ε_{min}.

$$\frac{1}{\varepsilon} - \ln\left(1 - \frac{1-e^{\varepsilon}}{\varepsilon^2}\right) = \frac{1}{\varepsilon_0} \tag{10}$$

5.2.2 Fine-tuning Privacy at Risk

For a fixed budget, say B, re-arrangement of Equation 9 gives us an upper bound on the privacy level ε. We use the cost model with $c = 1$ and $E_{min} = 0$ to derive the upper bound. If we have a maximum permissible expected mean absolute error T, we use Equation 12 to obtain a lower bound on the privacy at risk level. Equation 11 illustrates the upper and lower bounds that dictate the permissible range of ε that a data publisher can promise depending on the budget and the permissible error constraints.

$$\frac{1}{T} \leq \varepsilon \leq \left[\ln\left(\frac{\gamma E}{B - (1-\gamma)E^{dp}_{\varepsilon_0}}\right)\right]^{-1} \tag{11}$$

Thus, the privacy level is constrained by the effectiveness requirement from below and by the mone-

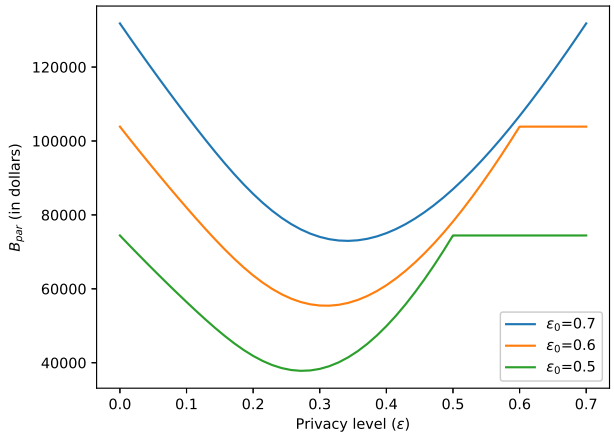


Fig. 4. Variation in the budget for Laplace mechanism $\mathcal{L}^1_{\varepsilon_0}$ under privacy at risk considering explicit randomness in the Laplace mechanism for the illustration in Section 5.3.

tary budget from above. [19] calculate upper and lower bound on the privacy level in the differential privacy. They use a different cost model owing to the scenario of research study that compensates its participants for their data and releases the results in a differentially private manner. Their cost model is different than our GDPR inspired modelling.

5.3 Illustration

Suppose that the health centre in a university that complies to GDPR publishes statistics of its staff health checkup, such as obesity statistics, twice in a year. In January 2018, the health centre publishes that 34 out of 99 faculty members suffer from obesity. In July 2018, the health centre publishes that 35 out of 100 faculty members suffer from obesity. An intruder, perhaps an analyst working for an insurance company, checks the staff listings in January 2018 and July 2018, which are publicly available on website of the university. The intruder does not find any change other than the recruitment of John Doe in April 2018. Thus, with high probability, the intruder deduces that John Doe suffers from obesity. In order to avoid such a privacy breach, the health centre decides to publish the results using the Laplace mechanism. In this case, the Laplace mechanism operates on the count query.

In order to control the amount of noise, the health centre needs to appropriately set the privacy level. Suppose that the health centre decides to use the expected mean absolute error, defined in Equation 12, as the measure of *effectiveness* for the Laplace mechanism.

$$\mathbb{E}\left[|\mathcal{L}^1_\varepsilon(x) - f(x)|\right] = \frac{1}{\varepsilon} \tag{12}$$

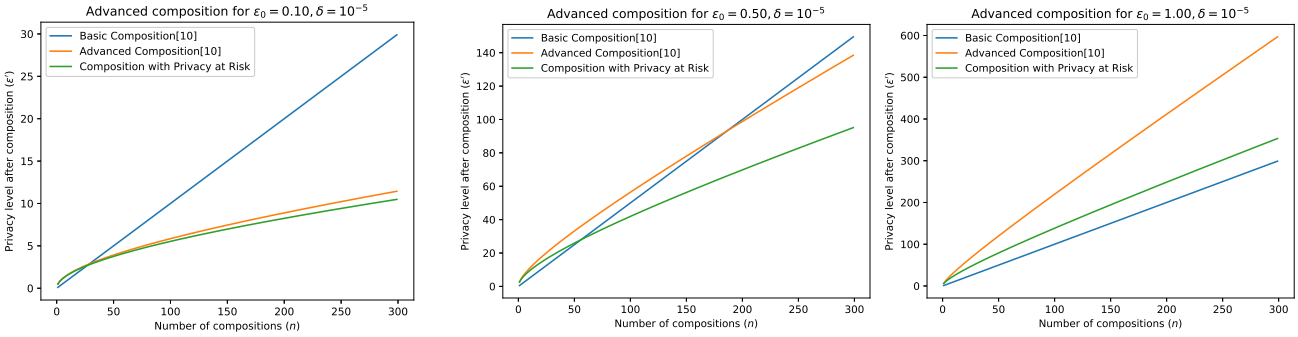
Equation 12 makes use of the fact that the sensitivity of the count query is one. Suppose that the health centre requires the expected mean absolute error of at most two in order to maintain the quality of the published statistics. In this case, the privacy level has to be at least 0.5.

In order to compute the budget, the health centre requires an estimate of E. Moriarty et al. [30] show that the incremental cost of premiums for the health insurance with morbid obesity ranges between \$5467 to \$5530. With reference to this research, the health centre takes \$5500 as an estimate of E. For the staff size of 100 and the privacy level 0.5, the health centre uses Equation 8 in its simplified setting to compute the total budget of \$74434.40.

Is it possible to reduce this budget without degrading the effectiveness of the Laplace mechanism? We show that it is possible by fine-tuning the Laplace mechanism. Under the consideration of the explicit randomness introduced by the Laplace noise distribution, we show that ε_0-differentially private Laplace mechanism also satisfies ε-differential privacy with risk γ, which is computed using the formula in Theorem 4. Fine-tuning allows us to get a stronger privacy guarantee, $\varepsilon < \varepsilon_0$ that requires a smaller budget. In Figure 4, we plot the budget for various privacy levels. We observe that the privacy level 0.274, which is same as ε_{min} computed by solving Equation 10, yields the lowest compensation budget of \$37805.86. Thus, by using privacy at risk, the health centre is able to save \$36628.532 without sacrificing the quality of the published results.

5.4 Cost Model and the Composition of Laplace Mechanisms

Convexity of the proposed cost function enables us to estimate the optimal value of the privacy at risk level. We use the optimal privacy value to provide tighter bounds on the composition of Laplace mechanism. In Figure 5, we compare the privacy guarantees obtained by using basic composition theorem [12], advanced composition theorem [12] and the composition theorem for privacy at risk. We comparatively evaluate them for composition of Laplace mechanisms with privacy levels $0.1, 0.5$ and 1.0. We compute the privacy level after composition by setting δ to 10^{-5}.



