INFINITELY DIVISIBLE PRIVACY AND BEYOND I RESOLUTION OF THE $s^2 = 2k$ CONJECTURE

Aaradhya Pandey[1], Arian Maleki[2], Sanjeev Kulkarni[1]
[1]Princeton University   [2] Columbia University

ABSTRACT

Differential privacy is increasingly formalized through the lens of hypothesis testing via the *robust* and *interpretable f*-DP framework, where privacy guarantees are encoded by a baseline *Blackwell* trade-off function $f_\infty = T(P_\infty, Q_\infty)$ involving a pair of distributions (P_∞, Q_∞) Dong et al. (2022). The problem of '*choosing the right privacy metric in practice*' Cummings & collaborators (2024) gives rise to a central question in this framework: *what is a statistically appropriate baseline f_∞ given some prior modeling assumptions*? The special case of Gaussian differential privacy (GDP) Dong et al. (2022) showed that, under compositions of nearly perfect DP mechanisms, these baseline trade-off functions exhibit a central limit behavior: $f_n = T(P_n^{\otimes n}, Q_n^{\otimes n}) \to f_\infty = T(P_\infty, Q_\infty)$, where the limiting baseline trade-off function f_∞ involves two shifted Gaussians P_∞, Q_∞. Inspired by Le Cam's theory of limits of statistical experiments Le Cam (1986), *we answer this question in full generality in its natural infinitely divisible setting.*

We show that the sequence of composition experiments $(P_n^{\otimes n}, Q_n^{\otimes n})$ '*converge*' in general to a binary limit experiment (P_∞, Q_∞) whose log-likelihood ratio $\log \frac{dQ_\infty}{dP_\infty}$ is an infinitely divisible distribution under the limiting null P_∞. So, any such limiting trade-off function f_∞ involves an infinitely divisible law P_∞ characterized by its *Lévy–Khintchine* triplet, and its *Esscher* tilt $dQ_\infty(x) = e^x dP_\infty(x)$. As a consequence, we determine all the limiting baseline trade-off functions f_∞ under the composition of nearly perfect differentially private procedures. Our framework recovers GDP as the purely Gaussian case, yields explicit many non-Gaussian limits including Poisson, and many others. Consequently, we *positively* resolve the $s^2 = 2k$ conjecture proposed in the GDP paper Dong et al. (2022), which curiously observed that in the Gaussian examples $P_\infty \overset{d}{=} -k + |s|Z$ with $Z \sim N(0,1)$, we always have $s^2 = 2k$. Then, we describe situations where one can goes beyond the infinite divisibility framework, such as $\log \frac{dQ_\infty}{dP_\infty}$ being a Gaussian mixture under P_∞. Finally, we describe a mechanism for count statistics that optimally achieves (asymmetric) Poisson differential privacy, going beyond the natural noise-adding mechanisms used in practice.

1 INTRODUCTION, PREVIOUS WORK, AND OUR CONTRIBUTIONS

Modern data analysis routinely draws on information about people: their location traces, search histories, media consumption, social graphs, alongside long-standing, highly sensitive records from hospitals, schools, and national censuses. This necessarily demands a framework of data analysis that preserves the privacy of individuals. Early *privacy* practices leaned on ad-hoc anonymization, but a series of high-profile re-identifications made it clear that simple de-identification is brittle at scale Narayanan & Shmatikov (2008). This motivated a rigorous foundation for privacy in statistical analysis. Differential privacy (DP) provides exactly that: a principled, mathematically verifiable guarantee against disclosure risks that persists under arbitrary post-processing and composition. Since its introduction, differential privacy has grown into a practical standard, with widespread industrial adoption Apple Inc. (2017) and deployment in official statistics Abowd (2018).

1.1 FROM (ε, δ) AND DIVERGENCE BASED DP TO HYPOTHESIS TESTING BASED DP

In its classical form, differential privacy formalizes protection via two real-valued parameters $\varepsilon \geq 0$, $0 \leq \delta \leq 1$, which jointly bound the advantage of any adversary attempting to distinguish between neighboring datasets from a mechanism's output Dwork et al. (2006a;b). The (ε, δ) formulation of differential privacy became the standard due to its robustness to post-processing, group privacy, and composability, but sharp accounting under composition proved to be technically delicate Dwork & Roth (2014). A line of *divergence-based* relaxations was subsequently developed to tighten composition analyses and enable modular accounting: concentrated DP (CDP and zCDP) Dwork & Rothblum (2016) Bun & Steinke (2016), and Rényi DP (RDP) Mironov (2017), with follow-ups on amplification by subsampling and analytic accountants Wang et al. (2019); Abadi et al. (2016).

In parallel, a practical, robust, and interpretable *hypothesis-testing* viewpoint of differential privacy was formalized, starting with Wasserman & Zhou (2010) and Kairouz et al. (2015). It describes privacy as the hypothesis testing problem of distinguishing between two neighboring datasets S, S' via the output distributions $M(S), M(S')$ of a private mechanism M. This was formalized in the framework of f-differential privacy (f-DP) Dong et al. (2022), which describes privacy of M by requiring that the optimal type I vs type II error curve for the distinguishing test for $M(S), M(S')$, for any two neighboring datasets, S and S', stays above a baseline *trade-off function* (TOF) f .

Immediately, the f-DP framework is statistically interpretable via hypothesis testing. It recovers the classical (ε, δ)-DP formulation for a certain choice of baseline trade-off function $f_{\varepsilon, \delta}$ Wasserman & Zhou (2010)[1], and it is a robust framework because it uniformly preserves privacy at all scales of type I error $0 \leq \alpha \leq 1$. Moreover, it is achievable by appropriate noise injection mechanisms Awan & Vadhan (2023)[2], and it is closed and tight under composition, thereby allowing for lossless reasoning about composition Dong et al. (2022). Finally, Dong et al. (2022) introduced Gaussian differential privacy (GDP) as a one-parameter family of baseline trade-off curves G_μ that require distinguishing between a pair of shifted normal distributions $N\left(-\frac{\mu^2}{2}, \mu^2\right)$ and $N\left(\frac{\mu^2}{2}, \mu^2\right)$ with the same variance μ^2 for a real-valued parameter $\mu \in \mathbb{R}^3$ and showed that it optimally captures the widely used Gaussian mechanism. Dong et al. (2022) further showed that under compositions of nearly perfect DP mechanisms $M_n \otimes \cdots \otimes M_1$, these baseline trade-off functions $f_n \otimes \cdots \otimes f_1$ converge to a limiting baseline trade-off function G_μ involving two shifted Gaussians as above.

1.2 OUR CONTRIBUTION: IN SEARCH OF A BASELINE TRADE-OFF FUNCTION

Dong et al. (2022) established a central limit behavior within differential privacy, indicating that the baseline trade off function f should be chosen as a trade off function $T(P_\infty, Q_\infty)$ (see definition 1) for a shifted Gaussian pair $P_\infty = N\left(-\frac{\mu^2}{2}, \mu^2\right)$ and $Q_\infty = N\left(\frac{\mu^2}{2}, \mu^2\right)$, for some $\mu \in \mathbb{R}$ under a larger number of (nearly perfect) differentially private outputs based on queries on the same dataset.

- **An infinitely divisible framework of privacy.** In this paper, we generalize this result in its natural infinitely divisible setting (see Theorem 6) and establish that under a large number of (nearly perfect) differentially private operations, if the deteriorated trade-off curves converge to a limiting trade-off curve, then the limiting trade-off curve has an infinitely divisible form (see Equation (18)), encompassing both Gaussian and Poisson components[4].

