# Laplace Transform Interpretation of Differential Privacy

**Rishav Chourasia**[1,2*] **Uzair Javaid**[1] **Biplap Sikdar**[2]

[1]Betterdata [2]National University of Singapore
{rishav, uzair}@betterdata.ai bsikdar@nus.edu.sg

**Abstract**

We introduce a set of useful expressions of Differential Privacy (DP) notions in terms of the Laplace transform of the privacy loss distribution. Its bare form expression appears in several related works on analyzing DP, either as an integral or an expectation. We show that recognizing the expression as a Laplace transform unlocks a new way to reason about DP properties by exploiting the duality between time and frequency domains. Leveraging our interpretation, we connect the $(q, \rho(q))$-Rényi DP curve and the $(\varepsilon, \delta(\varepsilon))$-DP curve as being the Laplace and inverse-Laplace transforms of one another. This connection shows that the Rényi divergence is well-defined for complex orders $q = \gamma + i\omega$. Using our Laplace transform-based analysis, we also prove an adaptive composition theorem for $(\varepsilon, \delta)$-DP guarantees that is exactly tight (i.e., matches even in constants) for all values of $\epsilon$. Additionally, we resolve an issue regarding symmetry of $f$-DP on subsampling that prevented equivalence across all functional DP notions.

## 1 Introduction

Differential privacy (DP) [13] has become a widely adopted standard for quantifying privacy of algorithms that process statistical data. In simple terms, differential privacy bounds the influence a single data-point may have on the outcome probabilities. Being a statistical property, the design of differentially private algorithms involves a *pen-and-paper analysis* of any randomness internal to the processing that obscures the influence a data-point might have on its output. A clear understanding of the nature of differential privacy notions is therefore tantamount to study and design of privacy-preserving algorithms.

Throughout its exploration, various functional interpretations of the concept of differential privacy have emerged over the years. These include the privacy-profile curve $\delta(\epsilon)$ [5] that traces the $(\epsilon, \delta)$-DP point guarantees, the $f$-DP [11] view of worst-case *trade-off curve* between type I and type II errors for hypothesis testing membership [19, 6], the Rényi DP [23] function of order $q$ that admits a natural analytical composition [1, 23], the view of the *privacy loss distribution (PLD)* [29] that allows for approximate numerical composition [20, 18], and the recent *characteristic function formulation* of the dominating privacy loss random variables Zhu et al. [32]. Each of these formalisms have their own properties and use-cases, and none of them seem to be superior in all aspects.

Regardless of their differences, they all have some shared difficulties—certain types of manipulations on them are harder to perform in the time-domain, but considerably simpler to do in the frequency-domain. For instance, Koskela et al. [20] noted that composing PLDs of two mechanisms involve convolving their probability densities, which can be numerically approximated efficiently

*Corresponding author

by multiplying their Discrete Fast-Fourier Transformations (DFFT) and then inverting it back to get the convolved density using Inverse-DFFT. Such maneuvers are also frequently performed for analytical reasons while proving properties of differential privacy, often without even realizing this detour through the frequency-domain. A notable example of this is the analysis of Moments' accountant by Abadi et al. [1], where the authors bound higher-order moments of subsampled Gaussian distributions, compose the moments through multiplication, and then derive the $(\varepsilon, \delta)$-DP bound on the DP-SGD mechanism. Their analysis goes the through frequency space, as the moment generating function of a random variable corresponds to the two-sided Laplace transform of its probability density function [22]. Often times when dealing with a functional notion of DP, expressing components like expectations or cumulative densities their integral form ends up being a Fourier or a Laplace transform. Realizing them as such can be tremendously useful in analysis.

In this paper, we formalize these time-frequency domain dualisms enjoyed by the functional representations into a new interpretation of differential privacy. In addition to augmenting existing perspectives on DP, this interpretation provides a flexible analytical toolkit that greatly extends our cognitive reach in reasoning about DP and its underpinnings. This interpretation is based on recognizing that the privacy-profile $\delta_{P|Q}(\varepsilon) := \sup_S P(S) - e^\varepsilon \cdot Q(S)$ and the Rényi-divergence $\mathrm{R}_q\left(P \| Q\right) := \frac{1}{q-1} \int_\Omega P^q Q^{1-q} \mathrm{d}\theta$ between any two distributions $P, Q$ on the same space $\Omega$ can be seen as a Laplace transform[1] of the privacy loss distribution $\mathrm{PLD}(P \| Q)$, the distribution of privacy loss random variable $Z = L_{P|Q}(\Theta)$ where $\Theta \sim P$:

$$\forall \varepsilon \in \mathbb{R} \ : \ \delta_{P|Q}(\varepsilon) = \underset{Z \leftarrow \mathrm{PLD}(P\|Q)}{\mathbb{E}} \left[\max\{0, 1 - e^{\varepsilon - Z}\}\right] = \mathcal{L}\left\{1 - F_Z(t + \varepsilon)\right\}(1), \tag{1}$$

$$\forall q \in \mathbb{C} \ : \ e^{(q-1)\cdot \mathrm{R}_q(P\|Q)} = \underset{Z \leftarrow \mathrm{PLD}(P\|Q)}{\mathbb{E}} \left[e^{(q-1)\cdot Z}\right] = \mathcal{B}\left\{f_Z(t)\right\}(1-q), \tag{2}$$

where $F_Z(t) = \Pr[Z < t]$ is the cumulative distribution function and $f_Z(t) = \int_{\{\theta \in \Omega : L_{P|Q}(\theta) = z\}} P \mathrm{d}\theta$ is the (generalized) density function of the privacy loss random variable $Z$. The first equality in (1) is a widely used way to represent the $(\varepsilon, \delta(\varepsilon))$-DP curve in literature [29, 6, 5, 20, 18, 30, 9]. Similarly, the first equality in (2) represents the well-known moment-generating function of privacy loss [23, 1, 6]. The second equalities above are part of a set of Laplace expressions presented in this paper. Together, these expressions unlock a formal approach to perform a wide-variety of manipulations and transformations on them using the fundamental properties of the Laplace functional (see Table 1). Using them, we show that the privacy-profile and Rényi divergence between any two distributions have the following equivalence.

$$\forall q \in \mathbb{C} \ : \ e^{(q-1)\cdot \mathrm{R}_q(P\|Q)} = q(q-1) \cdot \mathcal{B}\left\{\delta_{P|Q}(t)\right\}(1-q), \tag{3}$$

which again is a Laplace transform expression. Furthermore, Zhu et al. [32]'s characteristic function of the privacy loss $\phi_{P|Q}(q) := \underset{P}{\mathbb{E}}\left[e^{iq \log(P/Q)}\right]$ also turns out to be a Fourier transform, which is a special case of the bilateral Laplace transform:

$$\forall q \in \mathbb{R} \ : \ \phi_{P|Q}(q) = \underset{Z \leftarrow \mathrm{PLD}(P\|Q)}{\mathbb{E}} \left[e^{iqZ}\right] = \mathcal{B}\left\{f_Z\right\}(-iq). \tag{4}$$

These expressions can take advantage of the relationship between their time-domain and complex frequency-domain representations, as certain manipulations are more straightforward in one domain as compared to the other. Using the Laplace transform interpretations extensively, our paper presents the following findings.

[1]Laplace transform maps a time-domain function $g(t)$ with $t \in \mathbb{R}$ to a function $\mathcal{L}\left\{g\right\}(s) := \int_0^\infty e^{-st} g(t) \mathrm{d}t$ with $s \in \mathbb{C}$ in the complex space. Similarly, bilateral Laplace transform of $g(t)$ is defined as $\mathcal{B}\left\{g\right\}(s) := \int_{-\infty}^\infty e^{-st} g(t) \mathrm{d}t$.

1. We note that the Laplace transform expression of Rényi divergence permits the order $q$ to be a complex number in $\mathbb{C}$. Based on this observation, we revisit the discussion on equivalence and interconversion between $(q, \rho)$-Rényi DP and $(\varepsilon, \delta)$-DP in literature [6, 32, 2, 9]. We show that the privacy-profile curve $\delta_{P|Q}(\varepsilon)$ and the Rényi divergence $\mathrm{R}_q(P\|Q)$ as a function of $q$ are equivalent as long as either $P$ is absolutely continuous w.r.t. $Q$ (denote as $P \ll Q$) or $Q \ll P$; absolute continuity in both directions is not necessary. Moreover, we establish that while $\delta_{P_1|Q_1}(\varepsilon) \leq \delta_{P_2|Q_2}(\varepsilon)$ for all $\varepsilon$ implies that $\mathrm{R}_q(P_1\|Q_1) \leq \mathrm{R}_q(P_2\|Q_2)$ for all $q > 1$, the converse does not hold. This is due to the fact that the dominance relationship between privacy profiles $\delta_{P_1|Q_1}(\varepsilon)$ and $\delta_{P_2|Q_2}(\varepsilon)$ depends on how the Rényi divergence curves $\mathrm{R}_q(P_1\|Q_1)$ and $\mathrm{R}_q(P_2\|Q_2)$ behave along the complex line $\{q \in \mathbb{C} : \mathfrak{Re}(q) = c\}$ at any $c \in \mathbb{R}$ for which the two divergences exist; not along $(1, \infty)$.

2. Among all functional notions of DP, *exactly tight adaptive* composition theorem is only known for Rényi DP in an explicit form[2][23, 8]. And, for the PLD formalism, only non-adaptive composition theorems are known that are *exactly tight*[3] [18, 29, 20]. In this paper, we establish an exactly tight theorem for composing any two privacy profiles, $\delta_{P_1|Q_1}(\varepsilon)$ and $\delta_{P_2|Q_2}(\varepsilon)$, leveraging time-frequency dualities with their Rényi divergence curves. Our composition method also extends to adaptive scenarios, provided that the conditional distributions $P_2^\theta$ and $Q_2^\theta$, given an observation $\theta$ from the first distribution pair, are dominated by a privacy profile for all $\theta$, which is a standard assumption for adaptive composition guarantees.

3. We apply our composition theorem for privacy profiles to derive an optimal composition theorem for $(\epsilon_i, \delta_i)$-point guarantees of differential privacy. Our approach begins by determining the worst-case privacy profile $\delta_i(\epsilon)$ that *any* $(\epsilon_i, \delta_i)$-DP mechanism must satisfy. We then use our composition theorem to derive the combined privacy profile $\delta^{\otimes}(\epsilon)$. This provides the most precise composition guarantee possible when given only that a sequence of mechanisms each satisfies an $(\epsilon_i, \delta_i)$-DP point guarantee. Our bound surpasses the optimal composition result in Kairouz et al. [19, Theorem 3.3] because, whereas their result only provides a discrete set of $(\epsilon, \delta)$ values met by the composed curve, ours forms a continuous curve. This continuity enables us to determine the tightest $\epsilon$ value for any given $\delta$ budget. We also show that our results align with the bounds generated by numerical accountants such as Google's PLDAccountant [12] and Microsoft's PRVAccountant [18].

4. The concept of $f$-DP introduced by Dong et al. [11] provides a functional perspective on the indistinguishability between two distributions $P$ and $Q$ through hypothesis testing. The function $f : [0, 1] \to [0, 1]$ represents a bound on the trade-off $T(P, Q) : [0, 1] \to [0, 1]$ between Type-I and Type-II errors for any test aimed at determining whether a sample $\theta$ originates from $P$ or $Q$. Unlike other functional notions of differential privacy, $f$-DP is unique in being *not connected* to the rest via a Laplace transform. Instead, Dong et al. [11] establish that the privacy profile $\delta(\varepsilon)$ and the trade-off curve $f$ of a mechanism exhibit a convex-conjugate relationship, also known as Fenchel duality. However, Dong et al. [11, Proposition 2.12] confirm this functional equivalence only when $f$ is *symmetric*. With Poisson subsampling at probability $p$, the resulting amplified curve $f_p(x) = pf(x) + (1 - p) \cdot x$ becomes asymmetric.

[2]Dong et al. [11] examine tight composition under the $f$-DP framework by defining an abstract composition operation, denoted $f_1 \otimes f_2$. However, they do not provide an explicit form for this operator for a general trade-off function $f$, offering it only for the specific case of the Gaussian trade-off function $G_\mu$.

[3]Unlike under non-adaptivity, composing two privacy loss random variables $Z_1$ and $Z_2$ does not amount to convolving their privacy loss distributions (PLDs) $f_{Z_1} \circledast f_{Z_2}$ when the mechanisms are adaptive because $Z_1, Z_2$ become dependent. We note that Gopi et al. [18] seem to incorrectly assert their Theorem 5.5 to be valid under adaptivity.

To ensure symmetry, the subsampling result [11, Theorem 4.2] applies a $p$-sampling operator $C_p(f) : \min\{f_p, f_p^{-1}\}^{**}$ that overestimates the $f_p$-curve. We show that this symmetrization step disrupts the equivalence between $f$-DP and privacy profile formalisms (and thereby with other functional notions). To address this, we propose maintaining the natural asymmetry in $f$-DP and avoid the need for this symmetrization step by adopting a convention on the direction of skew. This completes the equivalences across all functional notions of DP.

**Related work:** Our work builds an interpretation of differential privacy by leveraging several works that appeared before. This includes, but not limited to the works on various interpretations of privacy by Dong et al. [11], Dwork and Rothblum [14], Dwork et al. [16], Mironov [23], Bun and Steinke [8], Sommer et al. [29], Gopi et al. [18], Koskela et al. [20], Zhu et al. [32]. In particular, the work of Zhu et al. [32] shares the most similarity with ours, as they were the first to observe that many functional notions of differential privacy appear to be linked via Laplace or Fourier transforms. However, their work centers on using the characteristic function of privacy loss (in (4)) as an intermediate functional representation connecting various DP notions. In contrast, we examine the nature of these connections themselves to harness the perspective of Laplace transformations as an analytical toolkit for differential privacy.

Relevant studies on composition theorems for differential privacy include Dwork et al. [15], Kairouz et al. [19], Murtagh and Vadhan [24], Bun and Steinke [8], Mironov [23], along with numerical accounting methods such as those by Gopi et al. [18], Koskela et al. [20], Doroshenko et al. [12].

**Paper structure:** After reviewing preliminaries on DP and Laplace transforms in Section 2, we present an equivalent description of both the $\delta_{P|Q}(\varepsilon)$ privacy profile and the $\mathrm{R}_q(P\|Q)$-Rényi divergence curve in terms of a set of Laplace transforms of the privacy loss distribution's probability function in Section 3 and use them to connect the two notions. After discussing the implications of these connections, we provide our composition results for privacy profiles and $(\varepsilon, \delta)$-DP point guarantees in Section 4. Finally, in Section 5 we discuss the problem of asymmetry in functional notions and an approach to handling it without breaking equivalences.

## 2 Preliminaries

Here we introduce our notations, provide some background related to Differential Privacy and a short introduction to Laplace transforms.

### 2.1 Background on Differential Privacy

For a data universe $\mathcal{X}$, we consider datasets $D$ of size $n$: $D = (x_0, \dots, x_n) \in \mathcal{X}^n$ and algorithms $\mathcal{M} : \mathcal{X}^n \to \Omega$ that return a random output in space $\Omega$. We assume that $\mathcal{X}$ includes a sentinel element $\perp$, representing an empty entry, to simulate 'add' and 'remove' adjacency within 'replace' adjacency. Two datasets $D$ and $D'$ in $\mathcal{X}^n$ are considered *neighboring* (denoted by $D \simeq D'$) if they differ by a single record replacement. Throughout the paper, we denote the distributions of the output random variables $\mathcal{M}(D)$ and $\mathcal{M}(D')$ as $P$ and $Q$, respectively, and focus our study on the indistinguishability behavior of these two distributions. For simplicity, we use the same symbols $P$ and $Q$ to refer to their probability mass or density functions.