(a) $\mathcal{L}^1_{0.1}$ satisfies $(0.08, 0.80)$-privacy at risk. **(b)** $\mathcal{L}^1_{0.5}$ satisfies $(0.27, 0.61)$-privacy at risk. **(c)** $\mathcal{L}^1_{1.0}$ satisfies $(0.42, 0.54)$-privacy at risk.

Fig. 5. Comparing the privacy guarantee obtained by basic composition and advanced composition [12] with the composition obtained using optimal privacy at risk that minimises the cost of Laplace mechanism $\mathcal{L}^1_{\varepsilon_0}$. For the evaluation, we set $\delta = 10^{-5}$.

We observe that the use of optimal privacy at risk provided significantly stronger privacy guarantees as compared to the conventional composition theorems. Advanced composition theorem is known to provide stronger privacy guarantees for mechanism with smaller εs. As we observe in Figure 5c and Figure 5b, the composition provides strictly stronger privacy guarantees than basic composition, in the cases where the advanced composition fails.

Comparison with the Moment Accountant

Papernot et al. [33, 34] empirically showed that the privacy guarantees provided by the advanced composition theorem are quantitatively worse than the ones achieved by the state-of-the-art *moment accountant* [1]. The moment accountant evaluates the privacy guarantee by keeping track of various moments of privacy loss random variables. The computation of the moments is performed by using numerical methods on the specified dataset. Therefore, despite the quantitative strength of privacy guarantee provided by the moment accountant, it is qualitatively weaker, in a sense that it is specific to the dataset used for evaluation, in constrast to advanced composition.

Papernot et al. [33] introduced the PATE framework that uses the Laplace mechanism to provide privacy guarantees for a machine learning model trained in an ensemble manner. We comparatively evaluate the privacy guarantees provided by their moment accountant on MNIST dataset with the privacy guarantees obtained using privacy at risk. We do so by using privacy at risk while computing a data dependent bound [33, Theorem 3]. Under the identical experimental setup, we use a 0.1-differentially private Laplace mechanism, which optimally satisfies $(0.08, 0.8)$-privacy at risk. We list the calculated privacy guarantees in Table 1. The reported privacy guarantee is the mean privacy guarantee over 30 experiments.

6 Balancing Utility and Privacy

In this section, we empirically illustrate and discuss the steps that a data steward needs to take and the issues that she needs to consider in order to realise a required privacy at risk level ε for a confidence level γ when seeking to disclose the result of a query.

We consider a query that returns the parameter of a ridge regression [31] for an input dataset. It is a basic and widely used statistical analysis tool. We use the privacy-preserving mechanism presented by Ligett et al. [26] for ridge regression. It is a Laplace mechanism that induces noise in the output parameters of the ridge regression. The authors provide a theoretical upper bound on the sensitivity of the ridge regression, which we refer as *sensitivity*, in the experiments.

6.1 Dataset and Experimental Setup.

We conduct experiments on a subset of the 2000 US census dataset provided by Minnesota Population Center in its Integrated Public Use Microdata Series [39]. The census dataset consists of 1% sample of the original census data. It spans over 1.23 million households with records of 2.8 million people. The value of several attributes is not necessarily available for every household. We have therefore selected 212,605 records, corresponding to the household heads, and 6 attributes, namely,

δ	#Queries	Privacy level for moment accountant(ε)	
		with differential privacy [33]	with privacy at risk
10^{-5}	100	2.04	**1.81**
10^{-5}	1000	8.03	**5.95**

Table 1. Comparative analysis of privacy levels computed using three composition theorems when applied to 0.1-differentially private Laplace mechanism, which optimally satisfies $(0.08, 0.8)$-privacy at risk. The observations for the moment accountant on MNIST datasets are taken from [33].

Age, Gender, Race, Marital Status, Education, Income, whose values are available for the $212,605$ records.

In order to satisfy the constraint in the derivation of the sensitivity of ridge regression [26], we, without loss of generality, normalise the dataset in the following way. We normalise *Income* attribute such that the values lie in $[0, 1]$. We normalise other attributes such that l_2 norm of each data point is unity.

All experiments are run on Linux machine with 12-core 3.60GHz Intel® Core i7™processor with 64GB memory. Python® 2.7.6 is used as the scripting language.

6.2 Result Analysis

We train ridge regression model to predict *Income* using other attributes as predictors. We split the dataset into the training dataset (80%) and testing dataset (20%). We compute the *root mean squared error (RMSE)* of ridge regression, trained on the training data with regularisation parameter set to 0.01, on the testing dataset. We use it as the metric of *utility loss*. Smaller the value of RMSE, smaller the loss in utility. For a given value of privacy at risk level, we compute 50 runs of an experiment of a differentially private ridge regression and report the means over the 50 runs of the experiment.

Let us now provide illustrative experiments under the three different cases. In every scenario, the data steward is given a privacy at risk level ε and the confidence level γ and wants to disclose the parameters of a ridge regression model that she trains on the census dataset. She needs to calibrate the Laplace mechanism by estimating either its privacy level ε_0 (Case 1) or sensitivity (Case 2) or both (Case 3) to achieve the privacy at risk required the ridge regression query.

The Case of Explicit Randomness (cf. Section 4.1). In this scenario, the data steward knows the sensitivity for the ridge regression. She needs to compute the privacy level, ε_0, to calibrate the Laplace mechanism. She uses Equation 3 that links the desired privacy at risk level ε, the confidence level γ_1 and the privacy level of noise ε_0. Specifically, for given ε and γ_1, she computes ε_0 by solving the equation:

$$\gamma_1 \mathbb{P}(T \leq \varepsilon_0) - \mathbb{P}(T \leq \varepsilon) = 0.$$

Since the equation does not give an analytical formula for ε_0, the data steward uses a root finding algorithm such as Newton-Raphson method [37] to solve the above equation. For instance, if she needs to achieve a privacy at risk level $\varepsilon = 0.4$ with confidence level $\gamma_1 = 0.6$, she can substitute these values in the above equation and solve the equation to get the privacy level of noise $\varepsilon_0 = 0.8$.

Figure 6 shows the variation of privacy at risk level ε and confidence level γ_1. It also depicts the variation of utility loss for different privacy at risk levels in Figure 6.

In accordance to the data steward's problem, if she needs to achieve a privacy at risk level $\varepsilon = 0.4$ with confidence level $\gamma_1 = 0.6$, she obtains the privacy level of noise to be $\varepsilon_0 = 0.8$. Additionally, we observe that the choice of privacy level 0.8 instead of 0.4 to calibrate the Laplace mechanism gives lower utility loss for the data steward. This is the benefit drawn from the risk taken under the control of privacy at risk.

Thus, she uses privacy level ε_0 and the sensitivity of the function to calibrate Laplace mechanism.