- **Resolution of the $s^2 = 2k$ conjecture.** As an application of Theorem 6, we resolve the $s^2 = 2k$ conjecture (see Theorem 3) proposed in Dong et al. (2022), which observed that in the Gaussian case $P_\infty \stackrel{d}{=} -k + |s|Z$ with $Z \sim N(0, 1)$, we always have $s^2 = 2k$.

---

[1]The expression of $f_{\varepsilon, \delta}$ is given by $f_{\varepsilon, \delta}(\alpha) = \max\left(0, 1 - \delta - e^\varepsilon \alpha, e^{-\varepsilon}(1 - \delta - \alpha)\right)$ for $\alpha \in [0, 1]$.

[2]Although very general, the noise adding mechanisms described here are not quite naturally implementable.

[3]From the basic symmetries of Gaussian trade-off functions G_μ Pandey et al. (2025), our description of G_μ is equivalent to the usual description of itself involving the pair of distributions $N(0, 1)$ and $N(\mu, 1)$. However, our formulation will make the resolution of the $s^2 = 2k$ conjecture of Dong et al. (2022) appear naturally.

[4]Inspired by Le Cam & Yang (2000) 'The gaussian ones are used everywhere because of mathematical tractability. Poisson experiments are less tractable and less studied, but they are probably more important. They will loom large in the new century.', and given that we are in the next century, we show how the Poisson baselines are extremely relevant for differentially private analysis on graph based relational datasets.

- **Random number of operations and beyond infinite divisibility.** In practice, we do not know the number of differentially private operations one is going the implement on the same dataset. In section 3.1 (see Equation (64)) we derive that under a large but random number of (nearly perfect) differentially private operations, the limiting trade-off functions can go beyond the class of infinitely divisible trade-off functions. Inspired by Le Cam (1986), we demonstrate this with a locally asymptotically mixed normal example 66.
- **Achieving Poisson differential privacy in practice.** In Theorem 13 , we describe a mechanism that optimally achieves (asymmetric) Poisson differential privacy real-valued statistics, including count statistics with baseline privacy $f = T(P(\lambda_1), P(\lambda_2))$ going beyond the popular noise -adding mechanisms.
- **Applications to graph-based relational datasets.** We demonstrate how Poisson differential privacy becomes extremely relevant in tightly capturing privacy on datasets that are relational or graph-based and only allow released statistics to be inherently discretized 4.
- **A coarsened Neyman-Pearson lemma.** Inspired by a Bayesian perspective of differential privacy Strack & Yang (2024), we prove (see Theorem 22) exactly how the trade-off functions $T(P, Q)$ on $(\Omega, \mathcal{F})$ degrade when we shrink the underlying the σ-algebra of information from $\mathcal{F}$ to $\mathcal{G}$, and demonstrate with a Gaussian example[5] (see Proposition 2).

2 HYPOTHESIS TESTING-BASED PRIVACY: DEFINITIONS AND MOTIVATIONS

We describe objects of the hypothesis-testing-based privacy framework and motivate our problem.

2.1 BASICS OF TRADE-OFF FUNCTIONS AND BLACKWELL ORDERING

Definition 1 (Trade-off function Dong et al. (2022)). *Given two probability distributions P, Q on a measurable space $(\mathcal{W}, \mathcal{F}_{\mathcal{W}})$, we define the* trade-off function *as the map $= T(P, Q) : [0, 1] \to [0, 1]$*

$$T(P,Q)(\alpha) := \inf_{\varphi}\Big\{\beta_\varphi := \mathbb{E}_Q[1 - \varphi] \ \Big| \ \alpha_\varphi := \mathbb{E}_P[\varphi] \le \alpha, \ \varphi : (\mathcal{W}, \mathcal{F}_{\mathcal{W}}) \to [0, 1] \text{ measurable}\Big\}. \tag{1}$$

Intuition, the Neyman-Pearson optimizer. For any type I error $\alpha \in [0, 1]$, the trade-off function (TOF) returns the smallest value of type II error β_φ over all possible Borel measurable test functions $\varphi : \mathcal{W} \to [0, 1]$. As a consequence of Neyman-Pearson lemma Polyanskiy & Wu (2025)[Thm 14.11] $T(P, Q)(\alpha)$ is achieved by the following likelihood ratio test for every $\alpha \in [0, 1]$

$$\varphi^*(w) := \mathbf{1}\left(\frac{dQ}{d\mu} > \tau^* \frac{dP}{d\mu}\right) + \lambda^* \mathbf{1}\left(\frac{dQ}{d\mu} = \tau^* \frac{dP}{d\mu}\right), \quad \frac{dQ}{dP} := \frac{dQ}{d\mu} / \frac{dP}{d\mu} : (\Omega, \mathcal{F}) \to [0, \infty] \tag{2}$$

for $2\mu = P + Q$, are the ratio of Radon–Nikodym derivatives, and τ^*, λ^* are determined as follows. Choose $\tau^* \in [0, \infty]$ as the unique number[6] such that $P\left(\frac{dQ}{dP} \ge \tau^*\right) \ge \alpha \ge P\left(\frac{dQ}{dP} > \tau^*\right)$, and

$$\text{define } \lambda^* := \frac{\alpha - P\left(\frac{dQ}{dP} > \tau^*\right)}{P\left(\frac{dQ}{dP} \ge \tau^*\right) - P\left(\frac{dQ}{dP} > \tau^*\right)} \mathbf{1}\left(\alpha - P\left(\frac{dQ}{dP} > \tau^*\right) > 0\right) \text{ to have } \mathbb{E}_P[\varphi^*] = \alpha.$$

Characterization of TOFS and their inverse. Dong et al. (2022)[Prop 1] proved that the collection of all trade-off functions $\mathcal{T} := \{T(P, Q) : (P, Q) \text{ on some measurable } (\mathcal{W}, \mathcal{F}_{\mathcal{W}})\}$ is the same as[7]

$$\mathcal{T} = \{f : [0, 1] \to [0, 1] \text{ convex, continuous, decreasing and } f(\alpha) \le 1 - \alpha \text{ for all } \alpha \in [0, 1]\} \tag{3}$$

Moreover, any such $f \in \mathcal{T}$ can be realized as $T(P, Q)$ on $\mathcal{W} = [0, 1], \mathcal{F}_{\mathcal{W}} = \mathcal{B}_{[0,1]}$ by taking P to be the standard uniform probability measure U on $([0, 1], \mathcal{B}_{[0,1]})$ and $Q = Q_f$ to be a Borel probability

---

[5]The results extend with appropriate modifications for any symmetric log concave shift family on $\mathbb{R}$.

[6]Such a τ^* uniquely exists as the CDF $F : \mathbb{R} \to [0, 1]$, defined as $F(t) := P\left(\frac{dQ}{dP} \le t\right)$ is monotone increasing as a function of t satisfying $(t_1 \le t_2 \to F(t_1) \le F(t_2))$, right continuous $(\lim_{t \downarrow t_0} F(t) = F(t_0))$ with $F(0^-) = P_0\left(\frac{dQ}{dP} < 0\right) = 0$ (lower limit of α) and $\lim_{t \to \infty} F(t) = 1$ (upper limit of α).