**Definition 2.1** ($(\varepsilon, \delta)$-Differential Privacy and Privacy Profiles [16, 5])**.** *Let* $\varepsilon \geq 0$ *and* $0 \leq \delta \leq 1$*. A randomized algorithm* $\mathcal{M}$ *is* $(\varepsilon, \delta)$-differentially private *(hereon* $(\varepsilon, \delta)$*-DP)*

$$\textit{if} \quad \forall D \simeq D' \quad \textit{and} \quad \forall S \subset \Omega, \quad \mathbb{P}[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \cdot \mathbb{P}[\mathcal{M}(D') \in S] + \delta. \tag{5}$$

*The* privacy profile *of $\mathcal{M}$ on a pair of neighbours $D \simeq D'$ is the function $\delta_{P|Q}(\varepsilon)$ for $\varepsilon \in \mathbb{R}$, where*

$$\delta_{P|Q}(\varepsilon) := \sup_{S \subset \Omega} P(S) - e^{\varepsilon} Q(S) = \int_{\Omega} \max\{0, P(\theta) - e^{\varepsilon} \cdot Q(\theta)\} d\theta. \tag{6}$$

*For any $\varepsilon > 0$, algorithm $\mathcal{M}$* tightly *satisfies $(\varepsilon, \delta)$-DP if $\delta = \sup_{D \simeq D'} \delta_{\mathcal{M}(D)|\mathcal{M}(D')}(\varepsilon)$.*

**Remark 2.1.** *The privacy profile $\delta_{P|Q}(\varepsilon)$ is equivalent to the Hockey-stick divergence $H_{e^{\epsilon}}(P\|Q)$. Our definition of the privacy profile allows $\varepsilon < 0$, in contrast to the original definitions of privacy profile by Balle et al. [5] and of Hockey-stick divergence by Sason and Verdú [28], which consider only non-negative $\varepsilon$. Although permitting $\varepsilon < 0$ might seem counterintuitive, it is both accurate and efficient, as negative values of $\varepsilon$ yield the privacy profile with $P$ and $Q$ reversed.*

$$\delta_{Q|P}(\varepsilon) = \sup_{S \subset \Omega} Q(S) - e^{\varepsilon} P(S) = 1 - e^{\varepsilon} + e^{\varepsilon}[\sup_{S \subset \Omega} P(S^{\complement}) - e^{-\varepsilon} Q(S^{\complement})] = 1 - e^{\varepsilon} + e^{\varepsilon} \delta_{P|Q}(-\varepsilon). \tag{7}$$

**Definition 2.2** (Rényi Differential Privacy [23])**.** *Let $q > 1$ and $\rho \geq 0$. A randomized algorithm $\mathcal{M}$ is $(q, \rho)$-*Rényi differentially private *(henceforth $(q, \rho)$-Rényi DP) if, for all datasets $D \simeq D'$, the $q$-Rényi divergence satisfies $\mathrm{R}_q(\mathcal{M}(D)\|\mathcal{M}(D')) \leq \rho$. For two distributions $P, Q$ over the same space, we define the Rényi divergence $\mathrm{R}_q(P\|Q)$ of any order $q \in \mathrm{ROC}_{P,Q}$ as*

$$\mathrm{R}_q(P\|Q) := \frac{1}{q-1} \log \mathrm{E}_q(P\|Q), \quad \text{where} \quad \mathrm{E}_q(P\|Q) := \int_{\theta \in \Omega} P(\theta)^q Q(\theta)^{1-q} d\theta, \tag{8}$$

*and $\mathrm{ROC}_{P,Q}$ is the region consisting of all orders $q \in \mathbb{C}$ where the integral is* conditionally convergent*.*

**Remark 2.2.** *It is known that Rényi divergence converges to Kullback-Leibler divergence as order $q$ tends to one*[4]*. Among real orders $q \in \mathbb{R}$, privacy researchers typically restrict themselves to $q > 1$ without thinking much about values smaller than 1. Just like privacy profile $\delta_{P|Q}(\varepsilon)$, Rényi divergence for orders $q < 1$ yield the Rényi divergence with $P$ and $Q$ reversed.*

$$e^{(q-1)\cdot \mathrm{R}_q(P\|Q)} = \mathrm{E}_q(P\|Q) = \int_{\theta \in \Omega} P(\theta)^q Q(\theta)^{1-q} d\theta = \mathrm{E}_{1-q}(Q\|P) = e^{-q \cdot \mathrm{R}_{1-q}(Q\|P)}. \tag{9}$$

**Definition 2.3** ($f$-Differential Privacy [11])**.** *Let $f : [0,1] \to [0,1]$ be a convex, continuous, non-increasing function such that $f(x) \leq 1 - x$ for all $x \in [0,1]$. A randomized algorithm $\mathcal{M}$ is* $f$-differentially private *(henceforth $f$-DP) if*

$$\forall \alpha \in [0,1] : f_{\mathcal{M}(D)|\mathcal{M}(D')}(\alpha) \geq f(\alpha), \tag{10}$$

*where for any distributions $P, Q$ on $\Omega$, the $f_{P|Q} : [0,1] \to [0,1]$ is the trade-off function defined as*

$$\forall \alpha \in [0,1] : f_{P|Q}(\alpha) := \inf_{\phi : \Omega \to [0,1]} \{\beta_\phi : \alpha_\phi \leq \alpha\}, \tag{11}$$

*where $\alpha_\phi := \mathbb{E}_P[\phi]$, and $\beta_\phi := 1 - \mathbb{E}_Q[\phi]$.*

**Remark 2.3.** *For a hypothesis test $\phi$ on a sample $\theta$ originating from either $P$ or $Q$, we follow the convention that $\theta \sim Q$ represents the positive event, and $\phi(\theta) = 1$ denotes a positive prediction. With this convention, $\alpha_\phi$ and $\beta_\phi$ represent the* false positive rate *(Type I error) and* false negative rate *(Type II error), respectively. Note that the left-continuous inverse of the trade-off function $f^{-1}_{P|Q}(\beta)$ gives the trade-off curve with $P$ and $Q$ reversed.*

$$f^{-1}_{P|Q}(\beta) := \{\inf \alpha \in [0,1] : f_{P|Q}(\alpha) \leq \beta\} = \inf_{\phi : \Omega \to [0,1]} \{\alpha_\phi : \beta_\phi \leq \beta\} = f_{Q|P}(\beta). \tag{12}$$

[4]It follows from the *replica trick*: $\mathbb{E}[\log X] = \lim_{n \to 0} \frac{1}{n} \log \mathbb{E}[X^n]$, when $\log X$ is the privacy loss rv $Z \leftarrow \mathrm{PLD}(P\|Q)$.

**Theorem 2.4** (Privacy of Gaussian Mechanism [16, 3, 23]). *If $P = \mathcal{N}(\mu, \sigma^2 I_d)$ and $Q = \mathcal{N}(\mu', \sigma^2 I_d)$ are two multivariate Gaussian distributions on $\mathbb{R}^d$ such that $\kappa = \|\mu - \mu'\|_2^2 / 2\sigma^2$, then $\mathrm{R}_q(P\|Q) = \kappa q$ for all $q > 1$ and for all $\varepsilon \in \mathbb{R}$* [5]*,*

$$\delta_{P|Q}(\varepsilon) = \overline{\Phi}\left(\frac{\varepsilon - \kappa}{\sqrt{2\kappa}}\right) - e^{\varepsilon}\overline{\Phi}\left(\frac{\varepsilon + \kappa}{\sqrt{2\kappa}}\right) = O(e^{-\varepsilon^2/4\kappa}), \text{ where } \overline{\Phi}(t) = \underset{G \sim \mathcal{N}(0,1)}{\mathbb{P}}[G > t]. \tag{13}$$

## 2.2 Background on Laplace Transforms

Laplace transform is an integral transform that maps time domain functions with real arguments ($t \in \mathbb{R}$) to frequency domain functions with complex arguments ($s \in \mathbb{C}$). The one-sided and two-sided Laplace transformations of a function $g(t)$ at complex frequency $s$ is defined respectively as

$$\mathcal{L}\{g(t)\}(s) := \int_{0^+}^{\infty} e^{-st} g(t) dt, \quad \text{and} \quad \mathcal{B}\{g(t)\}(s) := \int_{-\infty}^{\infty} e^{-st} g(t) dt, \tag{14}$$

where $0^+$ is the shorthand notation for limit approaching 0 from positive side. We can express two-sided Laplace transforms using one-sided transform as

$$\mathcal{B}\{f(t)\}(s) = \mathcal{L}\{f(t)\}(s) + \int_{0^-}^{0^+} f(t) dt + \mathcal{L}\{f(-t)\}(-s), \tag{15}$$

where $\int_{0^-}^{0^+} f(t)dt$ may not be 0 if $f$ has an *impulse* (aka. *integrable singularity*) at 0, informally defined to be an infinitely dense point but with a finite mass.

**Remark 2.5.** *Conventionally, one-sided Laplace transform is defined to include point mass located at* 0 *entirely (i.e., integration is from $0^-$ instead of $0^+$). For aligning with the conventions in differential privacy, our definition entirely excludes the point mass located at* 0 *by convention. The Laplace transform properties presented in this paper are adjusted to reflect the same.*

The values of $s \in \mathbb{C}$ for which the integrals in (14) converges conditionally[6] is referred to as the *region of (conditional) convergence* (ROC) of the respective transforms. This region is always a strip parallel to the imaginary axis $\omega i$ where $i = \sqrt{-1}$ and $\omega \in \mathbb{R}$, which follows from dominated convergence theorem [25]. We denote the region of convergence for a function $g$ as $\mathrm{ROC}_{\mathcal{L}}\{g\}$ in case of one-sided Laplace transform and as $\mathrm{ROC}_{\mathcal{B}}\{g\}$ for two-sided Laplace transforms.

**Uniqueness and Inversion.** Laplace transforms are unique [10]: if two continuous functions $g$ and $h$ share the same Laplace transform, then, $g(t) = h(t)$ holds for all $t \in \mathbb{R}^+$ in the case of a one-sided Laplace transform and for all $t \in \mathbb{R}$ for a two-sided Laplace transform. As a consequence of uniqueness, Laplace transform $\bar{g}(s) = \mathcal{L}\{g(t)\}(s)$ can be inverted to get back $g(t)$ by applying an *Inverse Laplace transform* [26], defined as

$$g(t) = \mathcal{L}^{-1}\{\bar{g}(s)\}(t) := \frac{1}{2\pi i} \lim_{\omega \to \infty} \int_{\gamma - i\omega}^{\gamma + i\omega} e^{st} \bar{g}(s) ds = \frac{1}{2\pi i} \int_{\gamma - i\infty}^{\gamma + i\infty} e^{st} \bar{g}(s) ds, \tag{16}$$

where the integral is taken over the line consisting of all points $s$ with $\mathfrak{Re}(s) = \gamma$ for any $\gamma$ lying in the ROC of $\bar{g}(s)$. The formula (16) also inverts two-sided Laplace transform $\mathcal{B}\{g(t)\}$, as long as we choose $\gamma$ in the ROC of the two-sided transform [25].

[5] While the original result by Balle and Wang [3] was stated only for non-negative $\varepsilon$, the expression remains identical for $\varepsilon < 0$ on invoking (7), thanks to the symmetry property $\overline{\Phi}(-x) = 1 - \overline{\Phi}(x)$ of normal distribution.

[6] Conditional convergence for $\mathcal{L}\{g\}$ at $s \in \mathbb{C}$ means that the limit $\lim_{\gamma \to \infty} \int_{0^+}^{\gamma} e^{-st} g(t)\, dt$ exists. Similarly, the limit $\lim_{\gamma \to \infty} \int_{-\gamma}^{\gamma} e^{-st} g(t)\, dt$ should exist for $\mathcal{B}\{g\}$ to be conditionally convergent at $s$.

**Remark 2.6.** *Uniqueness of Laplace transforms extends to discontinuous functions as well. If $g$ and $h$ are continuous* almost everywhere*, i.e. the set where either isn't continuous has a total Lebesgue measure of zero, then $g(t) = h(t)$ almost everywhere in the respective time-domains. In such cases, it can be shown [17] that the inversion formula* (16) *gives*

$$\frac{1}{2\pi i}\int_{\gamma-i\infty}^{\gamma+i\infty} e^{st}\bar{g}(s)ds = \frac{1}{2}[g(t^-) + g(t^+)]. \tag{17}$$

**Relation to Fourier transform.** The Fourier transform $G(\omega)$ of a function $g$ is defined as

$$G(\omega) = \int_{-\infty}^{\infty} e^{-i2\pi\omega t} g(t)dt, \tag{18}$$

which is the same as the two-sided laplace transform $\mathcal{B}\{g(t)\}(s)$ for a purely imaginary $s = i2\pi\omega$. As such, Fourier transform is seen as a *special case* of Laplace transforms.

**Properties of Laplace transforms.** The Laplace transformation is a very useful tool because a lot of operations in the time-domain correspond to simpler operations in the frequency domain and vice versa. For a detailed exposition on these properties, refer to Cohen [10] for one-sided Laplace transform and to Oppenhiem et al. [25] for the two-sided counterpart. In Appendix A.1, we provide a table summarizing all the properties that we rely on in this paper. We reference properties of Table 1 throughout the paper using the notation $\overset{(m)}{=}$, where $(m)$ is the equation number of the used property.

## 3 Laplace Transform Expressions of Differential Privacy

Differential privacy bounds the maximum divergence in the output distribution caused by including or omitting a data-point from the dataset. This principle is mirrored in the notion of *privacy loss distribution* which expresses how much an algorithm's output reveals about the inclusion of a specific data-point.

**Definition 3.1** (Privacy Loss Distribution [29])**.** *The* privacy loss *of an observation $\theta \in \Omega$ from an algorithm $\mathcal{M}$, when comparing datasets $D \simeq D'$, is defined as $L_{P|Q}(\theta) := \log(P(\theta)/Q(\theta))$, where $P$ and $Q$ are the probability mass/density functions of $\mathcal{M}(D)$ and $\mathcal{M}(D')$, respectively. The* privacy loss distribution $\mathrm{PLD}(P\|Q)$ *is the distribution of $L_{P|Q}(\Theta)$ when $\Theta \sim P$.*

**Remark 3.1.** *Similar to other functional privacy notions, the PLD formalism possesses a reversal property [18, Definition 3.1]. Let $f_Z$ and $f_{Z'}$ represent the generalized density functions of the random variables $Z \leftarrow \mathrm{PLD}(P\|Q)$ and $Z' \leftarrow \mathrm{PLD}(Q\|P)$, respectively. Then*

$$\forall z \in \mathbb{R}\ : f_Z(z) = e^z \cdot f_{Z'}(-z). \tag{19}$$

The $\mathrm{PLD}(P\|Q)$ describes how outputs arising from $D$ increase an observer's confidence that they did not come from $D'$. Many prior works make use of the following set of DP expressions in terms of the privacy loss distribution [29, 6, 5, 20, 18, 30, 9]. The privacy profile $\delta_{P|Q}(\varepsilon)$ is expressed as

$$\delta_{P|Q}(\varepsilon) = \mathop{\mathbb{E}}_{\mathrm{PLD}(P\|Q)}\left[1 - e^{\varepsilon - Z}\right]_+ = \mathop{\mathbb{E}}_{\mathrm{PLD}(Q\|P)}\left[e^{-Z'} - e^{\varepsilon}\right]_+ = \Pr_{\mathrm{PLD}(P\|Q)}[Z > \varepsilon] - e^{\varepsilon} \cdot \Pr_{\mathrm{PLD}(Q\|P)}[Z' < -\varepsilon], \tag{20}$$

and the Rényi divergence $\mathrm{R}_q(P\|Q)$ is expressed as

$$\mathrm{R}_q(P\|Q) = \frac{1}{q-1}\log \mathop{\mathbb{E}}_{\mathrm{PLD}(P\|Q)}\left[e^{(q-1)\cdot Z}\right] = \frac{1}{q-1}\log \mathop{\mathbb{E}}_{\mathrm{PLD}(Q\|P)}\left[e^{-q\cdot Z'}\right]. \tag{21}$$

In the following theorem, we present a more dynamic version of these relationships by expressing them as a set of *Laplace transforms* of the privacy loss distributions $\mathrm{PLD}(P\|Q)$ and $\mathrm{PLD}(Q\|P)$.

**Theorem 3.2.** *For a random variable $X$, let $F_X(t) := \Pr[X \leq t]$ denote its* cumulative distribution function *and $f_X(t)$ denote its* generalized probability density function[7]. *Let $P$ and $Q$ be probability distributions and $Z \sim \mathrm{PLD}(P\|Q)$ and $Z' \sim \mathrm{PLD}(Q\|P)$ denote their privacy loss random variables. If $Z \sim \mathrm{PLD}(P\|Q)$ and $Z' \sim \mathrm{PLD}(Q\|P)$, then for all $\varepsilon \in \mathbb{R}$,*

$$\begin{aligned}
\delta_{P|Q}(\varepsilon) &= \mathcal{L}\{1 - F_Z(t+\varepsilon)\}(1) && (22)\\
&= e^{\varepsilon}\cdot\mathcal{L}\{F_{Z'}(-t-\varepsilon)\}(-1) && (23)\\
&= e^{\varepsilon}\cdot\mathcal{L}\{f_{Z'}(-t-\varepsilon)\}(-1) - \mathcal{L}\{f_Z(t+\varepsilon)\}(1) && (24)\\
&= \mathcal{L}\{f_Z(t+\varepsilon)\}(0) - e^{\varepsilon}\cdot\mathcal{L}\{f_{Z'}(-t-\varepsilon)\}(0). && (25)
\end{aligned}$$

*And, for all $q \in \mathrm{ROC}_{\mathcal{B}}\{f_{Z'}\}$ (or equivalently, $1-q \in \mathrm{ROC}_{\mathcal{B}}\{f_Z\}$),*

$$e^{(q-1)\cdot\mathrm{R}_q(P\|Q)} = \mathrm{E}_q(P\|Q) = \mathcal{B}\{f_Z(t)\}(1-q) = \mathcal{B}\{f_{Z'}(t)\}(q). \tag{26}$$

Laplace expressions in Theorem 3.2 often arise in their explicit integral forms within several proofs in related works on differential privacy, for instance [9, Lemma 9], [30, Proposition 7], and [3, Theorem 5]. In their integral forms, they frequently undergo manipulations like integration-by-parts or change-of-variables which can quickly get complicated. Our Theorem 3.2 offers a way to simplify the complexity of such manipulations as one can express the concerned terms in their Laplace expressions and invoke its properties from Table 1, like (69) and (71) for shifting or scaling variable of integration and (73) and (74) for integrating-by-parts. Examples in this paper will illustrate that reasoning about privacy this way through its Laplace transform interpretation could be quite effective.

In the following theorem, we show that the Rényi divergence $\mathrm{R}_q(P\|Q)$ and the privacy profile $\delta_{P|Q}(\varepsilon)$ are connected through a Laplace transform as well. We can show this using only the expressions in Theorem 3.2.

**Theorem 3.3** (Rényi DP from privacy profile)**.** *Let $q > 1$. For any two distributions $P$ and $Q$,*

$$e^{(q-1)\cdot\mathrm{R}_q(P\|Q)} = \mathrm{E}_q(P\|Q) = q(q-1)\cdot\mathcal{B}\left\{\delta_{P|Q}(t)\right\}(1-q), \tag{27}$$

*for all orders $q$ such that $1-q \in \mathrm{ROC}_{\mathcal{B}}\{\delta_{P|Q}\}$.*

*Proof.* Let $Z \sim \mathrm{PLD}(P\|Q)$ and $Z' \sim \mathrm{PLD}(Q\|P)$. From (25),

$$\begin{aligned}
\delta_{P|Q}(\varepsilon) &= \mathcal{L}\{f_Z(t+\varepsilon)\}(0) - e^{\varepsilon}\cdot\mathcal{L}\{f_{Z'}(-t-\varepsilon)\}(0) && (28)\\
&= \int_{0^+}^{\infty} e^0\cdot f_Z(t+\varepsilon)\mathrm{d}t - e^{\varepsilon}\cdot\int_{0^+}^{\infty} e^0\cdot f_{Z'}(-t-\varepsilon)\mathrm{d}t && (29)\\
&= 1 - F_Z(\varepsilon) - e^{\varepsilon}\cdot F_{Z'}(-\varepsilon). && (30)
\end{aligned}$$

[7] We define density as $f_X(t)dt = \lim_{a\to 0^+}\int_{t-a}^{t+a} F_X(u)du$ to handle cases where $F_X$ isn't differentiable everywhere, such as when PLD is a discrete distribution. This density is expressible with *Dirac delta* $\triangle(t)$ as $f_X(t) = \begin{cases} \dot{F}_X(t) & \text{if derivative } \dot{F}_X \text{ exists at } t \\ [F_X(t^+) - F_X(t^-)]\triangle(t) & \text{otherwise} \end{cases}$, and satisfies $F_X(t) = \int_{-\infty}^{t^+} f_X(u)du$.