The Case of Implicit Randomness (cf. Section 4.2). In this scenario, the data steward does not know the sensitivity of ridge regression. She assesses that she can afford to sample at most n times from the population dataset. She understands the effect of the uncertainty introduced by the statistical estimation of the sensitivity. Therefore, she uses the confidence level for empirical privacy at risk $\hat{\gamma}_2$.

Given the value of n, she chooses the value of the accuracy parameter using Figure 2. For instance, if the number of samples that she can draw is 10^4, she chooses the value of the accuracy parameter $\rho = 0.01$. Next, she uses Equation 13 to determine the value of probabilistic tolerance, α, for the sample size n. For instance, if the data steward is not allowed to access more than $15,000$

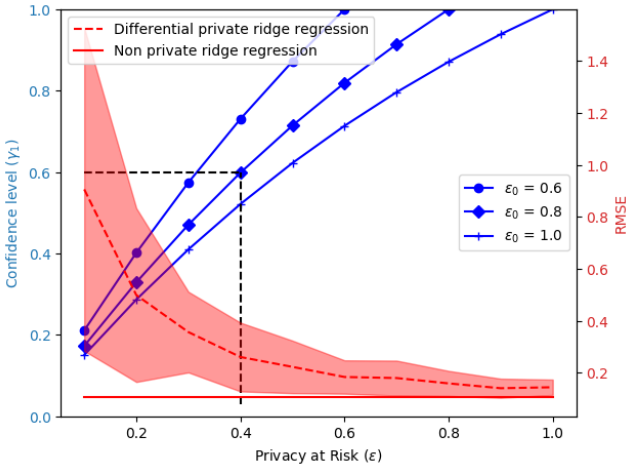


Fig. 6. Utility, measured by RMSE (right y-axis), and privacy at risk level ε for Laplace mechanism (left y-axis) for varying confidence levels γ_1.

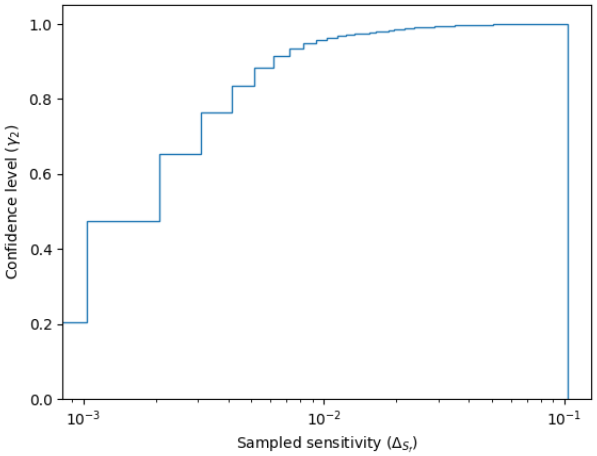


Fig. 7. Empirical cumulative distribution of the sensitivities of ridge regression queries constructed using 15000 samples of neighboring datasets.

samples, for the accuracy of 0.01 the probabilistic tolerance is 0.9.

$$\alpha = 1 - 2e^{(-2\rho^2 n)} \tag{13}$$

She constructs an empirical cumulative distribution over the sensitivities as described in Section 4.2. Such an empirical cumulative distribution is shown in Figure 7. Using the computed probabilistic tolerance and desired confidence level $\hat{\gamma_2}$, she uses equation in Theorem 5 to determine γ_2. She computes the sampled sensitivity using the empirical distribution function and the confidence level for privacy Δ_{S_f} at risk γ_2. For instance, using the empirical cumulative distribution in Figure 7 she calculates the value of the sampled sensitivity to be approximately 0.001 for $\gamma_2 = 0.4$ and approximately 0.01 for $\gamma_2 = 0.85$

Thus, she uses privacy level ε, sets the number of samples to be n and computes the sampled sensitivity Δ_{S_f} to calibrate the Laplace mechanism.

The Case of Explicit and Implicit Randomness (cf. Section 4.3). In this scenario, the data steward does not know the sensitivity of ridge regression. She is not allowed to sample more than n times from a population dataset. For a given confidence level γ_2 and the privacy at risk ε, she calibrates the Laplace mechanism using illustration for Section 4.3. The privacy level in this calibration yields utility loss that is more than her requirement. Therefore, she wants to re-calibrate the Laplace mechanism in order to reduce utility loss.

For the re-calibration, the data steward uses privacy level of the pre-calibrated Laplace mechanism, i.e. ε, as the privacy at risk level and she provides a new confidence level for empirical privacy at risk $\hat{\gamma_3}$. Using Equation 25 and Equation 23, she calculates:

$$\hat{\gamma_3}\mathbb{P}(T \leq \eta\varepsilon_0) - \alpha\gamma_2\ \mathbb{P}(T \leq \varepsilon) = 0$$

She solves such an equation for ε_0 using the root finding technique such as Newton-Raphson method [37]. For instance, if she needs to achieve a privacy at risk level $\varepsilon = 0.4$ with confidence levels $\hat{\gamma_3} = 0.9$ and $\gamma_2 = 0.9$, she can substitute these values and the values of tolerance parameter and sampled sensitivity, as used in the previous experiments, in the above equation. Then, solving the equation leads to the privacy level of noise $\varepsilon_0 = 0.8$.

Thus, she re-calibrates the Laplace mechanism with privacy level ε_0, sets the number of samples to be n and sampled sensitivity Δ_{S_f}.

7 Related Work

Calibration of mechanisms. Researchers have proposed different privacy-preserving mechanisms to make different queries differentially private. These mechanisms can be broadly classified into two categories. In one category, the mechanisms explicitly add calibrated noise, such as Laplace noise in the work of [11] or Gaussian noise in the work of [12], to the outputs of the query. In the other category, [2, 6, 17, 41] propose mechanisms that alter the query function so that the modified function satisfies differentially privacy. Privacy-preserving mechanisms in both of these categories perturb the original output of the query and make it difficult for a malicious data analyst to recover the original output of the query. These mechanisms induce randomness us-

ing the explicit noise distribution. Calibration of these mechanisms require the knowledge of the sensitivity of the query. Nissim et al. [32] consider the implicit randomness in the data-generation distribution to compute an estimate of the sensitivity. The authors propose the smooth sensitivity function that is an envelope over the local sensitivities for all individual datasets. Local sensitivity of a dataset is the maximum change in the value of the query over all of its neighboring datasets. In general, it is not easy to analytically estimate the smooth sensitivity function for a general query. Rubinstein et al. [38] also study the inherent randomness in the data-generation algorithm. We adopt their approach of sampling the sensitivity from the empirical distribution of the sensitivity. They use order statistics to choose a particular value of the sensitivity. We use the risk, which provides a mediation tool for business entities to assess the actual business risks, on the sensitivity distribution to estimate the sensitivity.