[7]As argued in Torgersen (1991) and is also immediate that the set $\mathcal{T}$ says the same, if we keep all the conditions intact, except one, where we replace the condition $f(x) \le 1 - x$ for all $x \in [0, 1]$ by just $f(1) = 0$.

measure on $[0,1]$ with cumulative distribution function $Q([0,x]) = f(1-x)$ for $x \in [0,1)$ and having an atom at $x = 1$ of mass $Q(\{1\}) = 1 - f(0)$ whenever $f(0) < 1$. Finally, if $f = T(P,Q)$ then its generalized inverse is a trade off function and $f^{-1} = T(Q,P)$ Dong et al. (2022)[Lem A.2]

$$\text{where } f^{-1} : [0,1] \to [0,1] \text{ defined as } f^{-1}(\alpha) = \inf\{t \in [0,1] : f(t) \leq \alpha\} \text{ for } \alpha \in [0,1]. \quad (4)$$

Functional ordering and indistinguishability (privacy). *One can define a functional ordering*

$$\textbf{\textit{Blackwell ordering:}} \quad T(P',Q')(\alpha) \geq T(P,Q)(\alpha) \text{ for all } \alpha \in [0,1] \quad (5)$$

capturing the intuition that the pair of distributions (P',Q') defined on some space $(\mathcal{W}', \mathcal{F}_{\mathcal{W}'})$ is more difficult to distinguish from a sample than the pair (P,Q) defined on another space $(\mathcal{W}, \mathcal{F}_{\mathcal{W}})$, uniformly across every scale $0 \leq \alpha \leq 1$. We refer to Torgersen (1991) for a comprehensive treatment of equivalent conditions for this ordering also known as Blackwell-Le-Cam equivalence.

Blackwell ordering. *Further, the functional ordering $f \geq g$ for $f, g \in \mathcal{T}$ (a partial ordering on $\mathcal{T}$ since not all pairs are comparable) described above is equivalent to the Blackwell ordering Dong et al. (2022)[Thm 2]. More precisely, $T(P',Q')(\alpha) \geq T(P,Q)(\alpha)$ for all $\alpha \in [0,1]$ if and only if there exists a Markov kernel $R : \mathcal{W} \to \mathcal{W}'$ Kallenberg (2021) such that $(P',Q') = (R(P), R(Q))$.*

2.2 BASICS OF f-DIFFERENTIAL PRIVACY, MOTIVATION OF OUR WORK

In this subsection, we introduce the f-differential privacy framework of Dong et al. (2022) based on Blackwell ordering and specialize to Gaussian differential privacy to motivate our generalization. Along the way, inspired by Torgersen (1991), we also state an equivalent way of describing the f-differential privacy of Dong et al. (2022) through their minimum Bayes risks.

Definition 2 (*f-differential privacy Dong et al. (2022)*). *For $k \in \mathbb{Z}_{\geq 1}$, given an input space $(\mathcal{X}, \mathcal{F}_{\mathcal{X}})$ and an output space $(\mathcal{Y}, \mathcal{F}_{\mathcal{Y}})$, a Markov kernel (randomized mechanism) $M : \mathcal{X}^k \to \mathcal{Y}$ is said to satisfy f-DP for some trade off function $f \in \mathcal{T}$ if minimizing over neighboring datasets $S, S' \in \mathcal{X}^k$[8]*

$$\inf_{d_H(S,S') \leq 1} T\big(M(S), M(S')\big)) \geq f \quad \text{pointwise on } [0,1]. \quad (6)$$

k **dependence, intuition, post processing.** For different values of k, one has different Markov kernels $M(k)$ and can allow different baseline trade off functions $f_k \in \mathcal{T}$. Let $f = T(P,Q)$ for probability measures P, Q on a space $(\mathcal{W}, \mathcal{F}_{\mathcal{W}})$. Then the above definition captures the intuition that distinguishing the outputs of M for any neighboring $S, S' \in \mathcal{X}^k$ is at least as hard as distinguishing between a sample of P or Q. Moreover, if $M : \mathcal{X}^k \to \mathcal{Y}$ is f-DP and $R : \mathcal{Y} \to \mathcal{Z}$ is a Markov kernel independent of the data given M, then $R \circ M : \mathcal{X}^k \to \mathcal{Z}$ is also f-DP Dong et al. (2022).

Convex duality and connections to (ε, δ)-DP and others. It follows from definition that if a Markov kernel M satisfy f_i DP for all $i \in I$, then it satisfy $\sup_{i \in I} f_i$ DP. Moreover, Wasserman & Zhou (2010) established that satisfying (ε, δ)-DP in the classical sense is equivalent to satisfying $f_{\varepsilon,\delta}$-DP in the sense above, with $f_{\varepsilon,\delta}(\alpha) = \max(0, 1 - \delta - e^\varepsilon \alpha, e^{-\varepsilon}(1 - \delta - \alpha))$ for $\alpha \in [0,1]$. As a consequence of the two above and convex duality[9], for a *symmetric* trade off function $f = f^{-1}$ we have a Markov kernel M satisfy f-DP if and only if it satisfy $(\varepsilon, \delta(\varepsilon))$-DP in the classical sense, for all $\varepsilon \geq 0$ with $\delta(\varepsilon) = 1 + f^*(-e^\varepsilon)$, where $f^*(y) = \sup_{\alpha \in [0,1]}(\alpha y - f(\alpha))$ is the convex conjugate of f. We refer to Dong et al. (2022)[Appendix B] for the conversion of f-DP to divergence-based DP. Inspired by Torgersen (1991), one can derive a duality relation between the trade-off curve

$$f(\alpha) = T(P,Q)(\alpha) = \inf_{\varphi} \Big\{ \mathbb{E}_Q[1-\varphi] \ \Big| \ \mathbb{E}_P[\varphi] \leq \alpha, \ \varphi : (\mathcal{W}, \mathcal{F}_{\mathcal{W}}) \to [0,1] \text{ measurable} \Big\},$$

and the minimum Bayes risk for the testing problem (P,Q) with prior $(1-\lambda, \lambda)$ for $\lambda \in [0,1]$

$$b(\lambda) = B(P,Q)(\alpha) = \min_{\varphi} \Big\{ (1-\lambda)\mathbb{E}_P[\varphi] + \lambda \mathbb{E}_Q[1-\varphi] \ \Big| \ \varphi : (\mathcal{W}, \mathcal{F}_{\mathcal{W}}) \to [0,1] \text{ measurable} \Big\}. \quad (7)$$

---

[8]Instead of the product metric space $(\mathcal{X}^k, d_H)$ with the Hamming metric $d_H(\underline{x}, \underline{y}) = \sum_{i=1}^k \mathbf{1}(x_i \neq y_i)$, one can consider any metric space $(\mathcal{X}, d)$ with the appropriate change in definitions. See 12 for more.

[9]A convex, continuous $f : [0,1] \to [0,1]$ is the pointwise supremum of all affine functions lying below it.

The precise duality relation for the pair of functions $f, b : [0, 1] \to [0, 1]$ is given by[10]

$$f(\alpha) = \sup_{0 < \lambda \leq 1} \left[\frac{b(\lambda) - (1 - \lambda)\alpha}{\lambda} \right], \quad b(\lambda) = \inf_{0 \leq \alpha \leq 1} \left[(1 - \lambda)\alpha + \lambda f(\alpha) \right]. \tag{8}$$

The equivalence between $f \leftrightarrow b$ gives rise to the an enumeration of all such functions $\{ B(P, Q) : [0, 1] \to [0, 1] : (P, Q) \text{ probability measures on some } (\Omega, \mathcal{F}) \}$ as the following class.

$$\mathcal{B} := \{ b : [0, 1] \to [0, 1] \quad | \quad b \text{ concave and } b(\lambda) \leq \min(\lambda, 1 - \lambda) \text{ for all } \lambda \in [0, 1] \} \tag{9}$$