We apply the linearity, time-shifting, and reversal properties of two-sided Laplace transforms to simplify the Laplace transform of privacy profile as follows

$$\mathcal{B}\left\{\delta_{P|Q}(t)\right\}(1-q) = \mathcal{B}\left\{1 - F_Z(t) - e^t \cdot F_{Z'}(-t)\right\}(1-q) \tag{31}$$

$$\overset{(68)}{=} \mathcal{B}\left\{1 - F_Z(t)\right\}(1-q) - \mathcal{B}\left\{e^t \cdot F_{Z'}(-t)\right\}(1-q) \tag{32}$$

$$\overset{(70)}{=} \mathcal{B}\left\{1 - F_Z(t)\right\}(1-q) - \mathcal{B}\left\{F_{Z'}(-t)\right\}(-q) \tag{33}$$

$$\overset{(72)}{=} \mathcal{B}\left\{1 - F_Z(t)\right\}(1-q) - \mathcal{B}\left\{F_{Z'}(t)\right\}(q). \tag{34}$$

Next, we apply the derivative property of Laplace transforms to get

$$\mathcal{B}\left\{1 - F_Z(t)\right\}(1-q) \overset{(73)}{=} \frac{\mathcal{B}\left\{f_Z(t)\right\}(1-q)}{q-1} \quad \text{and} \quad \mathcal{B}\left\{F_{Z'}(t)\right\}(q) \overset{(73)}{=} \frac{\mathcal{B}\left\{f_{Z'}\right\}(q)}{q}. \tag{35}$$

Finally, noting from Theorem 3.2 and (8) that $\mathcal{B}\left\{f_Z\right\}(1-q) = \mathcal{B}\left\{f_{Z'}\right\}(q) = \mathrm{E}_q\left(P\|Q\right)$, we have

$$\mathcal{B}\left\{\delta_{P|Q}(t)\right\}(1-q) = \mathrm{E}_q\left(P\|Q\right)\left[\frac{1}{q-1} - \frac{1}{q}\right] = \frac{e^{(q-1)\mathrm{R}_q(P\|Q)}}{q(q-1)}. \tag{36}$$

□

The privacy profile $\delta_{P|Q}$, for any pair of distributions $P, Q$, is a continuous function[8]. Therefore, by Lerch's theorem (cf. discussion on uniqueness in Section 2.2), *any two privacy profiles share the same Rényi divergence curve if and only if they are identical.* In other words, a privacy profile $\delta_{P|Q}$ is equivalent to its Rényi divergence curve $\mathrm{R}_q\left(P\|Q\right)$, provided the Laplace transform $\mathcal{B}\left\{\delta_{P|Q}\right\}(1-q)$ exists at some $q \in \mathbb{C}$, i.e., $\mathrm{ROC}_{\mathcal{B}}\{\delta_{P|Q}\} \neq \emptyset$.

Note that Theorem 3.3 establishes a connection between Rényi divergence and privacy profile that applies to *all distributions $P$ and $Q$*, whether $P$ is absolutely continuous[9] (denoted as $P \ll Q$) with respect to $Q$ or not. The choice of $P, Q$ however influences the region of convergence $\mathrm{ROC}_{\mathcal{B}}\{\delta_{P|Q}\}$ where the Rényi divergence can be defined. One can verify that if $P \ll Q$, the Laplace transform $\mathcal{B}\left\{\delta_{P|Q}\right\}(1-q)$ converges for all real $q > 1$; and if $Q \ll P$, it converges for all $q < 1$, except for $q = 0$. At $q = 0$ or $q = 1$, the transform does not converge as the expression for $\delta_{P|Q}(\varepsilon)$ has singularities at these points because numerator $\mathrm{E}_q\left(P\|Q\right) = \int P^q Q^{1-q} \mathrm{d}\theta = 1$ when $q = 0$ or 1 while denominator becomes zero. Since region of convergence is always a strip in the complex plane $\mathbb{C}$ parallel to the imaginary line, the Rényi divergence exists not just for real orders $q$, but also for imaginary orders: $\{q \in \mathbb{C} \,:\, \mathfrak{Re}(q) > 1\}$ when $P \ll Q$ and $\{q \in \mathbb{C} \,:\, \mathfrak{Re}(q) < 1 \text{ and } \mathfrak{Re}(q) \neq 0\}$ when $Q \ll P$. Therefore, as long as either $P \ll Q$ *or* $Q \ll P$, the region $\mathrm{ROC}_{\mathcal{B}}\{\delta_{P|Q}\}$ is not empty and a characterizing curve $\mathrm{R}_q\left(P\|Q\right)$ exists for the profile $\delta_{P|Q}(\varepsilon)$. That is to say, Theorem 3.3 can be used to derive the exact privacy profile $\delta_{P|Q}$ using the Rényi divergence $\mathrm{R}_q\left(P\|Q\right)$ function. We can do this by substituting $1 - q = s$ in (27), rearranging, and applying the *inverse Laplace transform* (16), to get the following explicit form:

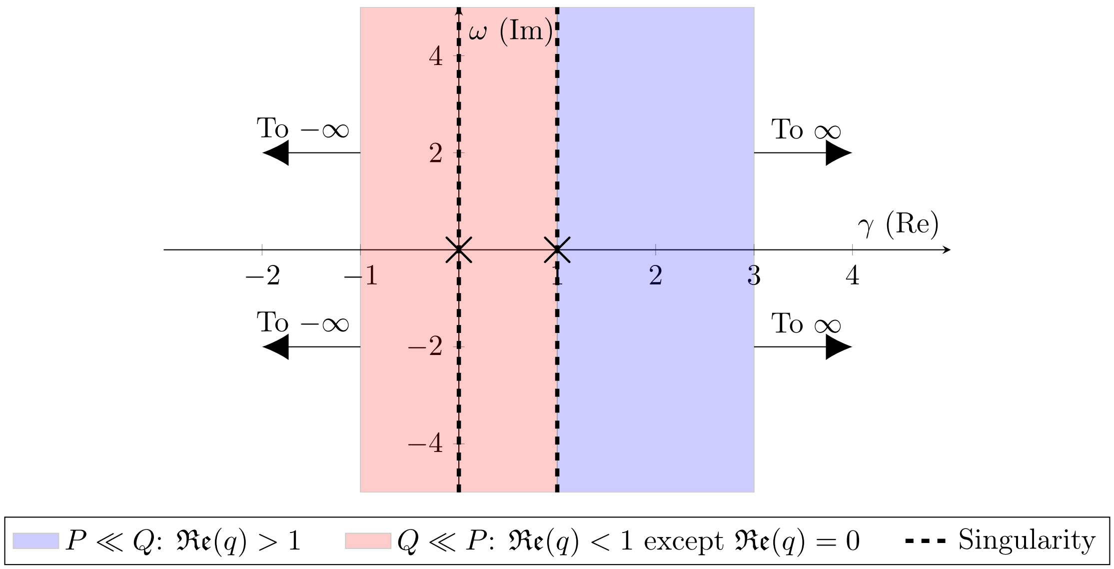
$$\delta_{P|Q}(\varepsilon) = \mathcal{L}^{-1}\left\{\frac{e^{-s\mathrm{R}_{1-s}(P\|Q)}}{s(s-1)}\right\}(\varepsilon) = \frac{1}{2\pi i}\lim_{\omega\to\infty}\int_{\gamma-i\omega}^{\gamma+i\omega} e^{s\varepsilon} \cdot \frac{e^{-s\mathrm{R}_{1-s}(P\|Q)}}{s(s-1)} ds, \tag{37}$$

where $\gamma \in \mathbb{R}$ can be *any* real point in $\mathrm{ROC}_{\mathcal{B}}\{\delta_{P|Q}\}$.

[8] We can write $\delta_{P|Q}(\varepsilon) = \sup_S \psi_S(\varepsilon)$ where $\psi_S(\varepsilon) := P(S) - e^{\varepsilon} \cdot Q(S)$ is a continuous decreasing function. Therefore, the supremum of $\phi_S$ over all $S$ is also a continuous decreasing function.

[9] $P$ is absolutely continuous with respect to $Q$ if for all measurable subsets $S \subset \Omega$, $P(S) > 0 \implies Q(S) > 0$. We say that distribution pair $(P, Q)$ is absolutely continuous (w.r.t. each other) if $P \ll Q$ and $Q \ll P$.

Region of Convergence of $\mathcal{B}\left\{\delta_{P|Q}\right\}(1-q)$ for $P \ll Q$ and $Q \ll P$ at Rényi order $q = \gamma + i\omega$




**Case study.** The following example demonstrates an application of the Laplace transform identity in (3.3). We first describe the privacy profile of the randomized response mechanism [19] and then use (27) on it to reason about its Rényi divergence characteristics.

**Theorem 3.4** (Privacy profile of randomized response)**.** *Fix $\varepsilon > 0$ and $\delta \in [0,1]$. Let $\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta} : \{0,1\} \to \{0,1\} \times \{\perp, \top\}$ be the randomized response mechanism, which has the following output probabilities.*

$$\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(0) = \begin{cases} (0,\perp) & \textit{with probability } \delta, \\ (0,\top) & \textit{with probability } \frac{(1-\delta)e^{\varepsilon}}{e^{\varepsilon}+1}, \\ (1,\top) & \textit{with probability } \frac{(1-\delta)}{e^{\varepsilon}+1}, \\ (1,\perp) & \textit{with probability } 0, \end{cases} \quad \mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(1) = \begin{cases} (0,\perp) & \textit{with probability } 0, \\ (0,\top) & \textit{with probability } \frac{(1-\delta)}{e^{\varepsilon}+1}, \\ (1,\top) & \textit{with probability } \frac{(1-\delta)e^{\varepsilon}}{e^{\varepsilon}+1}, \\ (1,\perp) & \textit{with probability } \delta. \end{cases} \tag{38}$$

*For $P = \mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(0)$ and $Q = \mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(1)$, the privacy profiles are*

$$\forall t \in \mathbb{R} \ : \ \delta_{\mathrm{RR}} := \delta_{P|Q}(t) = \delta_{Q|P}(t) = \begin{cases} \delta & \textit{if } \varepsilon < t, \\ 1 - \frac{e^t+1}{e^{\varepsilon}+1}(1-\delta) & \textit{if } -\varepsilon < t \leq \varepsilon, \\ 1 - e^t(1-\delta) & \textit{if } t \leq -\varepsilon. \end{cases} \tag{39}$$

Theorem 3.4 extends Balle et al. [5, Theorem 2], which provides the privacy profile for randomized response only in the case where $\delta = 0$ and for values $t \geq 0$. From (38), we can see that neither $P \ll Q$ nor $Q \ll P$ holds for the output distributions of randomized response mechanism when $\delta > 0$. As such, the Laplace transform $\mathcal{B}\left\{\delta_{P|Q}\right\}(1-q)$ must not converge for any $q \in \mathbb{R}$. One can check that when $q \geq 1$, the transform has a term $\int_{\varepsilon}^{\infty} e^{(q-1)t} \cdot \delta \mathrm{d}t$ which blows up to $\infty$, and when $q < 1$, the transform has a term $\int_{-\infty}^{-\varepsilon} e^{(q-1)t} \cdot (1 - e^t(1-\delta))\mathrm{d}t$ that blows up to $-\infty$. Hence, the Rényi divergence of the privacy profile (39) cannot be defined for any order $q \in \mathbb{R}$ when $\delta > 0$.

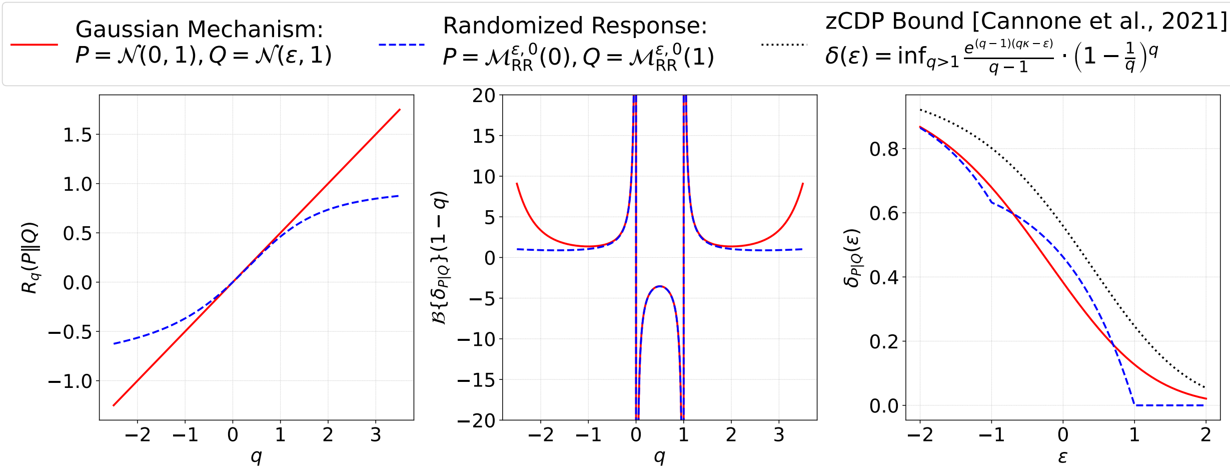



Figure 1: Comparison between the indistinguishability characteristic of Gaussian mechanism of Theorem 2.4 ($\kappa = \|\mu - \mu'\|^2/2\sigma^2 = \varepsilon^2/2$) and randomized response mechanism defined in Theorem 3.4 (with $\delta = 0$). This figure visualizes the singularities at $q = 0$ and $q = 1$ that exists for the Laplace transform $\mathcal{B}\left\{\delta_{P|Q}\right\}(1-q)$ disappears for Rényi divergence $\mathrm{R}_q(P\|Q)$, which is an effect of the replica trick unfolding. This figure also demonstrates that neither the dominance $\mathrm{R}_q(P_1\|Q_1) \leq \mathrm{R}_q(P_2\|Q_2)$ for all $q > 1$, nor the dominance $\mathcal{B}\left\{\delta_{P_1|Q_1}\right\}(1-q) \leq \mathcal{B}\left\{\delta_{P_2|Q_2}\right\}(1-q)$ for all $q \in \mathbb{R}\setminus\{0,1\}$ is enough to bound $\delta_{P_1|Q_1}(\varepsilon) \leq \delta_{P_2|Q_2}(\varepsilon)$ at all $\varepsilon \in \mathbb{R}$. Additionally, the black dotted line in the rightmost plot shows that even the tightest[10] conversion on the Rényi curve considering only real orders $q > 1$ fails to characterize its own privacy profile.

When $\delta = 0$, note that both $P \ll Q$ and $Q \ll P$. Therefore, the Laplace transform $\delta_{P|Q}(\varepsilon)$ must exist for all $q \in \mathbb{R} \setminus \{0,1\}$. The following theorem shows the resulting Rényi divergence curve, derived by computing the Laplace transform.

**Theorem 3.5** (Rényi DP of $(\varepsilon,0)$-Randomized Response)**.** *For any $\varepsilon > 0$ and $\delta = 0$, the output distributions of randomized response mechanism in Theorem 3.4 exhibit a Rényi divergence*

$$\forall q \in \mathbb{C} \text{ s.t. } \mathfrak{Re}(q) \notin \{0,1\} : \mathrm{R}_q(P\|Q) = \frac{1}{q-1} \log\left(\frac{e^{\varepsilon}}{1+e^{\varepsilon}} e^{-q\varepsilon} + \frac{1}{1+e^{\varepsilon}} e^{q\varepsilon}\right). \tag{40}$$

Theorem 3.5 generalizes Mironov [23, Proposition 5], which gives the Rényi divergence of randomized response only for real orders $q > 1$. In the following section, we elaborate on the significance of complex orders in Rényi divergence. Figure 1 visualizes the privacy profile $\delta_{P|Q}$, its Laplace transform $\mathcal{B}\left\{\delta_{P|Q}\right\}$, and the corresponding Rényi divergence $\mathrm{R}_q(P\|Q)$ of this randomized response mechanism, and compares it with that of Gaussian mechanism (cf. Theorem 2.4).

### 3.1 Dominance: Rényi Divergence vs. Privacy Profile

Differential privacy is a study of distributional divergence between output distributions $P$ and $Q$ not just for a pair of neighboring inputs $D, D'$ but across *all* neighboring inputs. When considering functional notions of DP, comparing indistinguishability characteristics of two output-distribution pairs, say $(P_1, Q_1)$ and $(P_2, Q_2)$, requires a notion of *dominance*. Zhu et al. [32] define a *dominating*

[10] The tightest conversion by Asoodeh et al. [2] lacks a closed-form expression and is challenging to approximate numerically in a stable way. Therefore, we compare with Canonne et al. [9, Corollary 13], which yields similar values.

*pair of distribution* specific to a mechanism $\mathcal{M}$ as a pair of distributions $P, Q$ such that

$$\forall \varepsilon \in \mathbb{R} \ : \ \sup_{D \simeq D'} \delta_{\mathcal{M}(D)|\mathcal{M}(D')}(\varepsilon) \leq \delta_{P|Q}(\varepsilon). \tag{41}$$

Following the definition by Zhu et al. [32], we define the following notions of dominance.

**Definition 3.2** (Dominance for Distribution Pairs)**.** *We say that the distribution pair $(P_2, Q_2)$ dominates $(P_1, Q_1)$ in privacy profile (denote as $(P_1, Q_1) \preceq_\delta (P_2, Q_2)$) if*

$$\forall \varepsilon \in \mathbb{R} \ : \ \delta_{P_1|Q_1}(\varepsilon) \leq \delta_{P_2|Q_2}(\varepsilon). \tag{42}$$

*And, we say $(P_2, Q_2)$ dominates $(P_1, Q_1)$ in Rényi divergence (denote as $(P_1, Q_1) \preceq_R (P_2, Q_2)$) if*

$$\forall q > 1 \ : \mathrm{R}_q \left(P_1 \| Q_1\right) \leq \mathrm{R}_q \left(P_2 \| Q_2\right). \tag{43}$$

The following theorem reveals the surprising fact that, while the Rényi divergence curve and privacy profile curve are equivalent for all absolutely continuous distribution pairs, this equivalence does not imply an identical dominance ordering across absolutely continuous distribution pairs.