Refinements of differential privacy. In order to account for both sources of randomness, refinements of ε-differential privacy are proposed in order to bound the probability of occurrence of worst case scenarios. Machanavajjhala et al. [27] propose probabilistic differential privacy that considers upper bounds of the worst case privacy loss for corresponding confidence levels on the noise distribution. Definition of probabilistic differential privacy incorporates the explicit randomness induced by the noise distribution and bounds the probability over the space of noisy outputs to satisfy the ε-differential privacy definition. Dwork et al. [13] propose Concentrated differential privacy that considers the expected values of the privacy loss random variables for the corresponding. Definition of concentrated differential privacy incorporates the explicit randomness induced by the noise distribution but considering only the expected value of privacy loss satisfying ε-differential privacy definition instead of using the confidence levels limits its scope.

Hall et al. [17] propose random differential privacy that considers the privacy loss for corresponding confidence levels on the implicit randomness in the data-generation distribution. Definition of random differential privacy incorporates the implicit randomness induced by the data-generation distribution and bounds the probability over the space of datasets generated from the given distribution to satisfy the ε-differential privacy definition. Dwork et al. [9] define approximate differential privacy by adding a constant bias to the privacy guarantee provided by the differential privacy. It is not a probabilistic refinement of the differential privacy.

Around the same time of our work, Triastcyn et al. [40] independently propose Bayesian differential privacy that takes into account both of the sources of randomness. Despite this similarity, our works differ in multiple dimensions. Firstly, they have shown the reduction of their definition to a variant of Renyi differential privacy. The variant depends on the data-generation distribution. Secondly, they rely on the moment accountant for the composition of the mechanisms. Lastly, they do not provide a finer case-by-case analysis of the source of randomness, which leads to analytical solutions for the privacy guarantee.

Kifer et al. [24] define Pufferfish privacy framework, and its variant by Bassily et al. [4], that considers randomness due to data-generation distribution as well as noise distribution. Despite the generality of their approach, the framework relies on the domain expert to define a set of *secrets* that they want to protect.

We refer interested readers to [8] for an extensive review of the differential privacy and its refinements.

Composition theorem. Recently proposed technique of the *moment accountant* [1] has become the state-of-the-art of composing mechanisms in the area of privacy-preserving machine learning. Abadi et al. show that the moment accountant provides much strong privacy guarantees than the conventional composition mechanisms. It works by keeping track of various moments of privacy loss random variable and use the bounds on them to provide privacy guarantees. The moment accountant requires access to data-generation distribution to compute the bounds on the moment. Hence, the privacy guarantees are specific to the dataset.

Cost models. [7, 15] propose game theoretic methods that provide the means to evaluate the monetary cost of differential privacy. Pejó et al. [36] also propose a game theoretic cost model in the setting of private collaborative learning. Our approach is inspired by the approach in the work of Hsu et al. [19]. They model the cost under a scenario of a research study wherein the participants are reimbursed for their participation. Our cost modelling is driven by the scenario of securing a compensation budget in compliance with GDPR. Our requirement differs from the requirements of [19], particularly in our model participants do not have any monetary incentive to share their data.

8 Conclusion and Future Works

In this paper, we provide a means to fine-tune the privacy level of a privacy-preserving mechanism by analysing various sources of randomness. Such a fine-tuning leads to probabilistic quantification on privacy levels with quantified risks, which we call as privacy at risk. We also provide composition theorem that leverages privacy at risk. We analytical calculate privacy at risk for Laplace mechanism. We propose a cost model that bridges the gap between the privacy level and the compensation budget estimated by a GDPR compliant business entity. Convexity of the cost function ensures existence of unique privacy at risk that minimises compensation budget. The cost model helps in not only reinforcing the ease of application in a business setting but also providing stronger privacy guarantees on the composition of mechanism.

It is possible to instantiate privacy at risk for Gaussian mechanism. The mechanism is (ε, δ)-differential private for $\gamma = 0$ and a non-zero risk calculated by accounting for the sources of randomness. We save it for future work. Privacy at risk may be fully analytically computed in cases where the data-generation, or the sensitivity distribution, the noise distribution and the query are analytically known and take convenient forms. We are now looking at such convenient but realistic cases.

Acknowledgement

This work was supported by the National Research Foundation (NRF) Singapore under its Corporate Laboratory@University Scheme, National University of Singapore, and Singapore Telecommunications Ltd. This research was also funded in part by the BioQOP project of the French ANR (ANR-17-CE39-0006). We thank Pierre Senellart for his help in reviewing derivations of privacy at risk for the Laplace mechanism.

References

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 308–318, 2016.

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

[3] RA Askey and AB Olde Daalhuis. Generalized hypergeometric functions and meijer g-function. *NIST handbook of mathematical functions*, pages 403–418, 2010.

[4] Raef Bassily, Adam Groce, Jonathan Katz, and Adam Smith. Coupled-worlds privacy: Exploiting adversarial uncertainty in statistical data privacy. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 439–448. IEEE, 2013.

[5] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. *CoRR*, 2016.

[6] Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(Mar):1069–1109, 2011.

[7] Yiling Chen, Stephen Chong, Ian A Kash, Tal Moran, and Salil Vadhan. Truthful mechanisms for agents that value privacy. *ACM Transactions on Economics and Computation (TEAC)*, 4(3):13, 2016.

[8] Damien Desfontaines and Balázs Pejó. Sok: Differential privacies. *Proceedings on Privacy Enhancing Technologies*, 2020(2):288–313, 2020.

[9] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Eurocrypt*, volume 4004, pages 486–503. Springer, 2006.

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

[11] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. *Calibrating Noise to Sensitivity in Private Data Analysis*, pages 265–284. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[12] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

[13] Cynthia Dwork and Guy N Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016.

[14] Simson L Garfinkel, John M Abowd, and Sarah Powazek. Issues encountered deploying differential privacy. *arXiv preprint arXiv:1809.02201*, 2018.

[15] Arpita Ghosh and Aaron Roth. Selling privacy at auction. *Games and Economic Behavior*, 91:334–346, 2015.

[16] Rob Hall, Alessandro Rinaldo, and Larry Wasserman. Random differential privacy. *Journal of Privacy and Confidentiality*, 4(2):43–59, 2012.

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

[18] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. In *The Collected Works of Wassily Hoeffding*, pages 409–426. Springer, 1994.

[19] Justin Hsu, Marco Gaboardi, Andreas Haeberlen, Sanjeev Khanna, Arjun Narayan, Benjamin C Pierce, and Aaron Roth. Differential privacy: An economic method for choosing epsilon. In *Computer Security Foundations Symposium (CSF), 2014 IEEE 27th*, pages 398–410. IEEE, 2014.

[20] Wolfram Research, Inc. Mathematica, Version 10. Champaign, IL, 2014.

[21] Philippe Jorion. Value at risk: The new benchmark for managing financial risk, 01 2000.

[22] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. In *International conference on machine learning*, pages 1376–1385, 2015.

[23] Daniel Kifer and Bing-Rong Lin. An axiomatic view of statistical privacy and utility. *Journal of Privacy and Confidentiality*, 4(1), 2012.

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

[25] Jaewoo Lee and Chris Clifton. How much is enough? choosing ε for differential privacy. In *International Conference on Information Security*, pages 325–340. Springer, 2011.