Moreover, Blackwell ordering $f_1 \geq f_2$ for $f_1, f_2 \in \mathcal{T}$ can be equivalently stated in terms of their corresponding minimum Bayes risks $b_1 \geq b_2$, pointwise on $[0, 1]$. So, the entire framework of f-differential privacy proposed in Dong et al. (2022) can also be written equivalently in terms of the minimum Bayes risk function b. However, working with b has the advantage that b is affine under mixtures of experiments. More precisely, given pairs of binary (Blackwell-Le-Cam) experiments $\mathcal{E}_1 = (\Omega_1, \mathcal{F}_1, P_1, Q_1)$ and $\mathcal{E}_2 = (\Omega_2, \mathcal{F}_2, P_2, Q_2)$, consider the mixture experiment[11]

$$\delta \mathcal{E}_1 + (1 - \delta)\mathcal{E}_2 := (\Omega_1 \bigsqcup \Omega_2, \sigma(\mathcal{F}_1 \sqcup \mathcal{F}_2), \delta P_1 + (1 - \delta)P_2, \delta Q_1 + (1 - \delta)Q_2). \tag{10}$$

$$\text{Then } b(\delta P_1 + (1 - \delta)P_2, \delta Q_1 + (1 - \delta)Q_2) \equiv \delta b(P_1, Q_1) + (1 - \delta)b(P_2, Q_2). \tag{11}$$

This affineness of the functional b is of importance in the sense that one can restrict the analysis of experiments that are only supported on $(\Omega, \mathcal{F}) \equiv (\{0, 1\}, 2^{\{0,1\}})$. It would be interesting to see the consequences of this in the practice of formalizing differential privacy through the b function.

G_μ-DP and motivation of our work. A central question in this hypothesis testing-based f-DP framework is to choose a statistically appropriate baseline trade-off function f depending on the situation at hand. As a consequence of a central limit behavior, the focus of the paper Dong et al. (2022) was on the special case of a mechanism M satisfying G_μ-DP by letting the baseline trade off function $f = G_\mu = T(P, Q)$ where $P = N\left(-\frac{\mu^2}{2}, \mu^2\right)$ and $Q = N\left(\frac{\mu^2}{2}, \mu^2\right)$ for some $\mu \in \mathbb{R}$. Our motivation is to search for a universal list of baseline TOFs $f = T(P, Q)$ beyond the Gaussian one.

2.3 COMPOSITION AND BEYOND THE UNIVERSAL CENTRAL LIMIT BEHAVIOUR

Now, our interests are in quantifying how privacy degrades when we compose a sequence of mechanisms $M_1, \cdots, M_n$ on a (fixed) private data set $S \in \mathcal{X}^k$ in which each analysis is informed by prior analyses on the same data set, and more and more analysis is publicly released with increasing n. For simplicity, we keep the σ-algebras and the measurability conditions implicit in definitions.

Definition 3 (Composition of private mechanisms Dong et al. (2022), Kallenberg (2021)). *Consider a Markov kernel $M_1 : \mathcal{X}^k \to \mathcal{Y}_1$, and conditional on the following pair (data, first output) = $(S, M_1(S))$, let $M_2 : \mathcal{X}^k \times \mathcal{Y}_1 \to \mathcal{Y}_2$ be another Markov kernel. The joint Markov kernel (randomized mechanism) $M = (M_1, M_2) : \mathcal{X}^k \to \mathcal{Y}_1 \times \mathcal{Y}_2$ is defined is $M(S) = (M_1(S), M_2(S, M_1(S)))$.*

Intuition and n step generalization. Given the dataset S, the distribution of output of the joint mechanism $M(S)$ is constructed from the marginal distribution of $M_1(S)$ on Y_1 and the conditional distribution of $M_2(S, y_1)$ on Y_2 given $M_1(S) = y_1$. More generally, given a sequence of Markov kernels $M_i : \mathcal{X}^k \times \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_{i-1} \to \mathcal{Y}_i$, we can recursively define the joint mechanism as their composition $M = (M_1, \cdots, M_n) : \mathcal{X}^k \to \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_n$. Now, with increasing n, we publicly release more and more analysis of the dataset S, first $M_1(S)$ then $M_2(S, M_1(S))$, then $M_3(S, M_1(S), M_2(S, M_1(S)))$, and continue. The key question of how the privacy degrades as the number of analyses increases, namely under composition leads us to the following definition.

Definition 4 (Tensor product Dong et al. (2022)). *Consider trade off functions $f = T(P, Q)$ for probability measures (P, Q) on some space $(\mathcal{W}, \mathcal{F}_\mathcal{W})$ and $g = T(P', Q')$ for probability measures (P', Q') on some space $(\mathcal{W}', \mathcal{F}_{\mathcal{W}'})$. We define the tensor product trade off function as $f \otimes g := T(P \otimes P', Q \otimes Q')$, where $(P \otimes P', Q \otimes Q')$ are the product measures on $(\mathcal{W} \times \mathcal{W}', \mathcal{F}_\mathcal{W} \otimes \mathcal{F}_{\mathcal{W}'})$.*

---

[10]This representation writes out f as a suprema over linear functions $f(\alpha) = \sup_{0 < \lambda \leq 1} \left[\frac{b(\lambda)}{\lambda} - \frac{(1-\lambda)}{\lambda}\alpha \right]$.

[11]The intuition is that in practice the mixture experiment is performed by deciding with probabilities $(\delta, 1 - \delta)$ whether to perform a sample from $\mathcal{E}_1$ or $\mathcal{E}_2$ and then a sample is drawn. The output of the experiment has the form (I, X_I), where I is the index $\{1, 2\}$ of which experiment is performed and X_I is the sample itself.

Well-definedness and basic properties. By definition $f \otimes g$ is a trade-off function. Dong et al. (2022)[Lem C.2] showed that $f \otimes g$ is well defined in the sense that if $f = T(P, Q) = T(\tilde{P}, \tilde{Q})$, then $T(P \otimes P', Q \otimes Q') = T(\tilde{P} \otimes P', \tilde{Q} \otimes Q')$. Moreover, we have Dong et al. (2022)[Prop D.1, D.2] *monotonicity:* $g_1 \geq g_2$ implies $f \otimes g_1 \geq f \otimes g_2$, *commutativity:* $f \otimes g = g \otimes f$, *associativity:* $(f \otimes g) \otimes h = f \otimes (g \otimes h)$, *identity:* $f \otimes I = I \otimes f = f$ with $I(\alpha) = 1 - \alpha$ for $\alpha \in [0, 1]$, *inverse:* $(f \otimes g)^{-1} = f^{-1} \otimes g^{-1}$, with a *trivial limit:* for any trade off function $\mathcal{T} \ni f \neq I$, we have $\lim_{n \to \infty} f^{\otimes n}(\alpha) = 0$ for all $\alpha \in (0, 1]$ as a consequence of Polyanskiy & Wu (2025)[Rem 7.6]. Now, we state the celebrated composition theorem that tightly quantifies the privacy degradation of sequential mechanisms with the tensor product construction of trade off functions.

Theorem 1 (Tightest composition theorem). *Dong et al. (2022)[Thm 4] For $1 \leq i \leq n$ consider Markov kernels M_i as above so that $M_i(\cdot, y_1, \cdots, y_{i-1}) : \mathcal{X}^k \to \mathcal{Y}_i$ satisfy f_i-DP for all $y_1 \in \mathcal{Y}_1, \cdots, y_{i-1} \in \mathcal{Y}_{i-1}$. Then the joint mechanism $M = (M_1, \cdots, M_n)$ satisfy $f_1 \otimes \cdots \otimes f_n$-DP.*

Before stating a central limit behavior of nearly perfect private mechanisms as the number of compositions $n \uparrow \infty$, we need to define a few moment functionals of the log-likelihood ratio (LLR) of P and Q such that $f = T(P, Q)$. The functionals $\mathrm{kl}(f) := -\int_0^1 \log |f'(x)| dx$, $\kappa_2(f) := \int_0^1 \log^2 |f'(x)| dx$, $\kappa_3(f) := \int_0^1 \left| \log |f'(x)| \right|^3 dx$, $\overline{\kappa}_3(f) := \int_0^1 \left| \log |f'(x)| + \mathrm{kl}(f) \right|^3 dx$ (this one requires $\mathrm{kl}(f) < \infty$) take values in $[0, +\infty]$. Now, consider a triangular array of Markov kernels (randomized mechanisms) $\{M_{n1}, \ldots, M_{nn}\}_{n=1}^\infty$ where M_{ni} satisfy f_{ni}-DP for $1 \leq i \leq n$ 2.