**Theorem 3.6.** *Consider two absolutely continuous distribution pairs $(P_1, Q_1)$ and $(P_2, Q_2)$. Then*

$$(P_1, Q_1) \preceq_\delta (P_2, Q_2) \implies (P_1, Q_1) \preceq_R (P_2, Q_2). \tag{44}$$

*However, the opposite direction does not hold:*

$$(P_1, Q_1) \preceq_R (P_2, Q_2) \not\Longrightarrow (P_1, Q_1) \preceq_\delta (P_2, Q_2). \tag{45}$$

*Proof.* First part directly follows from (26) by noting that

$$\mathcal{B}\left\{\delta_{P_1|Q_1}\right\}(1-q) = \int_{-\infty}^{\infty} e^{-(1-q)t} \cdot \delta_{P_1|Q_1}(t) \mathrm{d}t \leq \int_{-\infty}^{\infty} e^{-(1-q)t} \cdot \delta_{P_2|Q_2}(t) \mathrm{d}t = \mathcal{B}\left\{\delta_{P_2|Q_2}\right\}(1-q) \tag{46}$$

and that for $q > 1$, $q(q-1) > 0$.

The proof of the second statement is based on the example by Zhu et al. [32] comparing the Rényi DP curve and privacy profile of Gaussian mechanism described in Theorem 2.4 (denote with $(P_2, Q_2)$) with that of randomized response mechanism, defined in Theorem 3.4 (denote with $(P_1, Q_1)$). For any $\varepsilon > 0$, we set $\kappa = \varepsilon^2/2$ in the Gaussian mechanism where $(\varepsilon, 0)$ is the parameter of the randomized response and compare them in Figure 1. From the leftmost plot, we can see that $(P_1, Q_1) \preceq_R (P_2, Q_2)$ holds.[11] However, note that in the rightmost plot that their privacy profiles cross one another. Hence, the dominance $(P_1, Q_1) \preceq_\delta (P_2, Q_2)$ does not hold. □

Zhu et al. [32] point out that it is troubling that identifying a dominating pair of distributions in Rényi divergence does not guarantee a corresponding dominance in privacy profiles. Although the Rényi divergence curve $\mathrm{R}_q(P\|Q)$ is an equivalent representation of the privacy profile $\delta_{P|Q}(\varepsilon)$, converting the $\mathrm{R}_q(P\|Q)$ curve into a tight upper bound for privacy profile will *always* introduce a gap. This gap can be substantial, as shown in Figure 1, where we compare the (nearly) tight upper bound (black dotted line) derived for the privacy profile from the Rényi divergence curve (red line on the left) of the Gaussian mechanism with the actual privacy profile of this Gaussian mechanism.

[11] Dong et al. [11, Proposition B.7 (a)] show that $\mathrm{R}_q\left(\mathcal{M}_{\mathrm{RR}}^{\varepsilon,0}(0) \middle\| \mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(1)\right) \leq \mathrm{R}_q\left(\mathcal{N}(0,1) \| \mathcal{N}(\varepsilon,1)\right)$ for all $q > 1$, but the bound actually holds for all $q > 0$.

**Role of complex orders.** While recognizing the fundamental gap between the tightest achievable bound on the privacy profile derived from a bound on the Rényi divergence curve and the privacy profile itself, no such tight bound has yet been established. The best available bounds rely on taking an infimum over pointwise conversions from Rényi divergence at a single order $q > 1$ to $(\varepsilon, \delta)$-DP [9, 8]. However, since pointwise guarantees offer a lossy characterization of indistinguishability, taking an infimum over privacy profiles implied by these guarantees is unlikely to achieve tight upper bounds. Notably, from the Inverse Laplace Transform expression (37) for privacy profiles, we observe that the value of $\delta_{P|Q}$ at any $\epsilon$ depends on the behavior of the Rényi divergence $\mathrm{R}_q(P\|Q)$ along the complex line $\mathfrak{Re}(1-q) = \gamma$ for any $\gamma \in \mathbb{R} \setminus \{0, 1\}$, and not along the real line. In fact, the choice of $\gamma$ on $\mathbb{R}$ does not matter at all, as long as it lies in the $\mathrm{ROC}_{\mathcal{B}}\{\delta_{P|Q}\}$. Thus, we believe that establishing a tight functional conversion will require examining how the Rényi divergence *curls as we move the order $q$ along the complex line $\gamma + i\omega$.*

# 4 Exactly-Tight Composition Theorems

Unlike functional DP guarantees, point guarantees like $(\varepsilon, \delta)$-DP or $(q, \rho)$-Rényi DP are a lossy characterization of the indistinguishability between two distributions $P$ and $Q$. Despite this, using them for reporting or certifying an algorithm's worst-case privacy is generally acceptable as the privacy protection they guarantee, although conservative, is adequate. The main issue arises when attempting to compose point DP guarantees, as the quantification loss often compounds drastically, resulting in a significant overestimation of the actual privacy protection provided by an algorithm. Consider the $k$-fold (non-adaptive) composition of a one-dimensional Gaussian mechanism with $L_2$ sensitivity 1 and noise variance $\sigma^2 = 1$. For individual 1D Gaussian distributions $P = \mathcal{N}(0, 1)$ and $Q = \mathcal{N}(1, 1)$, the privacy profile is of the order $\delta_{P|Q}(\varepsilon) = O(e^{-\varepsilon^2/2})$ (cf. Theorem 2.4), indicating that the mechanism satisfies a point guarantee of $(O(\sqrt{\log(1/\delta)}), \delta)$-DP for any $\delta \in (0, 1]$. When we extend this to $k$-fold non-adaptive self-composition, the resulting output distribution is a $k$-dimensional Gaussian, specifically $P^{\otimes k} = \mathcal{N}(\vec{0}, I_k)$ and $Q^{\otimes k} = \mathcal{N}(\vec{1}, I_k)$. The privacy profile of this composition is of the asymptotic order $\delta_{P^{\otimes k}|Q^{\otimes k}}(\varepsilon) = O(e^{-\varepsilon^2/2k})$, yielding a point guarantee of $(O(\sqrt{k\log(1/\delta)}), \delta)$-DP. However, if we attempt to compose the individual $(O(\sqrt{\log(1/\delta)}), \delta)$-DP point guarantees for each Gaussian mechanism, even with the optimal composition theorem for $(\varepsilon, \delta)$-DP point guarantees from Kairouz et al. [19], the best achievable guarantee is $(O(\sqrt{k}\log(1/\delta)), (k+1)\delta)$-DP. This result is significantly more pessimistic than the true privacy profile—a factor of $O(\sqrt{\log(k/\delta)})$ in the $\varepsilon$ for the same $\delta$.

Functional notions of DP effectively address this problem by capturing the indistinguishability between any two distributions $P$ and $Q$ accurately [11, 20]. However, Rényi DP (as a function of order $q$) remains the only functional DP notion with an *exactly tight* composition theorem (i.e., matching even the constants) that can accommodate *adaptively chosen heterogeneous mechanisms.*[12]

**Remark 4.1.** *The previous statement requires some justifications. We note that the PLD formalism for composition is limited to non-adaptive mechanisms [29, 20, 18]. On the other hand for $f$-DP, there is no general composition formula for arbitrary trade-off curves $f_{P_i|Q_i}$. Instead, Dong et al. [11] provides an explicit composition operator expression applicable only to Gaussian trade-off functions (Corollary 3.3). Furthermore, Zhu et al. [32]'s characteristic function formalism* (4) *appears equivalent to Rényi divergence, as we have learned that the order $q$ in* (26) *can assume imaginary values.*

[12]Adaptive means each mechanism's output may depend on previous outputs; heterogeneous means the mechanisms need not be identical.

In the following theorem, we provide an *exactly tight* composition result that applies to arbitrary privacy profiles $\delta_{P_1|Q_1}$ and $\delta_{P_2|Q_2}$. Later we also extend this theorem to handle adaptivity.

**Theorem 4.2** (Exactly Tight Composition of Privacy Profiles)**.** *If $P = P_1 \times P_2$ and $Q = Q_1 \times Q_2$ are two product distributions on $\Omega_1 \times \Omega_2$ such that $(P_1, Q_1)$ and $(P_2, Q_2)$ are absolutely continuous at least in one direction, then*

$$\delta_{P|Q}(\varepsilon) = \left(\delta_{P_1|Q_1} \circledast \left(\ddot{\delta}_{P_2|Q_2} - \dot{\delta}_{P_2|Q_2}\right)\right)(\varepsilon) = \int_{-\infty}^{\infty} \delta_{P_1|Q_1}(\varepsilon - \tau) \cdot \left(\ddot{\delta}_{P_2|Q_2}(\tau) - \dot{\delta}_{P_2|Q_2}(\tau)\right) \mathrm{d}\tau, \quad (47)$$

*where $\dot{\delta}_{P_2|Q_2}$ and $\ddot{\delta}_{P_2|Q_2}$ are the first and second order gradient functions of $\delta_{P_2|Q_2}$.*

*Proof.* Let's consider the random variables $Z \sim \mathrm{PLD}(P\|Q)$, $Z_1 \sim \mathrm{PLD}(P_1\|Q_1)$, and $Z_2 \sim \mathrm{PLD}(P_2\|Q_2)$. For a pair $(\Theta_1, \Theta_2) \sim P$, the privacy loss random variable $Z$ is given by $L_{P|Q}(\Theta_1, \Theta_2)$, which simplifies to:

$$\log \frac{P_1(\Theta_1)P_2(\Theta_2)}{Q_1(\Theta_1)Q_2(\Theta_2)} = L_{P_1|Q_1}(\Theta_1) + L_{P_2|Q_2}(\Theta_2). \quad (48)$$

This decomposition implies that $Z$ can be expressed as the sum of $Z_1$ and $Z_2$, i.e., $Z = Z_1 + Z_2$. Consequently, the probability density of $Z$, $f_Z$, is the convolution of $f_{Z_1}$ and $f_{Z_2}$:

$$f_Z(t) = \int_{-\infty}^{\infty} f_{Z_1}(\tau) f_{Z_2}(t - \tau) d\tau = (f_{Z_1} \circledast f_{Z_2})(t). \quad (49)$$

Invoking Theorem 3.2, we then obtain the Laplace transform of $f_Z$ at $(1 - q)$:

$$\mathcal{B}\left\{f_Z(t)\right\}(1 - q) \overset{(75)}{=} \mathcal{B}\left\{f_{Z_1}(t)\right\}(1 - q) \cdot \mathcal{B}\left\{f_{Z_2}(t)\right\}(1 - q). \quad (50)$$

From Definition 2.2, this directly implies that the Rényi divergence of order $q$ for the pair $(P, Q)$ is the sum of the Rényi divergences for the pairs $(P_1, Q_1)$ and $(P_2, Q_2)$ as

$$\mathrm{R}_q\left(P\|Q\right) = \mathrm{R}_q\left(P_1\|Q_1\right) + \mathrm{R}_q\left(P_2\|Q_2\right). \quad (51)$$

Now suppose $s = 1 - q < 0$. Thanks to absolute continuity, we can express the Rényi divergences in terms of the Laplace transform using Theorem 3.3 as follows, cancelling out the common terms:

$$s(s-1)\mathcal{B}\left\{\delta_{P|Q}(t)\right\}(s) = s(s-1)\mathcal{B}\left\{\delta_{P_1|Q_1}(t)\right\}(s) \cdot s(s-1)\mathcal{B}\left\{\delta_{P_2|Q_2}(t)\right\}(s). \quad (52)$$

Since $q = 1 - s$ cannot be 0 or 1, we can divide by $s(s - 1)$ on both sides:

$$\begin{aligned}
\mathcal{B}\left\{\delta_{P|Q}(t)\right\}(s) &= \mathcal{B}\left\{\delta_{P_1|Q_1}(t)\right\}(s) \cdot s(s-1)\mathcal{B}\left\{\delta_{P_2|Q_2}(t)\right\}(s) && (53)\\
&= \mathcal{B}\left\{\delta_{P_1|Q_1}(t)\right\}(s) \cdot \left(s^2\mathcal{B}\left\{\delta_{P_2|Q_2}(t)\right\}(s) - s\mathcal{B}\left\{\delta_{P_2|Q_2}(t)\right\}(s)\right) && (54)\\
&\overset{(73)}{=} \mathcal{B}\left\{\delta_{P_1|Q_1}(t)\right\}(s) \cdot \left(\mathcal{B}\left\{\ddot{\delta}_{P_2|Q_2}(t)\right\}(s) - \mathcal{B}\left\{\dot{\delta}_{P_2|Q_2}(t)\right\}(s)\right) && (55)\\
&\overset{(68)}{=} \mathcal{B}\left\{\delta_{P_1|Q_1}(t)\right\}(s) \cdot \mathcal{B}\left\{\ddot{\delta}_{P_2|Q_2}(t) - \dot{\delta}_{P_2|Q_2}(t)\right\}(s) && (56)\\
&\overset{(75)}{=} \mathcal{B}\left\{\left(\delta_{P_1|Q_1} \circledast \left(\ddot{\delta}_{P_2|Q_2} - \dot{\delta}_{P_2|Q_2}\right)\right)(t)\right\}(s). && (57)
\end{aligned}$$

Hence, from the uniqueness of Laplace transform, we get

$$\forall \varepsilon \in \mathbb{R}, \quad \delta_{P|Q}(\varepsilon) = \left(\delta_{P_1|Q_1} \circledast \left(\ddot{\delta}_{P_2|Q_2} - \dot{\delta}_{P_2|Q_2}\right)\right)(\varepsilon). \quad (58)$$

□

Theorem 4.2 provides an *exactly tight* composition theorem—not only because the terms in (47) are equal, but also because the privacy profiles $\delta_{P_1|Q_1}$ and $\delta_{P_2|Q_2}$ precisely capture the indistinguishability of their respective distributions. This theorem mirrors the composition property of Rényi divergence but works in the time domain $\varepsilon$ instead of the frequency domain $q$. It assumes absolute continuity in at least one direction ($P_i \ll Q_i$ or $Q_i \ll P_i$) for both $i = 1$ and 2.

Interestingly however, our result in (47) appears to hold even when absolute continuity fails in either direction. We will see an example of this in the next section where we apply (47) to compose the privacy profile of randomized mechanisms $\delta_{\mathrm{RR}}^{\varepsilon_1,\delta_1} \otimes \delta_{\mathrm{RR}}^{\varepsilon_2,\delta_2}$ and get a tight expression for $\delta_1, \delta_2 > 0$ without running into a singularity. This happens because even when the Laplace transform is undefined everywhere, the frequency-domain manipulations performed on it still correspond to valid manipulation steps in the time domain. Since our main interest lies in the time domain function—the composed privacy profile—taking advantage of this flexibility proves beneficial. To emphasize the significance of this, Theorem 4.2 allows us to tightly derive composed privacy profiles with the same exact-tightness as Rényi DP even for mechanisms that do not satisfy Rényi-DP for any order $q \in \mathbb{R} \setminus \{0, 1\}$, opening exciting possibilities beyond the limitations of Rényi DP.

**Remark 4.3.** *Formally proving why dropping the absolute continuity assumption in Theorem 4.2 does not compromise the validity of* (47) *appears to be a challenging yet intriguing problem. Our efforts to resolve this suggests that a mathematical understanding of the number $0^i$ is necessary.*

**Adaptive Composition.** For two mechanisms $\mathcal{M}_1$ and $\mathcal{M}_2$, if mechanism $\mathcal{M}_2$ sees the output from $\mathcal{M}_1$, then the output distribution of $\mathcal{M}_1$ and $\mathcal{M}_2$ are no longer independent. Theorem 4.2 can still be applied as long as their exists a distribution pair $P_2, Q_2$ that *dominates the privacy profile* of output distribution pairs $\mathcal{M}_2(D, \theta)$ and $\mathcal{M}_2(D', \theta)$ across all $\theta$ for a given dataset pair $D \simeq D'$, which is a reasonable assumption for adaptively compositing functional guarantees for DP.[13] This is possible due to the following result.

**Lemma 4.4** (Zhu et al. [32, Theorem 27]). *Let $P(x, y) = P_1(x) \cdot P_2^x(y)$ and $Q(x, y) = Q_1(x) \cdot Q_2^x(y)$ be two joint distributions on $\Omega_1 \times \Omega_2$. Then for any distributions $P_2$ and $Q_2$ on $\Omega_2$ such that $\delta_{P_2^x|Q_2^x}(\varepsilon) \leq \delta_{P_2|Q_2}(\varepsilon)$ for all $\varepsilon \in \mathbb{R}$ and $x \in \Omega_1$, we have $\delta_{P|Q}(\varepsilon) \leq \delta_{P_1 \times P_2|Q_1 \times Q_2}(\varepsilon)$.*

## 4.1 Tight Composition for $(\varepsilon, \delta)$-DP

In this section, we use Theorem 4.2 to prove an exactly-tight composition theorem for adaptive composition of a sequence of $(\varepsilon_i, \delta_i)$-DP mechanisms. We begin with a variant of Kairouz et al. [19]'s result that the privacy profile $\delta_{P|Q}(\varepsilon)$ under an $(\varepsilon, \delta)$-DP point guarantee is dominated by the privacy profile of randomized response mechanism.

**Theorem 4.5** (Dominating Privacy Profile under $(\varepsilon, \delta)$-DP [19]). *Fix $\varepsilon \geq 0$ and $\delta \in [0, 1]$. Suppose distributions $P$ and $Q$ over $\Omega$ satisfy $(\varepsilon, \delta)$-differential privacy. Then,*

$$\forall \varepsilon \in \mathbb{R} \;:\; \delta_{P|Q}(\varepsilon) \leq \delta_{\mathrm{RR}}(\varepsilon) \quad \textit{and} \quad \delta_{Q|P}(\varepsilon) \leq \delta_{\mathrm{RR}}(\varepsilon), \tag{59}$$

*where $\delta_{\mathrm{RR}}(t)$ is the privacy profile of the randomized response mechanism $\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}$.*

Kairouz et al. [19] do not express their notion of *dominance* in the same way as we do in Definition 3.2—they say distribution pair $(P_1, Q_1)$ dominates $(P_2, Q_2)$ if their trade-off curves satisfy

$$\forall \alpha \in [0, 1] \;:\; f_{P_1|Q_1}(\alpha) \geq f_{P_2|Q_2}(\alpha). \tag{60}$$

[13] Adaptive composition for point DP guarantees requires that for fixed $D \simeq D'$, conditioned on any observation from $\mathcal{M}_1$, the point DP guarantee holds for $\mathcal{M}_2$. Adaptively composing DP curves needs a stronger assumption that conditioned on any output of $\mathcal{M}_1$, the privacy profile of $\mathcal{M}_2$ lies below a worst-case DP curve.