[26] Katrina Ligett, Seth Neel, Aaron Roth, Bo Waggoner, and Steven Z Wu. Accuracy first: Selecting a differential privacy level for accuracy constrained erm. In *Advances in Neural Information Processing Systems*, pages 2563–2573, 2017.

[27] Ashwin Machanavajjhala, Daniel Kifer, John Abowd, Johannes Gehrke, and Lars Vilhuber. Privacy: Theory meets practice on the map. In *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on*, pages 277–286. IEEE, 2008.

[28] Pascal Massart et al. The tight constant in the dvoretzky-kiefer-wolfowitz inequality. *The annals of Probability*, 18(3):1269–1283, 1990.

[29] Sebastian Meiser. Approximate and probabilistic differential privacy definitions. *IACR Cryptology ePrint Archive*, 2018:277, 2018.

[30] James P Moriarty, Megan E Branda, Kerry D Olsen, Nilay D Shah, Bijan J Borah, Amy E Wagie, Jason S Egginton, and James M Naessens. The effects of incremental costs of smoking and obesity on health care costs among adults: a 7-year longitudinal study. *Journal of Occupational and Environmental Medicine*, 54(3):286–291, 2012.

[31] Kevin P. Murphy. *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012.

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

[33] Nicolas Papernot, Martín Abadi, Úlfar Erlingsson, Ian J. Goodfellow, and Kunal Talwar. Semi-supervised knowledge transfer for deep learning from private training data. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*. OpenReview.net, 2017.

[34] Nicolas Papernot, Shuang Song, Ilya Mironov, Ananth Raghunathan, Kunal Talwar, and Úlfar Erlingsson. Scalable private learning with PATE. *CoRR*, abs/1802.08908, 2018.

[35] Athanasios Papoulis and S Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[36] Balazs Pejo, Qiang Tang, and Gergely Biczok. Together or alone: The price of privacy in collaborative learning. *Proceedings on Privacy Enhancing Technologies*, 2019(2):47–65, 2019.

[37] William H Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[38] Benjamin IP Rubinstein and Francesco Aldà. Pain-free random differential privacy with sensitivity sampling. In *International Conference on Machine Learning*, pages 2950–2959, 2017.

[39] Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. Integrated public use microdata series: Version 6.0 [dataset], 2015.

[40] Aleksei Triastcyn and Boi Faltings. Federated learning with bayesian differential privacy. *arXiv preprint arXiv:1911.10071*, 2019.

[41] Jun Zhang, Zhenjie Zhang, Xiaokui Xiao, Yin Yang, and Marianne Winslett. Functional mechanism: regression analysis under differential privacy. *Proceedings of the VLDB Endowment*, 5(11):1364–1375, 2012.

A Proof of Theorem 2 (Section 3)

Proof. Let us fix a pair of neighbouring datasets x and y, and also fix a subset of outputs $Z \subseteq Range(\mathcal{M})$. Choose OUT $\subseteq Z$ such that

$$\text{OUT} \triangleq \left\{ z \in Z \;\middle|\; \frac{\mathbb{P}(\mathcal{M}(f,\Theta)(x)=z)}{\mathbb{P}(\mathcal{M}(f,\Theta)(y)=z)} > e^{\varepsilon} \right\}$$

Thus, we also obtain a complementary set of outputs $Z \setminus$ OUT, where the privacy loss is upper bounded i.e. $\mathbb{P}(\mathcal{M}(f,\Theta)(x)=z) \leq e^{\varepsilon}\mathbb{P}(\mathcal{M}(f,\Theta)(y)=z)$. Thus,

$$\begin{aligned}
&\mathbb{P}(\mathcal{M}(f,\Theta)(x) \in Z) \\
&= \mathbb{P}(\mathcal{M}(f,\Theta)(x) \in Z \backslash \text{OUT}) + \mathbb{P}(\mathcal{M}(f,\Theta)(x) \in \text{OUT}) \\
&\underset{(a)}{\leq} e^{\varepsilon}\mathbb{P}(\mathcal{M}(f,\Theta)(y) \in Z \backslash \text{OUT}) + \mathbb{P}(\mathcal{M}(f,\Theta)(x) \in \text{OUT}) \\
&\underset{(b)}{\leq} e^{\varepsilon}\mathbb{P}(\mathcal{M}(f,\Theta)(y) \in Z \setminus \text{OUT}) + \gamma \\
&\leq e^{\varepsilon}\mathbb{P}(\mathcal{M}(f,\Theta)(y) \in Z) + \gamma
\end{aligned}$$

(a) is obtained from the fact that the privacy loss is upper bounded in the complimentary set of outputs $Z \setminus$ OUT. (b) is a consequence of the definition of (ε, γ) privacy at risk. The definition upper bounds the probability measures of the subset OUT by γ. □

B Proof of Theorem 4 (Section 4.1)

Although a Laplace mechanism $\mathcal{L}_{\varepsilon}^{\Delta_f}$ induces higher amount of noise on average than a Laplace mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$ for $\varepsilon < \varepsilon_0$, there is a non-zero probability that $\mathcal{L}_{\varepsilon}^{\Delta_f}$ induces noise commensurate to $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$. This non-zero probability guides us to calculate the privacy at risk γ_1 for the privacy at risk level ε. In order to get an intuition, we illustrate the calculation of the overlap between two Laplace distributions as an estimator of similarity between the two distributions.

Definition 8. *[Overlap of Distributions, [35]] The overlap, O, between two probability densities P_1, P_2 with support $\mathcal{X}$ is defined as*

$$O = \int_{\mathcal{X}} \min[P_1(x), P_2(x)]\, dx.$$

Lemma 5. *The overlap O between two probability distributions,* $\text{Lap}(\frac{\Delta_f}{\varepsilon_1})$ *and* $\text{Lap}(\frac{\Delta_f}{\varepsilon_2})$*, such that $\varepsilon_2 \leq \varepsilon_1$, is given by*

$$O = 1 - (\exp(-\mu\varepsilon_2/\Delta_f) - \exp(-\mu\varepsilon_1/\Delta_f)),$$

where $\mu = \frac{\Delta_f \ln(\varepsilon_1/\varepsilon_2)}{\varepsilon_1 - \varepsilon_2}$.

Using the result in Lemma 5, we note that the overlap between two distributions with $\varepsilon_0 = 1$ and $\varepsilon = 0.6$ is 0.81. Thus, $\mathcal{L}_{0.6}^{\Delta_f}$ induces noise that is more than 80% times similar to the noise induced by $\mathcal{L}_{1.0}^{\Delta_f}$. Therefore, we can loosely say that at least 80% of the times a Laplace Mechanism $\mathcal{L}_{1.0}^{\Delta_f}$ will provide the same privacy as a Laplace Mechanism $\mathcal{L}_{0.8}^{\Delta_f}$.