Theorem 2. *Dong et al. (2022)[Thm 6] Let $\{f_{ni} : 1 \leq i \leq n\}_{n=1}^\infty$ be a triangular array of symmetric $(f = f^{-1})$ trade-off functions and assume for constants $k \geq 0$ and $s > 0$ as $n \to \infty$ we have $\sum_{i=1}^n \mathrm{kl}(f_{ni}) \to k, \max_{1 \leq i \leq n} \mathrm{kl}(f_{ni}) \to 0, \sum_{i=1}^n \kappa_2(f_{ni}) \to s^2, \sum_{i=1}^n \kappa_3(f_{ni}) \to 0$. Then,*

$$\lim_{n \to \infty} f_{n1} \otimes \cdots \otimes f_{nn}(\alpha) = G_{2k/s}(\alpha) \text{ uniformly for all } \alpha \in [0, 1]. \tag{12}$$

Limitation of G_μ-DP and a non-Gaussian (Poisson) limit illustration. The above theorem shows (under its assumptions) that the composition of nearly perfect DP mechanisms $M_{n1} \otimes \cdots \otimes M_{nn}$ asymptotically satisfy $G_{2k/s}$-DP, where $G_\mu = T\left(N\left(-\frac{\mu^2}{2}, \mu^2\right), N\left(\frac{\mu^2}{2}, \mu^2\right)\right)$. But, consider the baseline TOFs $f_{ni} = T\left(\mathrm{Ber}\left(\frac{\lambda_1}{n}\right), \mathrm{Ber}\left(\frac{\lambda_2}{n}\right)\right)$ for $i \in [n]$, then it follows from the Neyman-Pearson lemma (2) and *law of small numbers* that uniformly for all $\alpha \in [0, 1]$ (see Corollary 9)

$$f_{n1} \otimes \cdots \otimes f_{nn}(\alpha) = T\left(\mathrm{Ber}\left(\frac{\lambda_1}{n}\right)^{\otimes n}, \mathrm{Ber}\left(\frac{\lambda_2}{n}\right)^{\otimes n}\right)(\alpha) \tag{13}$$

$$= T\left(\mathrm{Bin}\left(n, \frac{\lambda_1}{n}\right), \mathrm{Bin}\left(n, \frac{\lambda_2}{n}\right)\right)(\alpha) \to T(P(\lambda_1), P(\lambda_2))(\alpha), \tag{14}$$

where $P(\lambda_1), P(\lambda_2)$ denotes Poisson distributions with mean λ_1 and λ_2 for some $\lambda_1, \lambda_2 > 0$. This example is clearly beyond the reach of the framework of Gaussian differential privacy, and depicts that based on the situation at hand, the baseline trade-off functions could be non-Gaussian. We unify Gaussian and Poisson limits under the framework of infinite divisibility Janssen et al. (1985).

On the $s^2 = 2k$ conjecture of Dong et al. (2022). It was remarked in Dong et al. (2022) that 'in all examples of the application of Theorem 2' it is observed that $s^2 = 2k$. We resolve this conjecture positively in our infinitely divisible setting as a consequence of a contiguity requirement.

2.4 Infinitely divisible distributions and Lévy-Khintchine formula

Now, we define infinitely divisible distributions, give some examples, state basic properties including closure under scaling and convolution, and weak limits, describe their appearances as distributional limits of sums of triangular arrays of (row-wise) IID random variables Kallenberg (2021).

Definition 5 (Infinitely divisible distribution (IDD)). *A random variable X on $\mathbb{R}$ is called infinitely divisible if $\forall n \in \mathbb{Z}_{\geq 1}$, there exists IID random variables $X_{n1}, \cdots, X_{nn}$ such that $X \stackrel{d}{=} \sum_{i=1}^n X_{ni}$.*

Equivalent definitions. A probability measure $\mu \in \mathcal{P}(\mathbb{R}, \mathcal{B}_\mathbb{R})$ or its (equivalent) Fourier transform $\hat{\mu}(t) = \mu(e^{itx})$ for $t \in \mathbb{R}$ is called infinitely divisible if for every $n \in \mathbb{Z}_{\geq 1}$, there exists a probability

measure (which is also unique) $\mu_n \in \mathcal{P}(\mathbb{R}, \mathcal{B}_\mathbb{R})$ such that $\mu = \mu_n^{*n}$ or equivalently $\hat{\mu}(t) = \hat{\mu}_n(t)^n$ for $t \in \mathbb{R}$, where $\int_\mathbb{R} f(x)d\mu * \nu(x) := \int_{\mathbb{R}^2} f(x+y)d\mu(x)d\nu(y)$ defines the convolution of measures $\mu * \nu$ by varying f. We denote with $\mathcal{I}$ the collection of all infinitely divisible distributions on $\mathbb{R}$.

Intuition, Gaussian and Compound Poisson examples in $\mathcal{I}$. The notion of infinitely divisible distribution captures the intuition that the random variable X can be broken into smaller IID pieces of arbitrary order. Elements of $\mathcal{I}$ include Normal distribution $\mu_1 \overset{d}{=} m + \sigma Z \leftrightarrow \hat{\mu}_1(t) = e^{itm - \frac{t^2\sigma^2}{2}}, t \in \mathbb{R}$ for some $m, \sigma \in \mathbb{R}$, Compound Poisson distribution $\mu_2 \overset{d}{=} \sum_{i=1}^N X_i$, with IID $X_i \sim \frac{\nu}{\|\nu\|}$ for $i \in \mathbb{Z}_{\geq 1}$ and independent $N \sim P(\|\nu\|) \perp \{X_i\}_{i \geq 1}$, where ν is a (uniquely determined from μ_2) positive finite Borel measure on $\mathbb{R} \setminus \{0\}$ of total mass $\|\nu\| = \nu(\mathbb{R}) \leftrightarrow \log \hat{\mu}_2(t) = \int_\mathbb{R} (e^{itx} - 1)d\nu(x)$. One can also represent the measure $\mu_2(\nu) = \sum_{k \in \mathbb{Z}_{\geq 0}} e^{-\nu(\mathbb{R})} \frac{\nu^{*k}}{k!}$ as a Poisson mixture of measures.

Basic closure properties of $\mathcal{I}$. $\mathcal{I}$ is closed under some algebraic properties including *affine change:* $\mathcal{I} \ni X$ implies $c_1 X + c_2 \in \mathcal{I}$ for $c_1, c_2 \in \mathbb{R}$, *convolution:* $X_1 \perp X_2 \in \mathcal{I}$ implies $X = X_1 + X_2 \in \mathcal{I}$. As a consequence we have an independent sum of Gaussian and Compound Poisson $X = m + \sigma Z + \sum_{i=1}^N X_i$ as described above, is also infinitely divisible. $\mathcal{I}$ is also closed under analytic properties including *weak convergence:* $\mathcal{I} \ni \mu_n \overset{d}{\to} \mu$ implies $\mu \in \mathcal{I}$. Consequently, Poisson, Geometric, Negative Binomial, as well as Chi-square, Exponential, Gamma, and Cauchy are all elements of $\mathcal{I}$. Moreover, a probability measure μ on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ is infinitely divisible if and only if there is a sequence of positive finite Borel measures $\{\nu_n\}_{n \in \mathbb{Z}_{\geq 1}}$ on $\mathbb{R} \setminus \{0\}$ so that $\mu_2(\nu_n) \overset{d}{\to} \mu$ Kallenberg (2021).