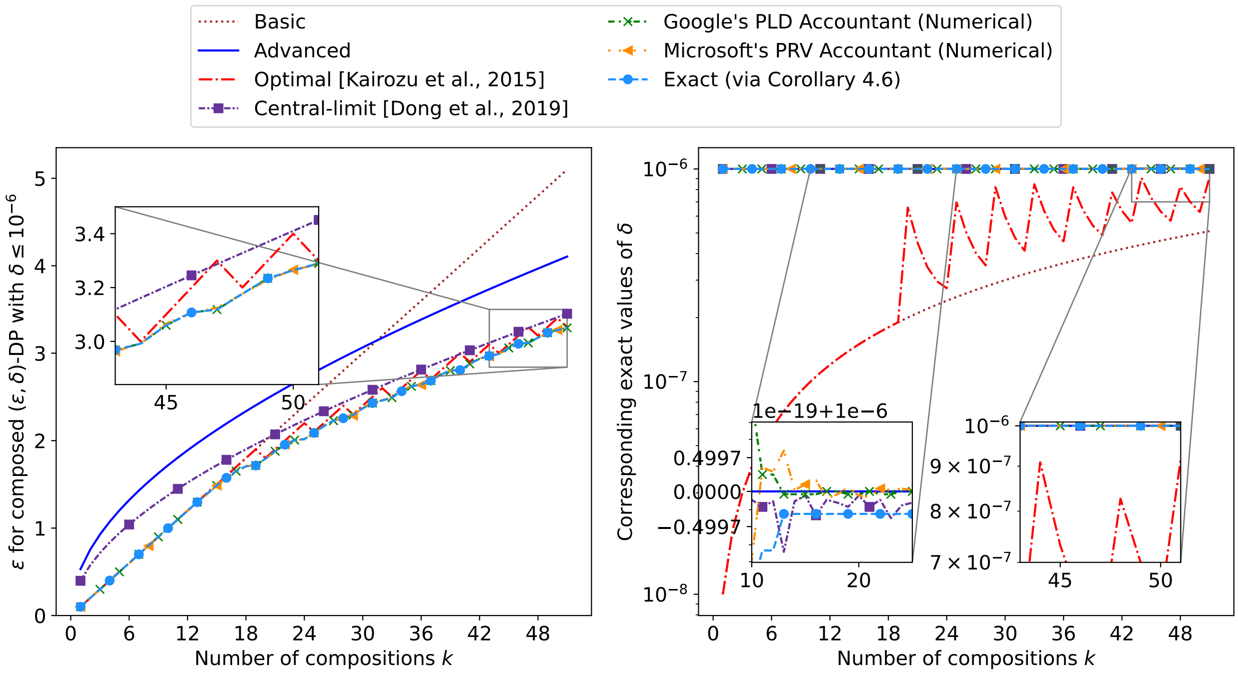



Figure 2: Comparison of $(\varepsilon, \delta)$-DP bounds from various composition theorems for $k$-fold composition of a $(0.1, 10^{-8})$-DP point guarantee, with the budget constraint $\delta < 10^{-6}$. The spikes in the right plot, showing exact $\delta < 10^{-6}$ values from Kairouz et al. [19, Theorem 3.3], occur because, out of a set of $\lfloor k/2 \rfloor$ DP point guarantees by their result, we select the smallest $\varepsilon$ corresponding to the largest $\delta < 10^{-6}$ in the set, which fluctuates as $k$ increases.

Nonetheless, our notion of dominance and their notion is equivalent, which was established by [11] by showing that the privacy profile $\delta_{P_i|Q_i}$ and the corresponding trade-off curve $f_{P_i|Q_i}$ are *primal* and *dual* with respect to Frenchel duality. We also provide a direct proof of Theorem 4.5 in the Appendix A.3.

As a side note, observe that combining Theorem 4.5 with Theorem 3.5 gives a tight Rényi DP guarantee for a pure $\varepsilon$-DP mechanism, which has recently attracted interest [27].

$$\delta_{P|Q}(\varepsilon) = \delta_{Q|P}(\varepsilon) = 0 \implies \forall q > 1 \ : \ \mathrm{R}_q\left(P \| Q\right) \leq \frac{1}{q-1} \log\left(\frac{e^{\varepsilon}}{e^{\varepsilon}+1} e^{-q\varepsilon} + \frac{1}{e^{\varepsilon}+1} e^{q\varepsilon}\right). \tag{61}$$

Following the objective of this section, we use our Theorem 4.2 on the above worst-case privacy profile under $(\varepsilon, \delta)$-DP point guarantees, resulting in an exactly-tight composition guarantee as stated below.

**Theorem 4.6** (Tight Composition for $(\varepsilon, \delta)$-DP)**.** *For any $\varepsilon_i \geq 0$, $\delta_i \in [0, 1]$ for $i \in \{1, \cdots, k\}$, the $k$-fold composition of $(\varepsilon_i, \delta_i)$-differentially private mechanisms satisfies $(\varepsilon, \delta^{\otimes k}(\varepsilon))$-DP for all $\varepsilon$, defined recursively as*

$$\forall t \in \mathbb{R} \ : \ \delta^{\otimes l}(t) = \delta_l + \frac{(1-\delta_l)}{e^{\varepsilon_l}+1}\left[e^{\varepsilon_l} \cdot \delta^{\otimes l-1}(t-\varepsilon_l) + \delta^{\otimes l-1}(t+\varepsilon_l)\right], \tag{62}$$

*with $\delta^{\otimes 0}(t) = [1 - e^t]_+$.*

Theorem 4.6 introduces a convenient recursive method for computing compositions of heterogeneous DP guarantees. While this bound matches each of the $\lfloor k \rfloor$ discrete $(\varepsilon, \delta)$-DP values given by the optimal composition theorem from Kairouz et al. [19, Theorem 3.3], our result offers a continuous curve over all $\varepsilon \in \mathbb{R}$. Consequently, for a given budget on $\delta$, our bound provides a tighter limit on $\varepsilon$ than that of [19], as shown in Figure 2. Furthermore, the recursion reduces to the following exact expression when composing homogeneous DP guarantees.

**Corollary 4.7.** *For any $\varepsilon \geq 0$, $\delta \in [0,1]$, the $k$-fold composition of $(\varepsilon, \delta)$-DP mechanisms satisfies $(\varepsilon, \delta^{\otimes k}(\varepsilon))$-DP for all $\varepsilon$, where*

$$\forall t \in \mathbb{R} \; : \; \delta^{\otimes k}(t) = 1 - (1-\delta)^k \left( 1 - \mathop{\mathbb{E}}_{Y \leftarrow \mathrm{Binomial}\left(k, \frac{e^{\varepsilon}}{1+e^{\varepsilon}}\right)} \left[ 1 - e^{t - \varepsilon \cdot (2Y - k)} \right]_+ \right). \tag{63}$$

Figure 2 illustrates the enhanced privacy quantification achieved by our bound for $k$-fold composition of $(\varepsilon, \delta)$-DP guarantees. We also compare our results with Google's PLDAccountant [12] and Microsoft's PRVAccountant [18], which utilize the Discrete Fast Fourier Transform. This comparison shows that our analytical approach closely matches the values approximated by these numerical methods. Additionally, through Figure 4 in the appendix, we show that these numerical methods can be unstable at edge case values and yield non-negligible gaps in their approximation.

# 5 Asymmetry and DP Notion Equivalences

An important characteristic of functional notions of DP is that they can be *asymmetric*, in the sense that switching $P \leftrightarrow Q$ may yield a different curve $\delta_{Q|P}$ than the original $\delta_{P|Q}$. In context of a mechanism $\mathcal{M}$, such an asymmetry between its output distribution $P = \mathcal{M}(D)$ and $Q = \mathcal{M}(D')$ means that a sample $\Theta \sim P$ might reveal more (or less) information that it came from $D$ than a sample $\Theta' \sim Q$ reveals about coming from $D'$. Since $D$ and $D'$ are neighboring datasets differing by the presence or absence of a single record, this asymmetry in indistinguishability means that an attacker might have an easier time trying to detect the presence of a record from the output of $\mathcal{M}$ than to detect its absence. In other words, optimal hypothesis tests would experience a skew in the trade-off between their false positive and false negative rates.

Such skewness often arises due to *subsampling*, which is a heavily used technique to boost an algorithm's intrinsic privacy properties [5, 7, 1, 31, 4]. For instance, Poisson subsampling in the context of the *add or remove* relationship between neighboring datasets, which is commonly used in DP-SGD [1] algorithm, skews the privacy profile of a base mechanism, as illustrated in the following example.

**Effect of Poisson Subsampling on $\delta_{P|Q}$.** Without loss of generality, assume datasets $D \simeq D'$ are such that the record at index $i$ is present in $D$ but empty in $D'$, i.e., $D[i] \neq D'[i] = \perp$. If we randomly filter the records using an iid selection mask $U \sim \mathrm{Bernoulli}(\lambda)^{\otimes n}$, the subsampled datasets $D_U$ and $D'_U$ are defined as follows for all $i \in [n]$:

$$D_U[i] := \begin{cases} D[i] & \text{if } U_i = 1 \\ \perp & \text{otherwise} \end{cases} \quad \text{and} \quad D'_U[i] := \begin{cases} D'[i] & \text{if } U_i = 1 \\ \perp & \text{otherwise} \end{cases}. \tag{64}$$

The distributions $P$ and $Q$ of the outputs $\mathcal{M}(D_U)$ and $\mathcal{M}(D'_U)$ for any algorithm $\mathcal{M}$ will be identical with probability $1 - \lambda$, which amplifies privacy considerably. More precisely, let $P_{\mathrm{IN}}$ and $Q_{\mathrm{IN}}$ be the distributions of $\mathcal{M}(D_U)$ and $\mathcal{M}(D'_U)$ conditioned on $i \in U$, and $P_{\mathrm{OUT}}$ and $Q_{\mathrm{OUT}}$ be the distributions conditioned on $i \notin U$. For any event $S \subseteq \Omega$, we have:

$$\begin{aligned} P(S) &= \Pr[i \notin U] \cdot P_{\mathrm{OUT}}(S) + \Pr[i \in U] \cdot P_{\mathrm{IN}}(S) \\ &= (1-\lambda) \cdot Q_{\mathrm{OUT}}(S) + \lambda \cdot P_{\mathrm{IN}}(S) && (65) \\ &= (1-\lambda) \cdot Q_{\mathrm{IN}}(S) + \lambda \cdot P_{\mathrm{IN}}(S). && (66) \end{aligned}$$

Equation (65) holds because if $i \notin U$, then $D_U = D'_U$, and equation (66) holds because the $i$th record in $D'$ is empty, so conditioning on $i \in U$ or $i \notin U$ does not affect the output distribution, i.e., $Q_{\text{IN}} = Q_{\text{OUT}} = Q$. Using this fact, we the following theorem shows the *exact effect* Poisson subsampling has on the privacy profile.

**Theorem 5.1** (Poisson subsampling on add/remove neighbours)**.** *Let $0 < \lambda \leq 1$. For any distributions $P$, $Q$, $P_{\text{IN}}$ and $Q_{\text{IN}}$ such that $P = \lambda P_{\text{IN}} + (1-\lambda)Q_{\text{IN}}$ and $Q = Q_{\text{IN}}$,*

$$\delta_{P|Q}(\varepsilon) = \begin{cases} \lambda\delta_{P_{\text{IN}}|Q_{\text{IN}}}(\log(1 + (e^{\varepsilon} - 1)/\lambda)) & \text{if } \varepsilon > \log(1-\lambda), \\ 1 - e^{\varepsilon} & \text{otherwise.} \end{cases} \tag{67}$$

This privacy amplification result was first provided by Li et al. [21], and since has appeared in several works [5, 4, 1, 30]. But unlike other works, we present this amplification effect as a *single curve* that exactly captures the impact of subsampling on both directions. This effect is visualized in the top-left plot in Figure 3 (solid orange curve vs. dashed orange curve).

**Imprecise Handling of Asymmetry.** We observe that several works on privacy handle asymmetric notions of DP in a somewhat imprecise manner, which can introduce significant slack in the analysis or numerical bounds. For example, Dong et al. [11, Theorem 4.2] uses the biconjugation operation $\min\{f_{P|Q}, f_{Q|P}\}^{**}$ to quantify amplification for Poisson subsampling on the trade-off curve $f_{P|Q}$ (corresponding to the subsampled profile $\delta_{P|Q}$ in Theorem 5.1), leading to an overestimation of the actual trade-off curve $f_{P|Q}$ (see the bottom-right plot in Figure 3 for a comparison of the overestimated trade-off after symmetrization in the red line versus the actual trade-off $f_{P|Q}$ in the orange dashed line). Similarly, Dong et al. [11, Proposition 30] defines a dominating profile under subsampling by taking the max of $\delta_{P|Q}$ and $\delta_{Q|P}$ (see the top-left plot in Figure 3 to compare the overestimated privacy profile in red with the actual profiles).

The reasoning behind these operations is that a DP guarantee function should hold in both directions, requiring consideration of the worst-case aspects from each direction. However, these symmetrization steps can introduce small gaps that may compound significantly when several privacy profiles are composed. Moreover, these operations disrupt the equivalence across different privacy notions, as after symmetrization, converting to another notion, such as the privacy profile, no longer aligns with the actual privacy profile.

**Proposed Solution.** We note that all these functional notions of privacy possess a reversal property (see Remarks 2.1, 2.2, 2.3, and 3.1), which allows us to bypass the need for symmetrization operations. The key idea is to retain the chosen characterizing function *in only one direction*, without attempting to control its asymmetry. Additionally, while composing functional notions, we refrain from changing the notion of adjacency (or its direction). Doing this enables lossless composition operations, thanks to Theorem 4.2 or Rényi DP composition [23], without running the risk of accidentally underestimating the privacy. When a pointwise DP guarantee is required, we can leverage the reversal properties to query the curve at a specific budget constraint in both directions, providing the max of the two, which results in an exactly-tight pointwise DP guarantee. Additionally, we also avoid the overhead of maintaining the functional representation, such as the PLD, in both directions as currently done in Google's PLDAccountant.

## 6 Conclusion

In summary, this paper presents a novel interpretation of differential privacy by leveraging time-frequency dualities across various functional representations, including privacy profiles, Rényi divergence, and privacy loss distributions. By framing these within the context of Laplace transforms,

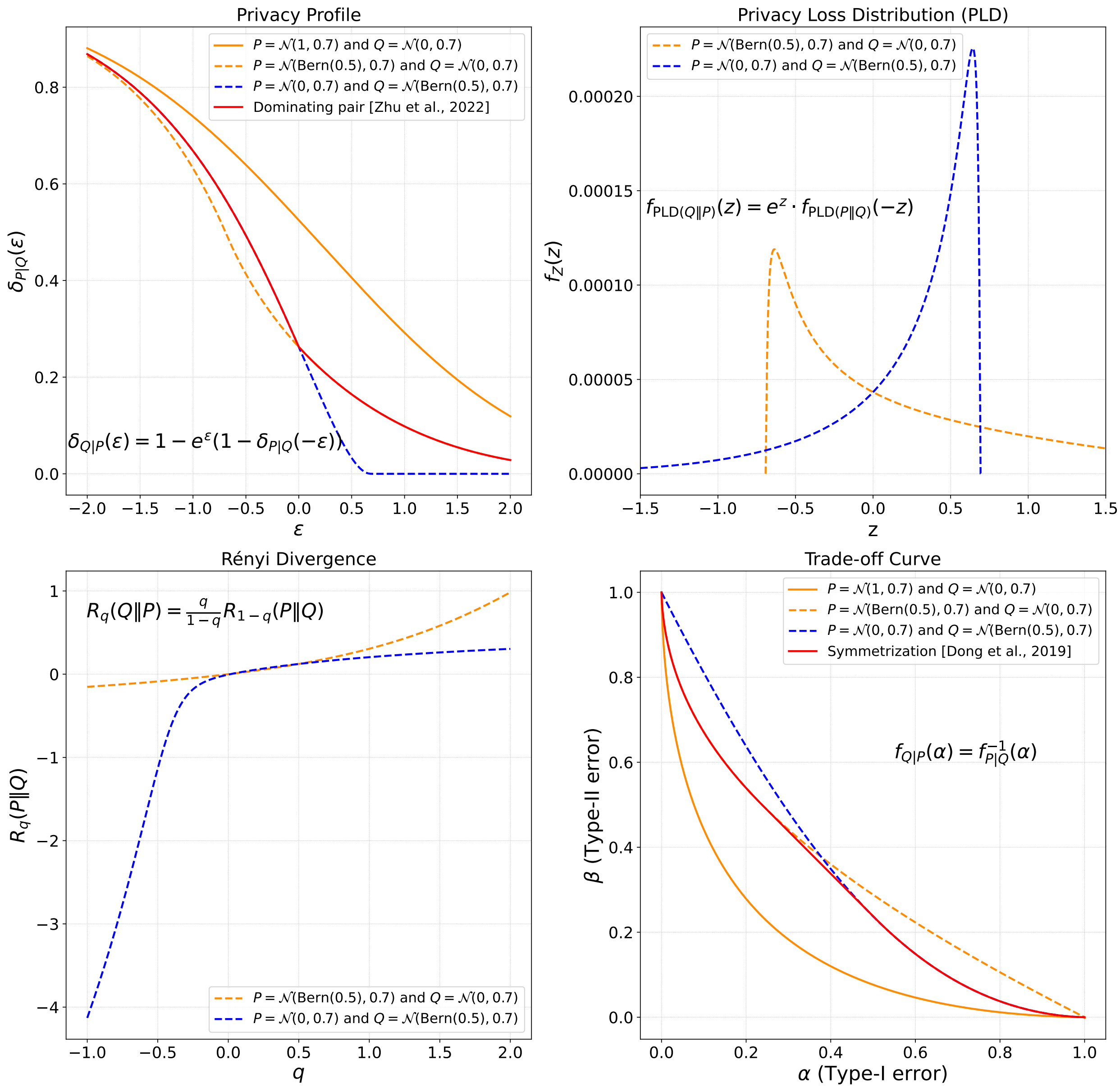



Figure 3: Visualization of the four functional notions of DP, namely privacy profile $\delta_{P|Q}(\varepsilon)$ as a function of $\varepsilon$, the generalized density function of privacy loss distribution $\mathrm{PLD}(P\|Q)$, the Rényi divergence $\mathrm{R}_q(P\|Q)$ as a function of order $q$, and the trade-off function $f_{P|Q}(\alpha)$ for hypothesis testing between $P$ and $Q$. We also provide the reversal theorems for each of the plots.

we develop a versatile analytical toolkit for DP, enhancing both theoretical understanding and practical composition methods. Our approach addresses limitations in existing adaptive composition bounds, provides continuous guarantees for composed privacy profiles, and bridges gaps between different DP frameworks without needing approximations or symmetrizations. Together, these results push forward the capabilities of differential privacy research, setting a foundation for more nuanced and robust privacy-preserving algorithms.