Although the overlap between Laplace distributions with different scales offers an insight into the relationship between different privacy levels, it does not capture the constraint induced by the *sensitivity*. For a given query f, the amount of noise required to satisfy differential privacy is commensurate to the sensitivity of the query. This calibration puts a constraint on the noise that is required to be induced on a pair of neighbouring datasets. We state this constraint in Lemma 6, which we further use to prove that the Laplace Mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$ satisfies (ε, γ_1)-privacy at risk.

Lemma 6. *For a Laplace Mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$, the difference in the absolute values of noise induced on a pair of neighbouring datasets is upper bounded by the sensitivity of the query.*

Proof. Suppose that two neighbouring datasets x and y are given input to a numeric query $f : \mathcal{D} \to \mathbb{R}^k$. For any output $z \in \mathbb{R}^k$ of the Laplace Mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$,

$$\begin{aligned}
\sum_{i=1}^{k} (|f(y_i) - z_i| - |f(x_i) - z_i|) &\leq \sum_{i=1}^{k} (|f(x_i) - f(y_i)|) \\
&\leq \Delta_f.
\end{aligned}$$

We use triangular inequality in the first step and Definition 2 of sensitivity in the second step. □

We write $\text{Exp}(b)$ to denote a random variable sampled from an *exponential distribution* with scale $b > 0$. We write $\text{Gamma}(k, \theta)$ to denote a random variable sampled from a *gamma distribution* with shape $k > 0$ and scale $\theta > 0$.

Lemma 7. *[[35]] If a random variable X follows Laplace Distribution with mean zero and scale b, $|X| \sim \text{Exp}(b)$.*

Lemma 8. *[[35]] If $X_1, ..., X_n$ are n i.i.d. random variables each following the Exponential Distribution with scale b, $\sum_{i=1}^{n} X_i \sim Gamma(n, b)$.*

Lemma 9. *If X_1 and X_2 are two i.i.d. Gamma(n,θ) random variables, the probability density function for the random variable $T = |X_1 - X_2|/\theta$ is given by*

$$P_T(t;n,\theta) = \frac{2^{2-n}t^{n-\frac{1}{2}}K_{n-\frac{1}{2}}(t)}{\sqrt{2\pi}\Gamma(n)\theta}$$

where $K_{n-\frac{1}{2}}$ is the modified Bessel function of second kind.

Proof. Let X_1 and X_2 be two i.i.d. Gamma(n,θ) random variables. Characteristic function of a Gamma random variable is given as

$$\phi_{X_1}(z) = \phi_{X_2}(z) = (1-\iota z\theta)^{-n}.$$

Therefore,

$$\phi_{X_1-X_2}(z) = \phi_{X_1}(z)\phi^*_{X_2}(z) = \frac{1}{(1+(z\theta)^2)^n}$$

Probability density function for the random variable $X_1 - X_2$ is given by,

$$\begin{aligned} P_{X_1-X_2}(x) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\mathrm{i}zx}\phi_{X_1-X_2}(z)dz \\ &= \frac{2^{1-n}|\frac{x}{\theta}|^{n-\frac{1}{2}}K_{n-\frac{1}{2}}(|\frac{x}{\theta}|)}{\sqrt{2\pi}\Gamma(n)\theta} \end{aligned}$$

where $K_{n-\frac{1}{2}}$ is the Bessel function of second kind. Let $T = |\frac{X_1-X_2}{\theta}|$. Therefore,

$$P_T(t;n,\theta) = \frac{2^{1-n}t^{n-\frac{1}{2}}K_{n-\frac{1}{2}}(t)}{\sqrt{2\pi}\Gamma(n)\theta}$$

We use Mathematica [20] to solve the above integral. □

Lemma 10. *If X_1 and X_2 are two i.i.d. Gamma(n,θ) random variables and $|X_1 - X_2| \leq M$, then $T' = |X_1 - X_2|/\theta$ follows the distribution with probability density function:*

$$P_{T'}(t;n,\theta,M) = \frac{P_T(t';n,\theta)}{P_T(T \leq M)},$$

where P_T is the probability density function of defined in Lemma 9.

Lemma 11. *For Laplace Mechanism $\mathcal{L}^{\Delta_f}_{\varepsilon_0}$ with query $f : \mathcal{D} \to \mathbb{R}^k$ and for any output $Z \subseteq Range(\mathcal{L}^{\Delta_f}_{\varepsilon_0})$, $\varepsilon \leq \varepsilon_0$,*

$$\gamma_1 \triangleq \mathbb{P}\left[\ln\left|\frac{\mathbb{P}(\mathcal{L}^{\Delta_f}_{\varepsilon_0}(x)\in Z)}{\mathbb{P}(\mathcal{L}^{\Delta_f}_{\varepsilon_0}(y)\in Z)}\right| \leq \varepsilon\right] = \frac{\mathbb{P}(T\leq\varepsilon)}{\mathbb{P}(T\leq\varepsilon_0)},$$

where T follows the distribution in Lemma 9, $P_T(t;k,\frac{\Delta_f}{\varepsilon_0})$.

Proof. Let, $x \in \mathcal{D}$ and $y \in \mathcal{D}$ be two datasets such that $x \sim y$. Let $f : \mathcal{D} \to \mathbb{R}^k$ be some numeric query. Let $\mathbb{P}_x(z)$ and $\mathbb{P}_y(z)$ denote the probabilities of getting the output z for Laplace mechanisms $\mathcal{L}^{\Delta_f}_{\varepsilon_0}(x)$ and $\mathcal{L}^{\Delta_f}_{\varepsilon_0}(y)$ respectively. For any point $z \in \mathbb{R}^k$ and $\varepsilon \neq 0$,

$$\begin{aligned} \frac{\mathbb{P}_x(z)}{\mathbb{P}_y(z)} &= \prod_{i=1}^{k}\frac{\exp\left(\frac{-\varepsilon_0|f(x_i)-z_i|}{\Delta_f}\right)}{\exp\left(\frac{-\varepsilon_0|f(y_i)-z_i|}{\Delta_f}\right)} \\ &= \prod_{i=1}^{k}\exp\left(\frac{\varepsilon_0(|f(y_i)-z_i| - |f(x_i)-z_i|)}{\Delta_f}\right) \\ &= \exp\left(\varepsilon\left[\frac{\varepsilon_0\sum_{i=1}^{k}(|f(y_i)-z_i| - |f(x_i)-z_i|)}{\varepsilon\Delta_f}\right]\right). \end{aligned} \tag{14}$$

By Definition 4,

$$(f(x)-z), (f(y)-z) \sim \mathrm{Lap}(\Delta_f/\varepsilon_0). \tag{15}$$

Application of Lemma 7 and Lemma 8 yields,

$$\sum_{i=1}^{k}(|f(x_i)-z_i|) \sim \mathrm{Gamma}(k,\Delta_f/\varepsilon_0). \tag{16}$$