Generalization of CLT and parametrizing limits as elements of IDD. A remarkable generalization of the central limit theorem says if $\mu \in \mathcal{P}(\mathbb{R}, \mathcal{B}_\mathbb{R})$ is such that there exists a triangular array of (row-wise) IID random variables $\{X_{ni} : 1 \leq i \leq n\}_{n \geq 1}$ such that $\sum_{i=1}^n X_{ni} \overset{d}{\to} \mu$, then $\mu \in \mathcal{I}$ Durrett (2019), Bose et al. (2002). In fact, something more general holds, where one can drop the identically distributed condition [12] on the triangular array, and replace it with a weaker condition on the triangular array also known as the *uniformly asymptotically negligible* condition given by

$$\lim_{n \to \infty} \max_{1 \leq i \leq n} \mathbb{P}[|X_{ni}| \geq \varepsilon] = 0 \text{ for all } \varepsilon > 0 \tag{15}$$

Even under such weaker row-wise independent triangular array, $\{X_{ni} : 1 \leq i \leq n\}_{n \in \mathbb{Z}_{\geq 1}}$, if $\sum_{i=1}^n X_{ni} \overset{d}{\to} \mu$, then $\mu \in \mathcal{I}$ Klenke (2008). This enumeration of all possible limit distributions as elements of $\mathcal{I}$ is also the key technical ingredient of our enumeration of all possible baseline trade-off functions. We will establish our results again in two such stages like the above.

Lévy-Khintchine representation of IDD. Every IDD μ on $\mathbb{R}$ has a unique Lévy-Khintchine parametrization (m, σ, ν) of their Fourier transform with Gaussian and Compound Poisson parts, with ν a positive Borel measure on $\mathbb{R} \setminus \{0\}$ satisfying $\int_\mathbb{R} (x^2 \wedge 1)d\nu(x) < \infty$ Kallenberg (2021)

$$\log \hat{\mu}(t) = itm - \frac{1}{2}t^2\sigma^2 + \int_\mathbb{R} (e^{itx} - 1 - itx\mathbf{1}(|x| \leq 1))d\nu(x). \tag{16}$$

Kolmogorov series representation of IDD. Consider $m \in \mathbb{R}, \sigma \in \mathbb{R}_{\geq 0}$. Then one can define

$$X = m + \sigma Z + X_0 + \sum_{k \in \mathbb{Z}_{\geq 1}} (X_k - \mathbb{E}[X_k]) \tag{17}$$

a random series with independent random variables $Z \overset{d}{=} N(0,1)$, $X_k \overset{d}{=} \mu_2(\nu_k)$ on a common probability space $(\Omega, \mathcal{F})$ where the positive finite Borel measures ν_k is supported on $I_k := \left(-\frac{1}{k}, -\frac{1}{k+1} \right] \cup \left[\frac{1}{k+1}, \frac{1}{k} \right)$ [13]. As a consequence of Kolmogorov three series theorem, it is necessary and sufficient that $\sum_{k \in \mathbb{Z}_{\geq 1}} \mathbb{E}\left[(X_k - \mathbb{E}[X_k])^2 \right] < \infty$, for the infinite series representation to converge almost surely. Moreover, $\mathbb{E}[X_k] = \int x\nu_k(dx)$, $\mathbb{E}\left[(X_k - \mathbb{E}[X_k])^2 \right] = \int x^2\nu_k(dx)$, and $\sum_{k \in \mathbb{Z}_{\geq 1}} \mathbb{E}\left[(X_k - \mathbb{E}[X_k])^2 \right] < \infty \leftrightarrow \int x^2 d\sum_{k \in \mathbb{Z}_{\geq 1}} \nu_k(x) < \infty \leftrightarrow \int (x^2 \wedge 1)d\nu(x) < \infty$ where $\nu = \sum_{k \in \mathbb{Z}_{\geq 0}} \nu_k$ is the canonical measure of the infinitely divisible distribution μ Klenke (2008).

---

[12] We still keep the independence of elements of the row $\{X_{ni} : 1 \leq i \leq n\}$ for every row.

[13] We follow the convention that $\frac{1}{0} = \infty$ here to also include the case $k = 0$ for ν_0.

2.5 Resolution of the $s^2 = 2k$ conjecture

One of the major conclusions of the Dong et al. (2022) paper was that under the limit of large number of compositions of private mechanisms, a natural baseline trade-off function is the Gaussian trade-off function $f = T\left(N\left(-\frac{\mu^2}{2}, \mu^2\right), N\left(\frac{\mu^2}{2}, \mu^2\right)\right)$ for some $\mu \in \mathbb{R}$. But, as we saw earlier in 2.3 that this framework does not capture the Poisson example $f = T(P(\lambda_1), P(\lambda_2))$. We show that under the limit of large number of compositions of private mechanisms, an enumeration of all the baseline trade-off functions (encompassing both Gaussian and Poisson case) is given by the following.

$$\mathcal{I}_T = \{f = T(P, Q) : P \in \mathcal{I} \text{ and } dQ(x) = e^x dP(x)\} \tag{18}$$

Theorem 3. *Consider the infinitely divisible distribution $P \overset{d}{=} -k + |s|Z$ with only Gaussian component $Z \sim N(0, 1), k, s \in \mathbb{R}$. If $dQ(x) = e^x dP(x)$ is a probability measure, then $s^2 = 2k$. More precisely, for a Gaussian $P \in \mathcal{I}$ so that its Esscher tilt $dQ(x) = e^x dP(x)$ is a probability measure and $T(P, Q) \in \mathcal{I}_f$ if and only if $P = N\left(-\frac{\mu^2}{2}, \mu^2\right)$.*

Proof. The proof follows immediately from a computation of the moment-generating function of a Gaussian distribution, saying that $1 = Q(\mathbb{R}) = \mathbb{E}_P[\exp(x)] = \exp\left(-k + \frac{s^2}{2}\right) \to s^2 = 2k$. $\qquad \square$

Now, in the following lemma, we describe the basic properties of trade-off functions in general and show that the class of infinitely divisible trade-off functions $\mathcal{I}_\mathcal{T}$ that captures both the Gaussian $T(N(0, 1), N(\mu, 1))$ and the Poisson $T(P(\lambda_1), P(\lambda_2))$ trade-off functions.