## References


[1] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 308–318, 2016.

[2] S. Asoodeh, J. Liao, F. P. Calmon, O. Kosut, and L. Sankar. A better bound gives a hundred rounds: Enhanced privacy guarantees via f-divergences. In *2020 IEEE International Symposium on Information Theory (ISIT)*, pages 920–925. IEEE, 2020.

[3] B. Balle and Y.-X. Wang. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *International Conference on Machine Learning*, pages 394–403. PMLR, 2018.

[4] B. Balle, G. Barthe, and M. Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in neural information processing systems*, 31, 2018.

[5] B. Balle, G. Barthe, and M. Gaboardi. Privacy profiles and amplification by subsampling. *Journal of Privacy and Confidentiality*, 10(1), 2020.

[6] B. Balle, G. Barthe, M. Gaboardi, J. Hsu, and T. Sato. Hypothesis testing interpretations and renyi differential privacy. In *International Conference on Artificial Intelligence and Statistics*, pages 2496–2506. PMLR, 2020.

[7] R. Bassily, A. Smith, and A. Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *2014 IEEE 55th annual symposium on foundations of computer science*, pages 464–473. IEEE, 2014.

[8] M. Bun and T. Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658. Springer, 2016.

[9] C. L. Canonne, G. Kamath, and T. Steinke. The discrete gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33:15676–15688, 2020.

[10] A. M. Cohen. *Numerical methods for Laplace transform inversion*, volume 5. Springer Science & Business Media, 2007.

[11] J. Dong, A. Roth, and W. J. Su. Gaussian differential privacy. *arXiv preprint arXiv:1905.02383*, 2019.

[12] V. Doroshenko, B. Ghazi, P. Kamath, R. Kumar, and P. Manurangsi. Connect the dots: Tighter discrete approximations of privacy loss distributions. *arXiv preprint arXiv:2207.04380*, 2022.

[13] C. Dwork. Differential privacy. In *International colloquium on automata, languages, and programming*, pages 1–12. Springer, 2006.

[14] C. Dwork and G. N. Rothblum. Concentrated differential privacy. *arXiv preprint arXiv:1603.01887*, 2016.

[15] C. Dwork, G. N. Rothblum, and S. Vadhan. Boosting and differential privacy. In *2010 IEEE 51st annual symposium on foundations of computer science*, pages 51–60. IEEE, 2010.

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

[17] P. P. Dyke and P. Dyke. *An introduction to Laplace transforms and Fourier series*. Springer, 2001.

[18] S. Gopi, Y. T. Lee, and L. Wutschitz. Numerical composition of differential privacy. *Advances in Neural Information Processing Systems*, 34:11631–11642, 2021.

[19] P. Kairouz, S. Oh, and P. Viswanath. The composition theorem for differential privacy. In *International conference on machine learning*, pages 1376–1385. PMLR, 2015.

[20] A. Koskela, J. Jälkö, and A. Honkela. Computing tight differential privacy guarantees using fft. In *International Conference on Artificial Intelligence and Statistics*, pages 2560–2569. PMLR, 2020.

[21] N. Li, W. Qardaji, and D. Su. On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy. In *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, pages 32–33, 2012.

[22] J. F. Miller. Moment-generating functions and laplace transforms. *Journal of the Arkansas Academy of Science*, 4(1):97–100, 1951.

[23] I. Mironov. Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pages 263–275. IEEE, 2017.

[24] J. Murtagh and S. Vadhan. The complexity of computing the optimal composition of differential privacy. In *Theory of Cryptography Conference*, pages 157–175. Springer, 2015.

[25] A. V. Oppenhiem, A. S. Willsky, and S. H. Nawab. Signals and systems, 1996.

[26] J. Orloff. Uniqueness of laplace transform, 2015.

[27] D. Privacy. Tight RDP & zCDP Bounds from Pure DP — differentialprivacy.org. `https://differentialprivacy.org/pdp-to-zcdp/`. [Accessed 12-11-2024].

[28] I. Sason and S. Verdú. $f$-divergence inequalities. *IEEE Transactions on Information Theory*, 62 (11):5973–6006, 2016.

[29] D. Sommer, S. Meiser, and E. Mohammadi. Privacy loss classes: The central limit theorem in differential privacy. *Cryptology ePrint Archive*, 2018.

[30] T. Steinke. Composition of differential privacy & privacy amplification by subsampling. *arXiv preprint arXiv:2210.00597*, 2022.

[31] Y.-X. Wang, S. Fienberg, and A. Smola. Privacy for free: Posterior sampling and stochastic gradient monte carlo. In *International Conference on Machine Learning*, pages 2493–2502. PMLR, 2015.

[32] Y. Zhu, J. Dong, and Y.-X. Wang. Optimal accounting of differential privacy via characteristic function. In *International Conference on Artificial Intelligence and Statistics*, pages 4782–4817. PMLR, 2022.

# A Appendix

## A.1 Table of Properties of Laplace Transform

Table 1: Properties of the Laplace Transform. Let $g(t)$ and $h(t)$ be two functions defined for $t \in \mathbb{R}$ and let $a, b \in \mathbb{R}$ be arbitrary constants.

| Property | Expression | |
|---|---|---|
| Linearity : | $\mathcal{L}\left\{ag(t)+bh(t)\right\}(s) = a\mathcal{L}\left\{g(t)\right\}(s) + b\mathcal{L}\left\{h(t)\right\}(s)$<br>$\mathcal{B}\left\{ag(t)+bh(t)\right\}(s) = a\mathcal{B}\left\{g(t)\right\}(s) + b\mathcal{B}\left\{h(t)\right\}(s)$ | (68) |
| Time-Shifting : | $\mathcal{L}\left\{g(t-a)\mathbb{I}\{t>a\}\right\}(s) = e^{-as}\mathcal{L}\left\{g(t)\right\}(s),\ \text{for } a>0$<br>$\mathcal{B}\left\{g(t-a)\right\}(s) = e^{-as}\mathcal{B}\left\{g(t)\right\}(s),\ \text{for } a \in \mathbb{R}$ | (69) |
| Frequency-Shifting : | $\mathcal{L}\left\{e^{at}g(t)\right\}(s) = \mathcal{L}\left\{g(t)\right\}(s-a)$<br>$\mathcal{B}\left\{e^{at}g(t)\right\}(s) = \mathcal{B}\left\{g(t)\right\}(s-a)$ | (70) |
| Time-Scaling : | $\mathcal{L}\left\{g(at)\right\}(s) = \frac{1}{a}\mathcal{L}\left\{g(t)\right\}\left(\frac{s}{a}\right) \text{ for } a>0$<br>$\mathcal{B}\left\{g(at)\right\}(s) = \frac{1}{\lvert a\rvert}\mathcal{B}\left\{g(t)\right\}\left(\frac{s}{a}\right) \text{ for } a \in \mathbb{R}$ | (71) |
| Reversal : | $\mathcal{B}\left\{g(-t)\right\}(s) = \mathcal{B}\left\{g(t)\right\}(-s)$ | (72) |
| Derivative : | $\mathcal{L}\left\{\dot{g}(t)\right\}(s) = s\mathcal{L}\left\{g(t)\right\}(s) - g(0^+)$<br>$\mathcal{B}\left\{\dot{g}(t)\right\}(s) = s\mathcal{B}\left\{g(t)\right\}(s)$ | (73) |
| Integration : | $\mathcal{L}\left\{\int_0^t g(t)dt\right\}(s) = \frac{1}{s}\mathcal{L}\left\{g(t)\right\}(s),\ \text{for } \mathfrak{Re}(s)>0$<br>$\mathcal{B}\left\{\int_{-\infty}^t g(t)dt\right\}(s) = \frac{1}{s}\mathcal{B}\left\{g(t)\right\}(s),\ \text{for } \mathfrak{Re}(s)>0$ | (74) |
| Convolution : | $\mathcal{L}\left\{\int_0^t g(\tau)h(t-\tau)d\tau\right\}(s) = \mathcal{L}\left\{g(t)\right\}(s)\cdot\mathcal{L}\left\{h(t)\right\}(s)$<br>$\mathcal{B}\left\{\int_{-\infty}^{\infty} g(\tau)h(t-\tau)d\tau\right\}(s) = \mathcal{B}\left\{g(t)\right\}(s)\cdot\mathcal{B}\left\{h(t)\right\}(s)$ | (75) |

### A.2 Deferred Proofs for Section 3

**Theorem 3.2.** *For a random variable $X$, let $F_X(t) := \Pr[X \le t]$ denote its* cumulative distribution function *and $f_X(t)$ denote its* generalized probability density function. *Let $P$ and $Q$ be probability distributions and $Z \sim \mathrm{PLD}(P\|Q)$ and $Z' \sim \mathrm{PLD}(Q\|P)$ denote their privacy loss random variables. If $Z \sim \mathrm{PLD}(P\|Q)$ and $Z' \sim \mathrm{PLD}(Q\|P)$, then for all $\varepsilon \in \mathbb{R}$,*

$$\delta_{P|Q}(\varepsilon) = \mathcal{L}\left\{1 - F_Z(t+\varepsilon)\right\}(1) \tag{76}$$
$$= e^{\varepsilon} \cdot \mathcal{L}\left\{F_{Z'}(-t-\varepsilon)\right\}(-1) \tag{77}$$
$$= e^{\varepsilon} \cdot \mathcal{L}\left\{f_{Z'}(-t-\varepsilon)\right\}(-1) - \mathcal{L}\left\{f_Z(t+\varepsilon)\right\}(1) \tag{78}$$
$$= \mathcal{L}\left\{f_Z(t+\varepsilon)\right\}(0) - e^{\varepsilon} \cdot \mathcal{L}\left\{f_{Z'}(-t-\varepsilon)\right\}(0). \tag{79}$$

*And, for all $q \in \mathrm{ROC}_{\mathcal{B}}\{f_{Z'}\}$ (or equivalently, $1-q \in \mathrm{ROC}_{\mathcal{B}}\{f_Z\}$),*

$$e^{(q-1)\cdot \mathrm{R}_q(P\|Q)} = \mathrm{E}_q\left(P\|Q\right) = \mathcal{B}\{f_Z(t)\}(1-q) = \mathcal{B}\{f_{Z'}(t)\}(q). \tag{80}$$

*Proof.* Denote the set where privacy loss $L_{P|Q}(\theta)$ exceeds $\varepsilon$ as

$$S^*_{>\varepsilon} = \{\theta \in \Omega : P(\theta) > e^{\varepsilon}Q(\theta)\}. \tag{81}$$

Note that for all $S \subset \Omega$, and all $\varepsilon \in \mathbb{R}$, we have

$$P(S) - e^{\varepsilon}Q(S) \le P(S^*_{>\varepsilon}) - e^{\varepsilon}Q(S^*_{>\varepsilon}) = \delta_{P|Q}(\varepsilon), \tag{82}$$

because $S^*_{>\varepsilon}$ includes any and all points where $P(\theta) > e^{\varepsilon}Q(\theta)$. We can express the probabilities $P(S^*_{>\varepsilon})$ and $Q(S^*_{>\varepsilon})$ using the Laplace transform as follows:

$$P(S^*_{>\varepsilon}) = \int_{S^*_{>\varepsilon}} P(\theta)d\theta \tag{83}$$
$$= \int_{S^*_{>\varepsilon}} \left(\frac{P(\theta)}{Q(\theta)}\right) Q(\theta)d\theta \tag{84}$$
$$= \int_{S^*_{>\varepsilon}} e^{-L_{Q|P}(\theta)}Q(\theta)d\theta \tag{85}$$
$$= \int_{\varepsilon^+}^{\infty} e^t \int_{\{\theta\in\Omega : L_{Q|P}(\theta)=-t\}} Q(\theta)d\theta \tag{86}$$
$$= \int_{\varepsilon^+}^{\infty} e^t \int_{-t^-}^{-t^+} F_{Z'}(u)du \tag{87}$$
$$= \int_{\varepsilon^+}^{\infty} e^t f_{Z'}(-t)dt \tag{88}$$
$$= e^{\varepsilon} \int_{0^+}^{\infty} e^{t'} f_{Z'}(-t'-\varepsilon)dt' \tag{89}$$
$$= e^{\varepsilon}\mathcal{L}\left\{f_{Z'}(-t-\varepsilon)\right\}(-1). \tag{90}$$

$$Q(S^*_{>\varepsilon}) = \int_{S^*_{>\varepsilon}} Q(\theta)d\theta \tag{91}$$
$$= \int_{S^*_{>\varepsilon}} \left(\frac{Q(\theta)}{P(\theta)}\right) P(\theta)d\theta \tag{92}$$
$$= \int_{S^*_{>\varepsilon}} e^{-L_{P|Q}(\theta)}P(\theta)d\theta \tag{93}$$
$$= \int_{\varepsilon^+}^{\infty} e^{-t} \int_{\{\theta\in\Omega : L_{P|Q}(\theta)=t\}} P(\theta)d\theta \tag{94}$$
$$= \int_{\varepsilon^+}^{\infty} e^{-t} \int_{t^-}^{t^+} F_Z(u)du \tag{95}$$
$$= \int_{\varepsilon^+}^{\infty} e^{-t} f_Z(t)dt \tag{96}$$
$$= e^{-\varepsilon} \int_{0^+}^{\infty} e^{-t'} f_Z(t'+\varepsilon)dt \tag{97}$$
$$= e^{-\varepsilon}\mathcal{L}\left\{f_Z(t+\varepsilon)\right\}(1). \tag{98}$$

Combining the two, we get equation (78):

$$\delta_{P|Q}(\varepsilon) = e^{\varepsilon} \cdot \mathcal{L}\left\{f_{Z'}(-t-\varepsilon)\right\}(-1) - \mathcal{L}\left\{f_Z(t+\varepsilon)\right\}(1). \tag{99}$$

Alternatively, we can express the profile directly as:

$$
\begin{aligned}
P(S^*_{>\varepsilon}) - e^{\varepsilon} Q(S^*_{>\varepsilon}) &= \int_{S^*_{>\varepsilon}} (1 - e^{\varepsilon} \frac{Q(\theta)}{P(\theta)}) P(\theta) d\theta && (100) \\
&= \int_{S^*_{>\varepsilon}} (1 - e^{\varepsilon - L_{P|Q}(\theta)}) P(\theta) d\theta && (101) \\
&= \int_{\varepsilon^+}^{\infty} (1 - e^{\varepsilon - t}) f_Z(t) dt && (102) \\
&= \int_{\varepsilon^+}^{\infty} f_Z(t) dt - \int_{0^+}^{\infty} e^{-t'} f_Z(t' + \varepsilon) dt' && (\text{Change } t = t' + \varepsilon) \\
&= 1 - F_Z(\varepsilon^+) - \mathcal{L}\left\{ f_Z(t + \varepsilon) \right\}(1) && (103) \\
&\overset{(73)}{=} \mathcal{L}\left\{ 1 - F_Z(t + \varepsilon) \right\}(1). && (104)
\end{aligned}
$$

Similarly, we can express it in terms of $Z'$ as

$$
\begin{aligned}
P(S^*_{>\varepsilon}) - e^{\varepsilon} Q(S^*_{>\varepsilon}) &= \int_{S^*_{>\varepsilon}} (\frac{P(\theta)}{Q(\theta)} - e^{\varepsilon}) Q(\theta) d\theta && (105) \\
&= \int_{S^*_{>\varepsilon}} (e^{L_{P|Q}(\theta)} - e^{\varepsilon}) Q(\theta) d\theta && (106) \\
&= \int_{\varepsilon^+}^{\infty} (e^t - e^{\varepsilon}) f_{Z'}(-t) dt && (107) \\
&= e^{\varepsilon} \left( \int_{0^+}^{\infty} e^{t'} f_{Z'}(-t' - \varepsilon) dt' - \int_{\varepsilon^+}^{\infty} f_{Z'}(-t) dt \right) && (\text{Change } t = t' + \varepsilon) \\
&= e^{\varepsilon} \left( \mathcal{L}\left\{ f_{Z'}(-t - \varepsilon) \right\}(-1) - F_{Z'}(-\varepsilon^-) \right) && (108) \\
&\overset{(73)}{=} e^{\varepsilon} \mathcal{L}\left\{ F_{Z'}(-t - \varepsilon) \right\}(-1). && (109)
\end{aligned}
$$

For showing (79), we apply the derivative property of Laplace transform to (76) and (77) to get

$$
\delta_{P|Q}(\varepsilon) = \mathcal{L}\left\{ 1 - F_Z(t + \varepsilon) \right\}(1) \overset{(73)}{=} -\mathcal{L}\left\{ f_Z(t + \varepsilon) \right\}(1) + 1 - F_Z(\varepsilon^+), \quad \text{and} \tag{110}
$$

$$
\delta_{P|Q}(\varepsilon) = e^{\varepsilon} \cdot \mathcal{L}\left\{ F_{Z'}(-t - \varepsilon) \right\}(-1) \overset{(73)}{=} e^{\varepsilon} \cdot \left( \mathcal{L}\left\{ f_{Z'}(-t - \varepsilon) \right\}(-1) - F_{Z'}(-\varepsilon^-) \right). \tag{111}
$$

Adding the above two equations and subtracting (78) from it, we get

$$
\begin{aligned}
\delta_{P|Q}(\varepsilon) &= 1 - F_Z(\varepsilon^+) - e^{\varepsilon} F_{Z'}(-\varepsilon^-) && (112) \\
&= \Pr[Z > \varepsilon] - e^{\varepsilon} \Pr[Z' < -\varepsilon] && (113) \\
&= \int_{0^+}^{\infty} e^{0 \cdot t} \cdot f_Z(t + \varepsilon) dt - e^{\varepsilon} \cdot \int_{0^+}^{\infty} e^{0 \cdot t} f_{Z'}(-t - \varepsilon) dt && (114) \\
&= \mathcal{L}\left\{ f_Z(t + \varepsilon) \right\}(0) - e^{\varepsilon} \cdot \mathcal{L}\left\{ f_{Z'}(-t - \varepsilon) \right\}(0). && (115)
\end{aligned}
$$