Using Equations 15, 16, and Lemma 6, 10, we get

$$\left(\frac{\varepsilon_0}{\Delta_f}\sum_{i=1}^{k}|(|f(y_i)-z| - |f(x_i)-z|)|\right) \sim P_{T'}(t;k,\Delta_f/\varepsilon_0,\Delta_f). \tag{17}$$

since, $\sum_{i=1}^{k}|(|f(y_i)-z| - |f(x_i)-z|)| \leq \Delta_f$. Therefore,

$$\mathbb{P}\left(\left[\frac{\varepsilon_0}{\Delta_f}\sum_{i=1}^{k}|(|f(y_i)-z| - |f(x_i)-z|)|\right] \leq \varepsilon\right) = \frac{\mathbb{P}(T\leq\varepsilon)}{\mathbb{P}(T\leq\varepsilon_0)}, \tag{18}$$

where T follows the distribution in Lemma 9. We use Mathematica [20] to analytically compute,

$$\mathbb{P}(T\leq x) \propto \left({}_1F_2(\frac{1}{2};\frac{3}{2}-k,\frac{3}{2};\frac{x^2}{4})\sqrt{\pi}4^k x]\right) - \left(2{}_1F_2(k;\frac{1}{2}+k,k+1;\frac{x^2}{4})x^{2k}\Gamma(k)\right)$$

where ${}_1F_2$ is the regularised generalised hypergeometric function as defined in [3]. From Equation 14 and 18,

$$\mathbb{P}\left[\ln\left|\frac{\mathbb{P}(\mathcal{L}^{\Delta_f}_{\varepsilon_0}(x)\in S)}{\mathbb{P}(\mathcal{L}^{\Delta_f}_{\varepsilon_0}(y)\in S)}\right| \leq \varepsilon\right] = \frac{\mathbb{P}(T\leq\varepsilon)}{\mathbb{P}(T\leq\varepsilon_0)}.$$

□

This completes the proof of Theorem 4.

Corollary 3. *Laplace Mechanism* $\mathcal{L}_{\varepsilon_0}^{\Delta_f}$ *with* $f : \mathcal{D} \to \mathbb{R}^k$ *satisfies* (ε, δ)*-probabilistic differentially private where*

$$\delta = \begin{cases} 1 - \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \varepsilon_0)} & \varepsilon \leq \varepsilon_0 \\ 0 & \varepsilon > \varepsilon_0 \end{cases}$$

and T *follows* $\text{BesselK}(k, \Delta_f/\varepsilon_0)$.

C Proof of Theorem 5 (Section 4.2)

Proof. Let, x and y be any two neighbouring datasets sampled from the data generating distribution $\mathcal{G}$. Let, Δ_{S_f} be the sampled sensitivity for query $f : \mathcal{D} \to \mathbb{R}^k$. Let, $\mathbb{P}_x(z)$ and $\mathbb{P}_y(z)$ denote the probabilities of getting the output z for Laplace mechanisms $\mathcal{L}_{\varepsilon}^{\Delta_{S_f}}(x)$ and $\mathcal{L}_{\varepsilon}^{\Delta_{S_f}}(y)$ respectively. For any point $z \in \mathbb{R}^k$ and $\varepsilon \neq 0$,

$$\begin{aligned} \frac{\mathbb{P}_x(z)}{\mathbb{P}_y(z)} &= \prod_{i=1}^{k} \frac{\exp\left(\frac{-\varepsilon|f(x_i)-z_i|}{\Delta_{S_f}}\right)}{\exp\left(\frac{-\varepsilon|f(y_i)-z_i|}{\Delta_{S_f}}\right)} \\ &= \exp\left(\frac{\varepsilon \sum_{i=1}^{k}(|f(y_i)-z_i| - |f(x_i)-z_i|)}{\Delta_{S_f}}\right) \\ &\leq \exp\left(\frac{\varepsilon \sum_{i=1}^{k}|f(y_i)-f(x_i)|}{\Delta_{S_f}}\right) \\ &= \exp\left(\frac{\varepsilon \|f(y)-f(x)\|_1}{\Delta_{S_f}}\right) \end{aligned} \tag{19}$$

We used triangle inequality in the penultimate step.

Using the trick in the work of [38], we define following events. Let, $B^{\Delta_{S_f}}$ denotes the set of pairs neighbouring dataset sampled from $\mathcal{G}$ for which the sensitivity random variable is upper bounded by Δ_{S_f}. Let, $C_{\rho}^{\Delta_{S_f}}$ denotes the set of sensitivity random variable values for which F_n deviates from the unknown cumulative distribution of S, F, at most by the accuracy value ρ. These events are defined in Equation 20.

$$\begin{aligned} B^{\Delta_{S_f}} &\triangleq \{x, y \sim \mathcal{G} \text{ such that } \|f(y)-f(x)\|_1 \leq \Delta_{S_f}\} \\ C_{\rho}^{\Delta_{S_f}} &\triangleq \left\{\sup_{\Delta} |F_S^n(\Delta) - F_S(\Delta)| \leq \rho\right\} \end{aligned} \tag{20}$$

Thus, we obtain

$$\begin{aligned} \mathbb{P}\left(B^{\Delta_{S_f}}\right) &= \mathbb{P}\left(B^{\Delta_{S_f}} \middle| C_{\rho}^{\Delta_{S_f}}\right)\mathbb{P}\left(C_{\rho}^{\Delta_{S_f}}\right) \\ &\quad + \mathbb{P}\left(B^{\Delta_{S_f}} \middle| \overline{C_{\rho}^{\Delta_{S_f}}}\right)\mathbb{P}\left(\overline{C_{\rho}^{\Delta_{S_f}}}\right) \\ &\geq \mathbb{P}\left(B^{\Delta_{S_f}} \middle| C_{\rho}^{\Delta_{S_f}}\right)\mathbb{P}\left(C_{\rho}^{\Delta_{S_f}}\right) \\ &= F_n(\Delta_{S_f})\mathbb{P}\left(C_{\rho}^{\Delta_{S_f}}\right) \\ &\geq \gamma_2 \cdot \left(1 - 2e^{-2\rho^2 n}\right) \end{aligned} \tag{21}$$

In the last step, we use the definition of the sampled sensitivity to get the value of the first term. The last term is obtained using DKW-inequality, as defined in [28], where the n denotes the number of samples used to build empirical distribution of the sensitivity, F_n.