Lemma 4. *Closure of trade-off functions under bijections and likelihood-ratios.*

1. *For two random variables defined X, Y on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and constants $0 \neq c_1, c_2 \in \mathbb{R}$ we have $T(X, Y) = T(c_1 X + c_2, c_1 Y + c_2)$, where we denote for any random elements $(\Omega, \mathcal{F}, \mathbb{P}) \overset{A,B}{\to} (\mathcal{X}, \mathcal{F}_\mathcal{X})$ trade-off function $T(A, B) := T(P_A, P_B)$ with $P_A = \mathbb{P} \circ A^{-1}$, $P_B = \mathbb{P} \circ B^{-1}$ are the corresponding distributions. More generally, for any bijective measurable map $(\mathcal{X}, \mathcal{F}_\mathcal{X}) \overset{h}{\to} (\mathcal{Y}, \mathcal{F}_\mathcal{Y})$ we have $T(A, B) = T(h(A), h(B))$.*

2. *For any two probability distributions P, Q on some measurable space $(\mathcal{X}, \mathcal{F}_\mathcal{X})$ we have $T(P, Q) = T(L_P, L_Q)$, where $L_P := P \circ \left(\log \frac{dQ}{dP}\right)^{-1}, L_Q := Q \circ \left(\log \frac{dQ}{dP}\right)^{-1}$ are the distributions of the likelihood ratio $\log \frac{dQ}{dP}$ under their respective probability measures.*

3. **Gaussian** *We observe that if $P \overset{d}{=} N\left(-\frac{\mu^2}{2}, \mu^2\right)$, then $dQ(x) = e^x dP(x) \overset{d}{=} N\left(\frac{\mu^2}{2}, \mu^2\right)$. So, the Gaussian trade-off curve as observed in Dong et al. (2022) belongs to $\mathcal{I}_T$.*

4. **Poisson.** *The Poisson curve $T(P(\lambda_1), P(\lambda_2)) \in \mathcal{I}_T$, since $T(P, Q) = T(P(\lambda_1), P(\lambda_2))$, where $P \overset{d}{=} \lambda_1 - \lambda_2 + N \log\left(\frac{\lambda_2}{\lambda_1}\right)$, with $N \sim P(\lambda_1)$ and $dQ(x) = e^x dP(x)$.*

Proof. 1. Observe that from Blackwell's theorem Dong et al. (2022) post-processing always increases the trade-off function. Therefore, for any measurable functions $(\mathcal{X}, \mathcal{F}_\mathcal{X}) \overset{h}{\to} (\mathcal{Y}, \mathcal{F}_\mathcal{Y})$, we have $T(h(A), h(B)) \geq T(A, B)$. Moreover, if h is bijective, reversing the process above yields $T(h(A), h(B)) \geq T(h^{-1}(h(A)), h^{-1}(h(B))) = T(A, B)$. Now, for scalar (or even vector) valued random variables, scaling ($x \to c_1 x, c_1 \neq 0$) and translations ($x \to x + c_2$) are measurable bijections, and hence the equality.

2. Let $X \sim P, Y \sim Q$ for random variables X, Y on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $L = \log \frac{dQ}{dP}$ the likelihood ratio, considered as a random variable $(\mathcal{X}, \mathcal{F}_\mathcal{X}) \to \mathbb{R}$. Again, by post–processing, $T(L(X), L(Y)) \geq T(X, Y) = T(P, Q)$. Now, for the reverse inequality, we apply the Neyman–Pearson likelihood ratio test (2) for the pair (P, Q) of size α.

$$\varphi^*(w) := \mathbf{1}\left(\log\left(\frac{dQ}{dP}\right) > \log \tau^*\right) + \lambda^* \mathbf{1}\left(\log\left(\frac{dQ}{dP}\right) = \log \tau^*\right), L = \log\left(\frac{dQ}{dP}\right)$$

8

with $\frac{dQ}{dP} := \frac{dQ}{d\mu} / \frac{dP}{d\mu} : (\mathcal{X}, \mathcal{F}_\mathcal{X}) \to [0, \infty]$, where the thresholds $\tau^* \in [0, \infty]$, $\lambda^* \in [0, 1]$ are chosen as in (2) so that $\mathbb{E}_P[\varphi^*] = \alpha$. Type I and type II errors are given by ($t^* = \log \tau^*$)

$$\alpha = \mathbb{E}_P[\varphi^*] = \mathbb{E}_P\big[\mathbf{1}\{L > t^*\} + \lambda^* \, \mathbf{1}\{L = t^*\}\big] = P[L > t^*] + \lambda^* P[L = t^*], \quad (19)$$

$$T(X, Y)(\alpha) = T(P, Q)(\alpha) = 1 - \mathbb{E}_Q[\varphi^*] = 1 - Q[L > t^*] - \lambda^* Q[L = t^*]. \quad (20)$$

Now, consider the trade-off function between $P_X \overset{d}{=} L(X)$ and $P_Y \overset{d}{=} L(Y)$ with the test $\varphi^*(x) = \mathbf{1}\{x > t^*\} + \lambda^* \, \mathbf{1}\{x = t^*\}$ to have the type I error $\mathbb{E}_{P_X}[\varphi^*] = \alpha$ and type II error $T(L(X), L(Y))(\alpha) \leq 1 - \mathbb{E}_{P_Y}[\varphi^*] = T(X, Y)(\alpha)$. Since $\alpha \in [0, 1]$ was arbitrary, we have $T(X, Y) = T(L(X), L(Y))$. Combining both inequalities proves the claim.

3. Observe that a shifted and scaled Gaussian $P = N\left(-\frac{\mu^2}{2}, \mu^2\right) \overset{d}{=} -\frac{\mu^2}{2} + |\mu|Z$ is indeed infinitely divisible and its Esscher tilt $\mathbb{Q}[\mathbb{R}] = \mathbb{E}_P[\exp(x)] = \exp\left(-\frac{\mu^2}{2} + \frac{\mu^2}{2}\right) = 1$, and therefore Q is a probability measure. Moreover, the moment generating function (MGF)

$$\mathbb{E}_Q[\exp(tx)] = \mathbb{E}_P[\exp((t+1)x)] = \exp\left((t+1)\left(-\frac{\mu^2}{2} + \frac{1}{2}(t+1)\mu^2\right)\right)$$

$$= \exp\left(\frac{t\mu^2}{2} + \frac{t^2\mu^2}{2}\right) = \mathbb{E}\big[\exp\left(t\left(\frac{\mu^2}{2} + |\mu|Z\right)\right)\big] \text{ for } t \in \mathbb{R}.$$

Since the moment generating function of Q matches that of a Gaussian, we have $Q \overset{d}{=} \frac{\mu^2}{2} + |\mu|Z$. Now the equality $T(P, Q) = T(Z, Z + \mu) = T(N(0, 1), N(\mu, 1))$ for $\mu > 0$ follows from the first part of the lemma about the invariance of trade-off curves under (common) translations and scaling, and additional translation symmetries of the Gaussian tradeoff functions $T(Z + c_1, Z + c_2) = T(Z, Z + c_2 - c_1)$ for $c_1, c_2 \in \mathbb{R}$.

$$T(P, Q) = T\left(-\frac{\mu^2}{2} + |\mu|Z, \frac{\mu^2}{2} + |\mu|Z\right) = T\left(-\frac{\mu}{2} + Z, \frac{\mu}{2} + Z\right) = T(Z, Z + \mu). \tag{21}$$

4. Consider a shifted and scaled Poisson $P \overset{d}{=} \lambda_1 - \lambda_2 + N_1 \log\left(\frac{\lambda_2}{\lambda_1}\right)$, with $N_1 \sim P(\lambda_1)$, and therefore infinitely divisible and $Q[\mathbb{R}] = \mathbb{E}_P[\exp(x)] = \exp\left(\lambda_1 - \lambda_2 + \lambda_1\left(\frac{\lambda_2}{\lambda_1} - 1\right)\right) = 1$, therefore Q is a probability measure. Moreover, the moment generating function (MGF)

$$\mathbb{E}_Q[\exp(tx)] = \mathbb{E}_P[\exp((t+1)x))] = \exp\left((t+1)(\lambda_1 - \lambda_2) + \lambda_1\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{t+1} - 1\right)\right)$$

$$= \exp\left(t(\lambda_1 - \lambda_2) + \lambda_2\left(\left(\frac{\lambda_2}{\lambda_1}\right)^t - 1\right)\right) = \mathbb{E}\big[\exp\left(t\left(\lambda_1 - \lambda_2 + N_2 \log\left(\frac{\lambda_2}{\lambda_1}\right)\right)\right)\big]$$

where $N_2 \sim P(\lambda_2)$. Since the moment generating function of Q matches that of a scaled and shifted Poisson, we have $Q \overset{d}{=} \lambda_1 - \lambda_2 + N_2 \log\left(\frac{\lambda_2}{\lambda_1}\right)$. Now the equality $T(P, Q) = T(P(\lambda_1), P(\lambda_2))$ for $\lambda_1, \lambda_2 > 0$ follows from the first part of the lemma about the invariance of trade-off curves under (common) translations and scaling[14].