For the last part, recall from definition that Rényi divergence $\mathrm{R}_q(P \| Q) = \frac{1}{q-1} \log \mathrm{E}_q(P \| Q)$, for which we show the following two equivalences:

$$\mathrm{E}_q\left(P\|Q\right) = \int_\Omega \left(\frac{P(\theta)}{Q(\theta)}\right)^{q-1} P(\theta)d\theta \tag{116}$$
$$= \int_\Omega e^{(q-1)L_{P|Q}(\theta)} P(\theta)d\theta \tag{117}$$
$$= \int_{-\infty}^{\infty} e^{(q-1)t} f_Z(t)dt \tag{118}$$
$$= \mathcal{B}\left\{f_Z(t)\right\}(1-q). \tag{119}$$

$$\mathrm{E}_q\left(P\|Q\right) = \int_\Omega \left(\frac{P(\theta)}{Q(\theta)}\right)^{q} Q(\theta)d\theta \tag{120}$$
$$= \int_\Omega e^{-qL_{Q|P}(\theta)} Q(\theta)d\theta \tag{121}$$
$$= \int_{-\infty}^{\infty} e^{-qt} f_{Z'}(t)dt \tag{122}$$
$$= \mathcal{B}\left\{f_{Z'}(t)\right\}(q). \tag{123}$$

□

**Theorem 3.4** (Privacy profile of randomized response). *Fix $\varepsilon > 0$ and $\delta \in [0, 1]$. Let $\mathcal{M}_{\mathrm{RR}} : \{0, 1\} \to \{0, 1\} \times \{\bot, \top\}$ be the randomized response mechanism, which has the following output probabilities.*

$$\mathcal{M}_{\mathrm{RR}}(0) = \begin{cases} (0,\bot) & \text{with probability } \delta, \\ (0,\top) & \text{with probability } \frac{(1-\delta)e^\varepsilon}{e^\varepsilon+1}, \\ (1,\top) & \text{with probability } \frac{(1-\delta)}{e^\varepsilon+1}, \\ (1,\bot) & \text{with probability } 0, \end{cases} \quad \mathcal{M}_{\mathrm{RR}}(1) = \begin{cases} (0,\bot) & \text{with probability } 0, \\ (0,\top) & \text{with probability } \frac{(1-\delta)}{e^\varepsilon+1}, \\ (1,\top) & \text{with probability } \frac{(1-\delta)e^\varepsilon}{e^\varepsilon+1}, \\ (1,\bot) & \text{with probability } \delta. \end{cases} \tag{124}$$

*For $P = \mathcal{M}_{\mathrm{RR}}(0)$ and $Q = \mathcal{M}_{\mathrm{RR}}(1)$, the privacy profiles are*

$$\forall t \in \mathbb{R} \ : \ \delta_{P|Q}(t) = \delta_{Q|P}(t) = \begin{cases} \delta & \text{if } \varepsilon < t, \\ 1 - \frac{(e^t+1)}{e^\varepsilon+1}(1-\delta) & \text{if } -\varepsilon < t \le \varepsilon, \\ 1 - e^t(1-\delta) & \text{if } t \le -\varepsilon. \end{cases} \tag{125}$$

*Proof.* Let $S_1 = \{(0,\bot)\}$, $S_2 = S_1 \cup \{(1,\bot)\}$, and $S_3 = S_2 \cup \{(1,\top)\}$. From (30),

$$\delta_{P|Q}(t) = \Pr_{Z\leftarrow\mathrm{PLD}(P\|Q)}[Z > t] - e^\varepsilon \cdot \Pr_{Z'\leftarrow\mathrm{PLD}(Q\|P)}[Z' < -t] \tag{126}$$
$$= \Pr_P[\log\frac{P(\Theta)}{Q(\Theta)} > t] - e^\varepsilon \cdot \Pr_Q[\log\frac{Q(\Theta)}{P(\Theta)} < -t] \tag{127}$$
$$= \Pr_P[P(\Theta) > e^t \cdot Q(\Theta)] - e^\varepsilon \cdot \Pr_Q[P(\Theta) < e^t \cdot Q(\Theta)] \tag{128}$$
$$= \begin{cases} P(S_1) - e^t \cdot Q(S_1) & \text{if } \varepsilon < t, \\ P(S_2) - e^t \cdot Q(S_2) & \text{if } -\varepsilon < t \le \varepsilon, \\ P(S_3) - e^t \cdot Q(S_3) & \text{otherwise} \end{cases} \tag{129}$$
$$= \begin{cases} \delta & \text{if } \varepsilon < t, \\ \delta + \frac{1-\delta}{e^\varepsilon+1}\cdot(e^\varepsilon - e^t) & \text{if } -\varepsilon < t \le \varepsilon, \\ 1 - e^t\cdot(1-\delta) & \text{otherwise} \end{cases} \tag{130}$$

The same holds for $\delta_{Q|P}$ due to symmetry of (124). □

**Theorem 3.5** (Rényi DP of $(\varepsilon, 0)$-Randomized Response). *For any $\varepsilon > 0$ and $\delta = 0$, the output distributions of randomized response mechanism in Theorem 3.4 exhibit a Rényi divergence*

$$\forall q \in \mathbb{C} \ s.t.\ \mathfrak{Re}(q) \notin \{0,1\} : \mathrm{R}_q\left(P\|Q\right) = \frac{1}{q-1}\log\left(\frac{e^\varepsilon}{1+e^\varepsilon}e^{-q\varepsilon} + \frac{1}{1+e^\varepsilon}e^{q\varepsilon}\right). \tag{131}$$

*Proof.* From Theorem 3.4, when $\delta = 0$, the privacy profile of randomized response algorithm's output-distributions $P$ and $Q$ is

$$\delta_{P|Q}(t) = \begin{cases} 0 & \text{if } \varepsilon < t, \\ \frac{e^\varepsilon - e^t}{1+e^\varepsilon} & \text{if } -\varepsilon < t \le \varepsilon, \\ 1 - e^t & \text{otherwise.} \end{cases} \tag{132}$$

From the equivalence (27) of Theorem 3.3,

$$\frac{e^{(q-1)R_q(P\|Q)}}{q(q-1)} = \mathcal{B}\left\{\delta_{P|Q}(t)\right\}(1-q) \tag{133}$$

$$= \int_{-\infty}^{\infty} e^{(q-1)t}\delta_{P|Q}(t)\mathrm{d}t \tag{134}$$

$$= \int_{-\infty}^{-\varepsilon} e^{(q-1)t} \cdot (1-e^t)\mathrm{d}t + \int_{-\varepsilon}^{\varepsilon} e^{(q-1)t} \cdot \frac{e^\varepsilon - e^t}{1+e^\varepsilon}\mathrm{d}t \tag{135}$$

$$= \left[\frac{e^{(q-1)t}}{q-1} - \frac{e^{qt}}{q}\right]_{-\infty}^{-\varepsilon} + \frac{1}{1+e^\varepsilon}\left[\frac{e^\varepsilon \cdot e^{(q-1)t}}{q-1} - \frac{e^{qt}}{q}\right]_{-\varepsilon}^{\varepsilon} \tag{136}$$

$$= \left(\frac{e^{-(q-1)\varepsilon}}{q-1} - \frac{e^{-q\varepsilon}}{q}\right) + \frac{1}{1+e^\varepsilon}\left[\left(\frac{e^{q\varepsilon}}{q-1} - \frac{e^{q\varepsilon}}{q}\right) - \left(\frac{e^\varepsilon \cdot e^{-(q-1)\varepsilon}}{q-1} - \frac{e^{-q\varepsilon}}{q}\right)\right] \tag{137}$$

$$= \frac{e^{-(q-1)\varepsilon}}{q-1} \cdot \left(1 - \frac{e^\varepsilon}{1+e^\varepsilon}\right) - \frac{e^{-q\varepsilon}}{q} \cdot \left(1 - \frac{1}{1+e^\varepsilon}\right) + \frac{e^{q\varepsilon}}{1+e^\varepsilon} \cdot \left(\frac{1}{q-1} - \frac{1}{q}\right) \tag{138}$$

$$= \frac{e^{-(q-1)\varepsilon}}{1+e^\varepsilon} \cdot \left(\frac{1}{q-1} - \frac{1}{q}\right) + \frac{e^{q\varepsilon}}{1+e^\varepsilon} \cdot \left(\frac{1}{q-1} - \frac{1}{q}\right) \tag{139}$$

$$= \frac{e^\varepsilon \cdot e^{-q\varepsilon} + e^{q\varepsilon}}{1+e^\varepsilon} \cdot \frac{1}{q(q-1)}. \tag{140}$$

Therefore, for any $q \in \mathbb{R} \setminus \{0, 1\}$ we can cancel $q(q-1)$ from the denominator in both sides, which proves the theorem statement for real orders. From dominated convergence theorem the theorem statement holds for complex orders as well on corresponding real orders (cf. Section 2.2). □

### A.3 Deferred Proofs for Section 4

**Lemma A.1** ([30, Corollary 24])**.** *Let $P$ and $Q$ be probability distributions over $\Omega$. Fix $\varepsilon \ge 0$ and $\delta \in [0, 1]$. Suppose $P, Q$ satisfy $(\epsilon, \delta)$-differential privacy. Then there exists distributions $A, B, P', Q'$ over $\Omega$ such that*

$$P = (1-\delta)\frac{e^\varepsilon}{e^\varepsilon + 1}A + (1-\delta)\frac{1}{e^\varepsilon + 1}B + \delta P', \tag{141}$$

$$Q = (1-\delta)\frac{e^\varepsilon}{e^\varepsilon + 1}B + (1-\delta)\frac{1}{e^\varepsilon + 1}A + \delta Q'. \tag{142}$$

**Theorem 4.5** (Dominating Privacy Profile under $(\varepsilon, \delta)$-DP)**.** *Fix $\varepsilon \ge 0$ and $\delta \in [0, 1]$. Suppose distributions $P$ and $Q$ over $\Omega$ satisfy $(\varepsilon, \delta)$-differential privacy. Then,*

$$\forall t \in \mathbb{R} \;:\; \delta_{P|Q}(t) \le \delta_{\mathrm{RR}}(t) \quad \textit{and} \quad \delta_{Q|P}(t) \le \delta_{\mathrm{RR}}(t), \tag{143}$$

*where $\delta_{\mathrm{RR}}(t)$ is the privacy profile of the randomized response mechanism $\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}$.*

*Proof.* Since the output distributions $P$ and $Q$ are $(\varepsilon, \delta)$-differentially private, Lemma A.1 from Steinke [30] tells us that we can simulate these two distributions as post-processing of the randomized response mechanism $\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}$. To see this, imagine that $P = \mathcal{M}(D)$ and $Q = \mathcal{M}(D')$ are the output distributions of some mechanism $\mathcal{M}$. Define another mechanism $\mathcal{G} : \{0,1\} \times \{\perp, \top\} \to \Omega$ with output distribution:

$$\mathcal{G}(u) = \begin{cases} P' & \text{if } u = (0, \perp), \\ A & \text{if } u = (0, \top), \\ B & \text{if } u = (1, \top), \\ Q' & \text{if } u = (1, \perp). \end{cases} \tag{144}$$

From Lemma A.1, see that $\mathcal{G}(\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(0)) = \mathcal{M}(D)$ and $\mathcal{G}(\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}(1)) = \mathcal{M}(D')$. Therefore, by expressing $P$ and $Q$ in terms of distributions $P', A, B, Q'$, we can conclude that for all $t \in \mathbb{R}$,

$$\delta_{P|Q}(t) = \sup_{S \subset \Omega} P(S) - e^t \cdot Q(S) \tag{145}$$

$$= \sup_{S \subset \Omega} \left( \delta \cdot P'(S) + \frac{(1-\delta)(e^\varepsilon - e^t)}{e^\varepsilon + 1} \cdot A(S) + \frac{(1-\delta)(1 - e^{\varepsilon + t})}{e^\varepsilon + 1} \cdot B(S) - \delta e^t \cdot Q'(S) \right) \tag{146}$$

$$\leq \delta + \begin{cases} 0 & \text{if } \varepsilon < t, \\ \frac{(1-\delta)(e^\varepsilon - e^t)}{e^\varepsilon + 1} & \text{if } -\varepsilon < t \leq \varepsilon, \\ (1-\delta)(1-e^t) & \text{if } t \leq -\epsilon, \end{cases} \tag{147}$$

$$= \begin{cases} \delta & \text{if } \varepsilon < t, \\ 1 - \frac{(e^t+1)(1-\delta)}{e^\varepsilon + 1} & \text{if } -\varepsilon < t \leq \varepsilon, \\ 1 - e^t(1-\delta) & \text{if } t \leq -\varepsilon. \end{cases} \tag{148}$$

Note that the expression on the right is the privacy profile of the randomized response mechanism $\mathcal{M}_{\mathrm{RR}}^{\varepsilon,\delta}$. An identical bound follows for $\delta_{Q|P}(t)$ as well, with the switched roles: $P' \leftrightarrow Q'$ and $A \leftrightarrow B$. ☐

**Theorem 4.6** (Tight Composition for $(\varepsilon, \delta)$-DP)**.** *For any $\varepsilon_i \geq 0$, $\delta_i \in [0,1]$ for $i \in \{1, \cdots, k\}$, the $k$-fold composition of $(\varepsilon_i, \delta_i)$-differentially private mechanisms satisfies $(\varepsilon, \delta^{\otimes k}(\varepsilon))$-DP for all $\varepsilon$, defined recursively as*

$$\forall t \in \mathbb{R} \ : \ \delta^{\otimes l}(t) = \delta_l + \frac{(1-\delta_l)}{e^{\varepsilon_l} + 1} \left[ e^{\varepsilon_l} \cdot \delta^{\otimes l-1}(t - \varepsilon_l) + \delta^{\otimes l-1}(t + \varepsilon_l) \right], \tag{149}$$

*with $\delta^{\otimes 0}(t) = [1 - e^t]_+$.*

*Proof.* Let $P_{1:k}, Q_{1:k}$ be the joint output distributions of the $k$-fold composed mechanism on neighboring inputs. To prove the statement, we need to show that

$$\forall t \in \mathbb{R} \ : \ \delta_{P_{1:k}|Q_{1:k}}(t) \leq \delta^{\otimes}(t). \tag{150}$$

Let's define $P_i^{x_{<i}}, Q_i^{x_{<i}}$ be the output distributions of the $i^{th}$ mechanism, conditioned on the preceding $i-1$ mechanisms' output being $x_{<i}$. From Theorem 4.5 and from Theorem 4.4, we know that under adaptive $(\varepsilon_i, \delta_i)$-DP, the privacy profiles for conditional distributions are dominated as follows.

$$\forall i \in \{1, \cdots, k\} \ : \ \sup_{x_{<i} \in \Omega_1 \times \cdots \times \Omega_{i-1}} \delta_{P_i^{x_{<i}} | Q_i^{x_{<i}}}(t) \leq \delta_{\mathrm{RR}}^{\varepsilon_i, \delta_i}(t), \tag{151}$$

where $\delta_{\mathrm{RR}}^{\varepsilon_i,\delta_i}(t)$ is the privacy profile of $\mathcal{M}_{\mathbb{RR}}^{\varepsilon_i,\delta_i}$.

Using this, we prove the theorem statement inductively.

**Base step.** Let's denote the Heaveside step function as $H(t) := \mathbb{I}\{t > 0\}$. Then, we can write $\delta^{\otimes 0}(t) = (1 - H(t)) \cdot (1 - e^t)$. Using this, we can express

$$\delta^{\otimes 1}(t) = \delta_1 + \frac{1-\delta_1}{e^{\varepsilon_1}+1}\left[e^{\varepsilon_1} \cdot \delta^{\otimes 0}(t-\varepsilon_1) + \delta^{\otimes 0}(t+\varepsilon_1)\right] \tag{152}$$

$$= \delta_1 + \frac{1-\delta_1}{e^{\varepsilon_1}+1}\left[e_1^{\varepsilon} \cdot (1 - H(t-\varepsilon_1)) \cdot (1 - e^{t-\varepsilon_1}) + (1 - H(t+\varepsilon_1)) \cdot (1 - e^{t+\varepsilon_1})\right] \tag{153}$$

$$= 1 + e^x(1-\delta_1) + H(t-\varepsilon_1) \cdot \frac{(1-\delta_1)(e^t - e^{\varepsilon_1})}{e^{\varepsilon_1}+1} + H(t+\varepsilon_1) \cdot \frac{(1-\delta_1)(e^{t+\varepsilon_1}-1)}{e^{\varepsilon_1}+1} \tag{154}$$

$$= \begin{cases} \delta_1 & \text{if } \varepsilon_1 < t, \\ 1 - \frac{(e^t+1)(1-\delta_1)}{e_1^{\varepsilon}+1} & \text{if } -\varepsilon_1 < t \le \varepsilon_1, \\ 1 - e^t(1-\delta_1) & \text{if } t \le -\varepsilon_1. \end{cases} \tag{155}$$

$$= \delta_{\mathrm{RR}}^{\varepsilon_1,\delta_1}(x) \tag{156}$$

From (151), we therefore get that $\delta_{P_1|Q_1}(t) \le \delta^{\otimes 1}(t)$ for all $t \in \mathbb{R}$.