From Equation 19, we understand that if $\|f(y) - f(x)\|_1$ is less than or equals to the sampled sensitivity then the Laplace mechanism $\mathcal{L}_{\varepsilon}^{\Delta_{S_f}}$ satisfies ε-differential privacy. Equation 21 provides the lower bound on the probability of the event $\|f(y) - f(x)\|_1 \leq \Delta_{S_f}$. Thus, combining Equation 19 and Equation 21 completes the proof. □

D Proof of Theorem 6 (Section 4.3)

Proof of Theorem 6 builds upon the ideas from the proofs for the rest of the two cases. In addition to the events defined in Equation 20, we define an additional event $A_{\varepsilon_0}^{\Delta_{S_f}}$, defined in Equation 22, as a set of outputs of Laplace mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_{S_f}}$ that satisfy the constraint of ε-differential privacy for a specified privacy at risk level ε.

$$A_{\varepsilon_0}^{\Delta_{S_f}} \triangleq \left\{ z \sim \mathcal{L}_{\varepsilon_0}^{\Delta_{S_f}} \;:\; \ln\left|\frac{\mathcal{L}_{\varepsilon_0}^{\Delta_{S_f}}(x)}{\mathcal{L}_{\varepsilon_0}^{\Delta_{S_f}}(y)}\right| \leq \varepsilon, x, y \sim \mathcal{G} \right\} \tag{22}$$

Corollary 4.

$$\mathbb{P}(A_{\varepsilon_0}^{\Delta_{S_f}} | B^{\Delta_{S_f}}) = \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)}$$

where T *follows the distribution* $P_T(t; \Delta_{S_f}/\varepsilon_0)$ *in Lemma 9 and* $\eta = \frac{\Delta_f}{\Delta_{S_f}}$.

Proof. We provide the sketch of the proof. Proof follows from the proof of Lemma 11. For a Laplace mechanism calibrated with the sampled sensitivity Δ_{S_f} and privacy level ε_0, Equation 17 translates to,

$$\left(\frac{\varepsilon_0}{\Delta_{S_f}} \sum_{i=1}^{k} |(|f(y_i) - z| - |f(x_i) - z|)|\right) \sim$$

$$P_{T'}(t; k, \Delta_{S_f}/\varepsilon_0, \Delta_{S_f}).$$

since, $\sum_{i=1}^{k} |(|f(y_i) - z| - |f(x_i) - z|)| \leq \Delta_f$. Using Lemma 10 and Equation 18,

$$\mathbb{P}(A_{\varepsilon_0}^{\Delta_{S_f}}) = \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)}$$

where T follows the distribution $P_T(t; \Delta_{S_f}/\varepsilon_0)$ and $\eta = \frac{\Delta_f}{\Delta_{S_f}}$. □

For this case, we do not assume the knowledge of the sensitivity of the query. Using the empirical estimation presented in Section 4.2, if we choose the sampled sensitivity for privacy at risk $\gamma_2 = 1$, we obtain an approximation for η.

Lemma 12. *For a given value of accuracy parameter ρ,*

$$\frac{\Delta_f}{\Delta^*_{S_f}} = 1 + \mathcal{O}\left(\frac{\rho}{\Delta^*_{S_f}}\right)$$

where $\Delta^*_{S_f} = F_n^{-1}(1)$. $\mathcal{O}\left(\frac{\rho}{\Delta^*_{S_f}}\right)$ denotes order of $\frac{\rho}{\Delta^*_{S_f}}$, i.e., $\mathcal{O}\left(\frac{\rho}{\Delta^*_{S_f}}\right) = k\frac{\rho}{\Delta^*_{S_f}}$ for some $k \geq 1$.

Proof. For a given value of accuracy parameter ρ and any $\Delta > 0$,

$$F_n(\Delta) - F(\Delta) \leq \rho$$

Since above inequality is true for any value of Δ, let $\Delta = F^{-1}(1)$. Therefore,

$$\begin{aligned} F_n(F^{-1}(1)) - F(F^{-1}(1)) &\leq \rho \\ F_n(F^{-1}(1)) &\leq 1 + \rho \end{aligned} \tag{23}$$

Since a cumulative distribution function is 1-Lipschitz [[35]],

$$\begin{aligned} |F_n(F_n^{-1}(1)) - F_n(F^{-1}(1))| &\leq |F_n^{-1}(1) - F^{-1}(1)| \\ |F_n(F_n^{-1}(1)) - F_n(F^{-1}(1))| &\leq |\Delta^*_{S_f} - \Delta_f| \\ \rho &\leq \Delta_f - \Delta^*_{S_f} \\ 1 + \frac{\rho}{\Delta^*_{S_f}} &\leq \frac{\Delta_f}{\Delta^*_{S_f}} \end{aligned}$$

where we used result from Equation 23 in step 3. Introducing $\mathcal{O}\left(\frac{\rho}{\Delta^*_{S_f}}\right)$ completes the proof. □

Lemma 13. *For Laplace Mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_{S_f}}$ with sampled sensitivity Δ_{S_f} of a query $f : \mathcal{D} \to \mathbb{R}^k$ and for any $Z \subseteq Range(\mathcal{L}_{\varepsilon}^{\Delta_{S_f}})$,*

$$\mathbb{P}\left[\ln\left|\frac{\mathbb{P}(\mathcal{L}_{\varepsilon_0}(x) \in Z)}{\mathbb{P}(\mathcal{L}_{\varepsilon_0}(y) \in Z)}\right| \leq \varepsilon\right] \geq \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)}\gamma_2(1-2e^{-2\rho^2 n})$$

where n is the number of samples used to find sampled sensitivity, $\rho \in [0, 1]$ is a accuracy parameter and $\eta = \frac{\Delta_f}{\Delta_{S_f}}$. The outer probability is calculated with respect to support of the data-generation distribution $\mathcal{G}$.

Proof. The proof follows from the proof of Lemma 11 and Lemma 13. Consider,

$$\begin{aligned} \mathbb{P}(A_{\varepsilon_0}^{\Delta_{S_f}}) &\geq \mathbb{P}(A_{\varepsilon_0}^{\Delta_{S_f}}|B^{\Delta_{S_f}})\mathbb{P}(B^{\Delta_{S_f}}|C_{\rho}^{\Delta_{S_f}})\mathbb{P}(C_{\rho}^{\Delta_{S_f}}) \\ &\geq \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)} \cdot \gamma_2 \cdot (1 - 2e^{-2\rho^2 n}) \end{aligned} \tag{24}$$

The first term in the final step of Equation 24 follows from the result in Corollary 4 where T follows BesselK$(k, \frac{\Delta_{S_f}}{\varepsilon_0})$. It is the probability with which the Laplace mechanism $\mathcal{L}_{\varepsilon_0}^{\Delta_{S_f}}$ satisfies ε-differential privacy for a given value of sampled sensitivity. □

Probability of occurrence of event $A_{\varepsilon_0}^{\Delta_{S_f}}$ calculated by accounting for both explicit and implicit sources of randomness gives the risk for privacy level ε. Thus, the proof of Lemma 13 completes the proof for Theorem 6.

Comparing the equations in Theorem 6 and Lemma 13, we observe that

$$\gamma_3 \triangleq \frac{\mathbb{P}(T \leq \varepsilon)}{\mathbb{P}(T \leq \eta\varepsilon_0)} \cdot \gamma_2 \tag{25}$$

The privacy at risk, as defined in Equation 25, is free from the term that accounts for the accuracy of sampled estimate. If we know cumulative distribution of the sensitivity, we do not suffer from the uncertainty of introduced by sampling from the empirical distribution.