**Induction step.** Suppose for any $l \in \{2, \cdots, k\}$ the composition of first $l-1$ mechanisms have a privacy profile dominated by $\delta^{\otimes l-1}$. More precisely,

$$\forall t \in \mathbb{R} \ : \ \delta_{P_{1:l-1}|Q_{1:l-1}}(t) \le \delta^{\otimes l-1}(t). \tag{157}$$

We need to show that

$$\forall t \in \mathbb{R} \ : \ \delta_{P_{1:l}|Q_{1:l}}(t) \le \delta^{\otimes l}(t). \tag{158}$$

Recall that from the adaptive $(\varepsilon_l, \delta_l)$-DP assumption on the $l^{th}$ mechanism, (151) says that

$$\sup_{x_{<l} \in \Omega_1 \times \cdots \times \Omega_{l-1}} \delta_{P_l^{x_{<l}}|Q_l^{x_{<l}}}(t) \le \delta_{\mathrm{RR}}^{\varepsilon_l,\delta_l}(t). \tag{159}$$

Therefore, from Theorem 4.2, Theorem 4.4 and the induction assumption, we have that

$$\delta_{P_{1:l}|Q_{1:l}}(t) \le \left(\delta_{P_{1:l-1}|Q_{1:l-1}} \circledast \left(\ddot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l} - \dot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l}\right)\right)(t) \tag{160}$$

$$\le \left(\delta^{\otimes l-1} \circledast \left(\ddot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l} - \dot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l}\right)\right)(t) \tag{161}$$

$$= \int_{-\infty}^{\infty} \delta^{\otimes l-1}(t-\tau) \times \left(\ddot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l}(\tau) - \dot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l}(\tau)\right) \mathrm{d}\tau. \tag{162}$$

To differentiate properly, in a manner that handles discontinuity, we state the function $\delta_{\mathrm{RR}}^{\varepsilon,\delta}(t)$ in terms of Heaviside functions as in (154):

$$\delta_{\mathrm{RR}}^{\varepsilon,\delta}(t) = \underbrace{1 - e^t(1-\delta)}_{I_1(t)} + H(t-\varepsilon) \cdot \underbrace{\frac{(e^t - e^{\varepsilon})(1-\delta)}{e^{\varepsilon}+1}}_{I_2(t)} + H(t+\varepsilon) \cdot \underbrace{\frac{(e^{t+\varepsilon}-1)(1-\delta)}{e^{\varepsilon}+1}}_{I_3(t)}. \tag{163}$$

Then, by chain rule, its first and second derivatives are:

$$\dot{\delta}_{\mathrm{RR}}^{\varepsilon,\delta}(t) = \dot{I}_1(t) + \left(H(t-\varepsilon)\cdot \dot{I}_2(t) + \underbrace{\triangle(t-\varepsilon)\cdot I_2(t)}_{J_2(t)}\right) + \left(H(t+\varepsilon)\cdot \dot{I}_3(t) + \underbrace{\triangle(t+\varepsilon)\cdot I_3(t)}_{J_3(t)}\right) \tag{164}$$

$$\begin{aligned}\ddot{\delta}_{\mathrm{RR}}^{\varepsilon,\delta}(t) &= \ddot{I}_1(t) + \Big(H(t-\varepsilon)\cdot \ddot{I}_2(t) + \triangle(t-\varepsilon)\cdot \dot{I}_2(t) + \dot{J}_2(t)\Big)\\ &\quad + \Big(H(t+\varepsilon)\cdot \ddot{I}_3(t) + \triangle(t+\varepsilon)\cdot \dot{I}_3(t) + \dot{J}_3(t)\Big) \end{aligned}\tag{165}$$

Note that $\dot{I}_1(t) = \ddot{I}_1(t) = -e^t(1-\delta)$, $\dot{I}_2(t) = \ddot{I}_2(t) = e^t \cdot \frac{(1-\delta)}{e^\varepsilon+1}$ and $\dot{I}_3(t) = \ddot{I}_3(t) = e^t \cdot \frac{e^\varepsilon(1-\delta)}{e^\varepsilon+1}$. Therefore, on subtracting the two, a lot of terms cancel out, and we get:

$$\begin{aligned}\ddot{\delta}_{\mathrm{RR}}^{\varepsilon,\delta}(t) - \dot{\delta}_{\mathrm{RR}}^{\varepsilon,\delta}(t) &= \triangle(t-\varepsilon)\cdot(\dot{I}_2(t) - I_2(t)) + \dot{J}_2(t) + \triangle(t+\varepsilon)\cdot(\dot{I}_3(t) - I_3(t)) + \dot{J}_3(t) &(166)\\ &= \triangle(t-\varepsilon)\cdot\frac{(1-\delta)e^\varepsilon}{e^\varepsilon+1} + \dot{J}_2(t) + \triangle(t+\varepsilon)\cdot\frac{(1-\delta)}{e^\varepsilon+1} + \dot{J}_3(t). &(167)\end{aligned}$$

Note that $J_2(t) = 0$ everywhere except at $t = \varepsilon$. Similarly, $J_3(t) = 0$ everywhere except $t = -\varepsilon$.

On substituting and convolving, we get

$$\begin{aligned}\delta_{P_{1:l}|Q_{1:l}}(t) &\le \int_{-\infty}^{\infty} \delta^{\otimes l-1}(t-\tau)\times\Big(\ddot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l}(\tau) - \dot{\delta}_{\mathrm{RR}}^{\varepsilon_l,\delta_l}(\tau)\Big)\,\mathrm{d}\tau &(168)\\ &= \int_{-\infty}^{\infty} \delta^{\otimes l-1}(t-\tau)\times\left(\triangle(\tau-\varepsilon_l)\cdot\frac{(1-\delta_l)e^{\varepsilon_l}}{e^{\varepsilon_l}+1} + \triangle(\tau+\varepsilon_l)\cdot\frac{(1-\delta_l)}{e^{\varepsilon_l}+1} + \dot{J}_2(\tau) + \dot{J}_3(\tau)\right)\mathrm{d}\tau &(169)\\ &= \delta^{\otimes l-1}(t-\varepsilon_l)\cdot\frac{(1-\delta_l)e^{\varepsilon_l}}{e^{\varepsilon_l}+1} + \delta^{\otimes l-1}(t+\varepsilon_l)\cdot\frac{1-\delta_l}{e^{\varepsilon_l}+1} &(170)\\ &\quad + \underbrace{\int_{-\infty}^{\infty}\delta^{\otimes l-1}(t-\tau)\times\dot{J}_2(\tau)\mathrm{d}\tau}_{K_2(t)} + \underbrace{\int_{-\infty}^{\infty}\delta^{\otimes l-1}(t-\tau)\times\dot{J}_3(\tau)\mathrm{d}\tau}_{K_3(t)} &(171)\end{aligned}$$

For the last integral, apply integration by parts to get

$$\begin{aligned}K_2(t) &= \int_{-\infty}^{\infty}\delta^{\otimes l-1}(t-\tau)\dot{I}_2(\tau)\cdot\triangle(\tau-\varepsilon_l)\mathrm{d}\tau + \int_{-\infty}^{\infty}\delta^{\otimes l-1}(t-\tau)I_2(\tau)\cdot\dot{\triangle}(\tau-\varepsilon_l)\mathrm{d}\tau &(172)\\ &= \delta^{\otimes l-1}(t-\varepsilon_l)\dot{I}_2(\varepsilon_l) - \Big(\delta^{\otimes l-1}(t-\varepsilon_l)\dot{I}_2(\varepsilon_l) + \dot{\delta}^{\otimes l-1}(t-\varepsilon_l)I_2(\varepsilon_l)\Big) &(173)\\ &= 0 - 0 &(174)\end{aligned}$$

since $I_2(\varepsilon_l) = 0$ and for any function $f$ on $\mathbb{R}$ it holds that, $\int_{\mathbb{R}} f\cdot\dot{\triangle}\mathrm{d}\tau = -\int_{\mathbb{R}}\dot{f}\cdot\triangle\mathrm{d}\tau$. Similarly,

$$\begin{aligned}K_3(t) &= \int_{-\infty}^{\infty}\delta^{\otimes l-1}(t-\tau)\dot{I}_3(\tau)\cdot\triangle(\tau+\varepsilon_l)\mathrm{d}\tau + \int_{-\infty}^{\infty}\delta^{\otimes l-1}(t-\tau)I_3(\tau)\cdot\dot{\triangle}(\tau+\varepsilon_l)\mathrm{d}\tau &(175)\\ &= \delta^{\otimes l-1}(t+\varepsilon_l)\dot{I}_3(-\varepsilon_l) - \delta^{\otimes l-1}(t+\varepsilon_l)\dot{I}_3(-\varepsilon_l) - \dot{\delta}^{\otimes l-1}(t-\varepsilon_l)I_3(-\varepsilon_l) &(176)\\ &= 0 - 0 &(177)\end{aligned}$$

since $I_3(-\varepsilon_l) = 0$. Therefore, we have

$$\delta_{P_{1:l}|Q_{1:l}}(t) \le \frac{(1-\delta_l)}{e^{\varepsilon_l}+1}\Big[e^{\varepsilon_l}\cdot\delta^{\otimes l-1}(t-\varepsilon_l) + \delta^{\otimes l-1}(t+\varepsilon_l)\Big] \le \delta^{\otimes l}(t). \tag{178}$$

Hence, the induction statement holds. □

**Corollary 4.7.** *For any $\varepsilon \geq 0$, $\delta \in [0,1]$, the k-fold composition of $(\varepsilon, \delta)$-DP mechanisms satisfies $(\varepsilon, \delta^{\otimes k}(\varepsilon))$-DP for all $\varepsilon$, where*

$$\forall t \in \mathbb{R} \;:\; \delta^{\otimes k}(t) = 1 - (1-\delta)^k \left(1 - \mathop{\mathbb{E}}_{Y \leftarrow \text{Binomial}\left(k, \frac{e^\varepsilon}{1+e^\varepsilon}\right)} \left[1 - e^{t - \varepsilon\cdot(2Y-k)}\right]_+\right). \tag{179}$$

*Proof.* We just have to show that recurrence relationship in Theorem 4.6, that is $\forall l \in \{1, \cdots, k\}$

$$\forall t \in \mathbb{R} \;:\; \delta^{\otimes l}(t) = \delta_l + \frac{(1-\delta_l)}{e^{\varepsilon_l}+1}\left[e^{\varepsilon_l} \cdot \delta^{\otimes l-1}(t - \varepsilon_l) + \delta^{\otimes l-1}(t+\varepsilon_l)\right], \quad \text{where} \quad \delta^{\otimes 0}(t) = [1-e^t]_+, \tag{180}$$

simplifies to the theorem statement when $\varepsilon_i = \varepsilon_j = \varepsilon$ and $\delta_i = \delta_j = \delta$ for all $i, j \in \{1, \cdots, k\}$. Let's define $p = \frac{e^\varepsilon}{e^\varepsilon+1}$. The recurrence can then be stated as

$$\delta^{\otimes l}(t) = \delta + (1-\delta) \mathop{\mathbb{E}}_{Y_l \leftarrow \text{Bernoulli}(p)} \left[\delta^{\otimes l-1}(t - \varepsilon(2Y_l - 1))\right] \tag{181}$$

$$= \delta + (1-\delta) \mathop{\mathbb{E}}_{Y_l \leftarrow \text{Bernoulli}(p)} \left[\delta + (1-\delta) \mathop{\mathbb{E}}_{Y_{l-1} \leftarrow \text{Bernoulli}(p)} \left[\delta^{\otimes l-2}(t - \varepsilon(2Y_l - 1) - \varepsilon(2Y_{l-1}-1))\right]\right] \tag{182}$$

$$= \delta \sum_{i=1}^{2} (1-\delta)^{i-1} + (1-\delta)^2 \mathop{\mathbb{E}}_{\substack{Y_l \leftarrow \text{Bernoulli}(p) \\ Y_{l-1} \leftarrow \text{Bernoulli(p)}}} \left[\delta^{\otimes l-2}(t - \sum_{i=l-1}^{l} \varepsilon(2 \cdot Y_i - 1))\right] \tag{183}$$

$$= \delta \sum_{i=1}^{l} (1-\delta)^{i-1} + (1-\delta)^l \mathop{\mathbb{E}}_{\substack{Y_l \leftarrow \text{Bernoulli}(p) \\ Y_{l-1} \leftarrow \text{Bernoulli(p)}}} \left[\delta^{\otimes 0}(t - \sum_{i=1}^{l} \varepsilon(2 \cdot Y_i - 1))\right] \tag{184}$$

$$= 1 - (1-\delta)^l + (1-\delta)^l \mathop{\mathbb{E}}_{Y \leftarrow \text{Binomial}(k,p)} \left[\delta^{\otimes 0}(t - \varepsilon(2 \cdot Y - l))\right] \tag{185}$$

$$= 1 - (1-\delta)^l \left(1 - \mathop{\mathbb{E}}_{Y \leftarrow \text{Binomial}(k,p)} \left[1 - e^{t-\varepsilon\cdot(2Y-k)}\right]_+\right). \tag{186}$$

□

### A.4 Deferred Proofs for Section 5

**Theorem 5.1** (Poisson subsampling)**.** *Let $0 \leq \lambda \leq 1$. For any two distributions $P$ and $Q$ on $\Omega$,*

$$\delta_{\lambda P + (1-\lambda)|Q}(\varepsilon) = \begin{cases} \lambda \delta_{P|Q}(\log(1 + (e^\varepsilon - 1)/\lambda)) & \text{if } \varepsilon > \log(1-\lambda) \\ 1 - e^\varepsilon & \text{otherwise} \end{cases}. \tag{187}$$

*Proof.* Recall from Definition 3.1 that $\text{PLD}(Q\|P)$ and $\text{PLD}(Q\|\lambda P + (1-\lambda)Q)$ are the distributions of $L_{Q|P}(\Theta)$ and $L_{Q|P\lambda+(1-\lambda)Q}(\Theta)$ respectively with $\Theta \sim Q$. Since for any $\theta \in \Omega$,

$$L_{Q|\lambda P + (1-\lambda)Q}(\theta) = -\log \frac{\lambda P(\theta) + (1-\lambda)Q(\theta)}{Q(\theta)} = -\log(1 - \lambda + \lambda \cdot e^{-L_{Q|P}(\theta)}), \tag{188}$$

the random variables $Z'_\lambda \sim \text{PLD}(Q\|\lambda P + (1-\lambda)Q)$ and $Z' \sim \text{PLD}(Q\|P)$ are related as

$$Z'_\lambda = -\log(1 + \lambda(e^{-Z'} - 1)). \tag{189}$$

Using Theorem 3.2 that we can express the privacy profile as:

$$\begin{aligned}
\delta_{\lambda P+(1-\lambda)Q|Q}(\varepsilon) &= e^{\varepsilon}\mathcal{L}\left\{F_{Z'_\lambda}(-t-\varepsilon)\right\}(-1) && (190)\\
&\overset{(73)}{=} e^{\varepsilon}\left[\mathcal{L}\left\{f_{Z'_\lambda}(-t-\varepsilon)\right\}(-1) - F_{Z'_\lambda}(\varepsilon^-)\right] && (191)\\
&= e^{\varepsilon}\left[\int_{0^+}^{\infty} e^t f_{Z'_\lambda}(-t-\varepsilon)dt - \int_{-\infty}^{0^-} f_{Z'_\lambda}(t-\varepsilon)dt\right] && (192)\\
&= e^{\varepsilon}\int_{-\infty}^{0^-}(e^{-t}-1)f_{Z'_\lambda}(t-\varepsilon)dt && (193)\\
&= \int_{-\infty}^{-\varepsilon^-}\left(e^{-t'}-e^{\varepsilon}\right)f_{Z'_\lambda}(t')dt' && (194)\\
&= \mathbb{E}\left[e^{-Z'_\lambda}-e^{\varepsilon}\right]_+, && (195)
\end{aligned}$$

where $[x]_+ := \max\{0, x\}$. On substituting $Z'_\lambda$, we get:

$$\begin{aligned}
\delta_{\lambda P+(1-\lambda)Q|Q}(\varepsilon) &= \mathbb{E}\left[e^{-Z'_\lambda}-e^{\varepsilon}\right]_+ && (196)\\
&= \mathbb{E}\left[1-\lambda+\lambda e^{-Z'}-e^{\varepsilon}\right]_+ && (197)\\
&= \lambda\mathbb{E}\left[e^{-Z'}-\frac{e^{\varepsilon}+\lambda-1}{\lambda}\right]_+ && (198)\\
&= \begin{cases} \lambda\mathbb{E}\left[e^{-Z'}-e^{\log(1+(e^{\varepsilon}-1)/\lambda)}\right]_+ & \text{if } \varepsilon > \log(1-\lambda)\\ \lambda\mathbb{E}\left[e^{-Z'}\right]+1-\lambda-e^{\varepsilon} & \text{otherwise}\end{cases} && (199)\\
&= \begin{cases} \lambda\delta_{P|Q}(\log(1+(e^{\varepsilon}-1)/\lambda)) & \text{if } \varepsilon > \log(1-\lambda)\\ 1-e^{\varepsilon} & \text{otherwise}\end{cases}. && \left(\because \mathbb{E}\left[e^{-Z'}\right]=1\right)
\end{aligned}$$

□

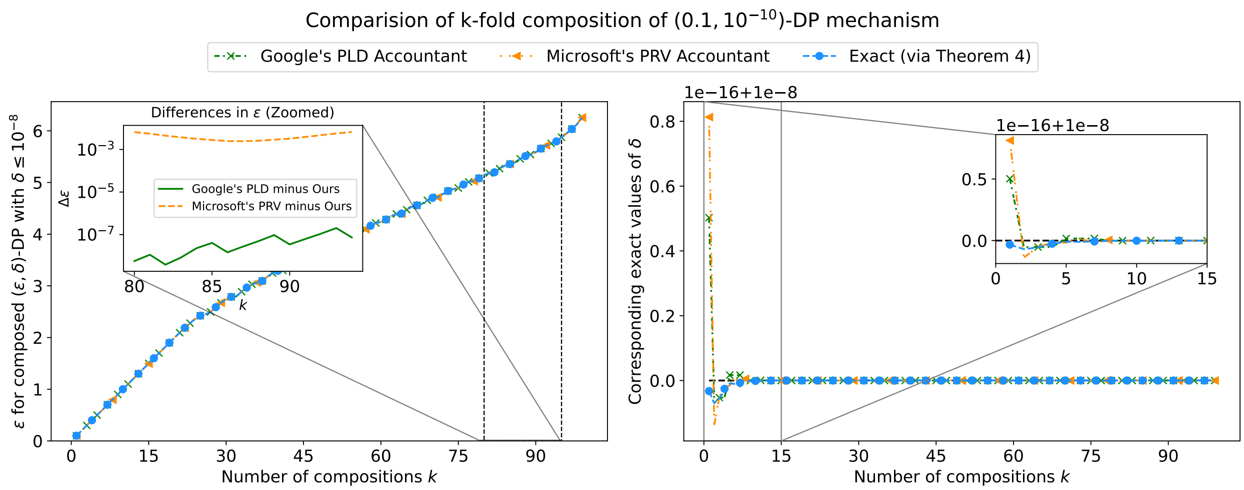



Figure 4: Comparison of $(\varepsilon, \delta)$-DP bounds between numerical accountants and our Theorem 4.7 for 100-fold composition of a $(0.1, 10^{-10})$-DP point guarantee, with the budget constraint $\delta < 10^{-8}$. Note that at $k = 100$, the constraint on $\delta$ cannot be satisfied and so the corresponding $\varepsilon = \infty$ at that value. We note that at smaller values of $k$, numerical accountants can sometimes over exceed the budget constraints on $\delta$. Additionally, the gap for $\varepsilon$ between our exact bound and those approximated by numerical accountant tend to be of the order $\approx 10^{-7}$ for Google's PLDAccountant and $\approx 10^{-3}$ for Microsoft's PRVAccountant.