$$T(P, Q) = T\left(\lambda_1 - \lambda_2 + N_1 \log\left(\frac{\lambda_2}{\lambda_1}\right), \lambda_1 - \lambda_2 + N_2 \log\left(\frac{\lambda_2}{\lambda_1}\right)\right) = T(N_1, N_2)$$

$\square$

Now, in the following lemma, we describe the basic properties of infinitely divisible trade-off functions, from which a closure of $\mathcal{I}_\mathcal{T}$ under pointwise limits $\mathcal{I}_\mathcal{T} \supset f_n \to f \in \mathcal{I}_T$ can also be inferred.

---

[14]Observe that equality holds if $\lambda_1 = \lambda_2 > 0$, and we do not need to apply any invariance properties.

Lemma 5. *Basic properties of the infinitely divisible trade-off functions.*

1. ***Tensor product.*** *Consider $P_1, P_2 \in \mathcal{I}$ with their corresponding Esscher tilts Q_1, Q_2. Then the tensor product $f = T(P_1 \otimes P_2, Q_1 \otimes Q_2) = T(P,Q) \in \mathcal{I}_T$ with $P = P_1 * P_2$ and $dQ = e^x dP$. This immediately extends to $(P_i, Q_i)_{i=1}^n$ for $P_i \in \mathcal{I}$ for any $n \in \mathbb{Z}_{\geq 1}$ as well.*

2. ***Infinite divisibility.*** *For any trade off function $f = T(P,Q) \in \mathcal{I}_T$ and any $n \in \mathbb{Z}_{\geq 1}$ there exists a trade off function $f_n = T(P_n, Q_n) \in \mathcal{I}_T$ such that $f = f_n^{\otimes n}$.*

Proof.   1. Observe that the likelihood ratio of the tensor product pairs of distributions satisfies

$$L_{P_1 \otimes P_2} := (P_1 \otimes P_2) \circ \left(\log \frac{dQ_1 \otimes Q_2}{dP_1 \otimes P_2} \right)^{-1} = P_1 * P_2 =: P, \text{ and} \qquad (22)$$

$$dL_{Q_1 \otimes Q_2} = d(Q_1 \otimes Q_2) \circ \log \left(\frac{dQ_1 \otimes Q_2}{dP_1 \otimes dP_2} \right)^{-1} = e^x dP_1 * P_2 = e^x dP =: dQ. \quad (23)$$

Now, using the lemma above, we have $T(P_1 \otimes P_2, Q_1 \otimes Q_2) = T(L_{P_1 \otimes P_2}, L_{Q_1 \otimes Q_2}) = T(P,Q)$. Finally, observe that $P_1 * P_2$ is infinitely divisible if P_1 and P_2 are Kallenberg (2021) since the addition of two independent infinitely divisible random variables is also an infinitely divisible random variable. Moreover, $\mathbb{E}_{P_1 * P_2}[\exp(x)] = \mathbb{E}_{P_1}[\exp(x)]\mathbb{E}_{P_2}[\exp(x)] = 1$ satisfies the normalization requirement.

2. Observe that by the definition of infinite divisibility of $P \stackrel{d}{=} X$ satisfying $\mathbb{E}_P[\exp(x)] = \mathbb{E}[\exp(X)] = 1$, for every $n \in \mathbb{Z}_{\geq 1}$ we have $P = P_n^{*n}$ for some probability measure P_n on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ such that $1 = \mathbb{E}_P[\exp(x)] = (\mathbb{E}_{\mathbb{P}_n}[\exp(x)])^n$. Then, from the positivity of the quantity $\mathbb{E}_{\mathbb{P}_n}[\exp(x)] > 0$, we have $\mathbb{E}_{\mathbb{P}_n}[\exp(x)] = 1$, and so $dQ_n = e^x dP_n$ is a probability measure. Moreover, from the Levy-Khintchine characterization (m, σ, ν) of the characteristic function $\hat{P}(t)$ of an infinitely divisible distribution P, it follows that P_n itself is an infinitely divisible distribution with its Levy-Khintchine triplet $\left(\frac{m}{n}, \frac{\sigma}{\sqrt{n}}, \frac{\nu}{n} \right)$. Now, $f_n^{\otimes n} = T(P_n^{\otimes n}, Q_n^{\otimes n}) = T(P,Q)$, where $P = P_n^{*n}$, and $dQ = e^x dP$.

$\square$

3   Technical results

In this section, we prove the main technical result of the paper, establishing why is the collection of infinitely divisible trade-off functions $\mathcal{I}_T$ (18) appears under the limit of a large number of compositions of (nearly perfect) differentially private mechanisms, as a converse of the result 5.

Theorem 6. *Consider a sequence of trade off functions $\{f_n\}_{n \in \mathbb{Z}_{\geq 1}}$ such that $f_n(0) = 1$ for all $n \in \mathbb{Z}_{\geq 1}$ and $f_n^{\otimes n}(\alpha) \to f_\infty(\alpha)$ pointwise on $[0,1]$ for some trade-off function f_∞ (1), then f_∞ is an infinitely divisible trade off function. More precisely $f_\infty \in \mathcal{I}_T$.*

Proof. We prove it in a few steps. First, recall that any trade-off function $f \in \mathcal{T}$ can be realized as $f = T(P,Q)$ on the measurable space $(\mathcal{W}, \mathcal{F}_{\mathcal{W}}) = ([0,1], \mathcal{B}_{[0,1]})$ by taking P to be the standard uniform probability measure U on $([0,1], \mathcal{B}_{[0,1]})$ and $Q = Q_f$ to be a Borel probability measure on $[0,1]$ with a cumulative distribution function $Q([0,x]) = f(1-x)$ for $x \in [0,1)$ and having an atom at $x = 1$ of mass $Q(\{1\}) = 1 - f(0)$ whenever $f(0) < 1$[15]. So, for tradeoff functions f_n with $f_n(0) = 1$ we have $f_n = T(P, Q_n)$, where $P = U$ as above and $Q_n := Q_{f_n} \ll P$[16] since

---

$f_n(0) = 1$[17]. More precisely, $f_n(0) = 1$ is equivalent to the fact $Q_n \ll P$[18]. Therefore, $f_n(0) = 1$ implies Q_n has no atom or, more precisely, no singular component with respect to $P = U$.

Second, $f_n^{\otimes n} = T(U^{\otimes n}, Q_n^{\otimes n})$ converges to $f_\infty = T(P_\infty, Q_\infty)$ pointwise (hence uniformly Dong et al. (2022),Lehmann et al. (1986))[19] for some pairs of probability measures (P_∞, Q_∞) on a measurable space $(\Omega, \mathcal{F})$[20]. Now, $f_n(0) = 1$ for all $n \in \mathbb{Z}_{\geq 1} \rightarrow f_\infty(0) = 1$ which is equivalent to $Q_\infty \ll P_\infty$. This is an equivalent reformulation of contiguity for $(U^{\otimes n}, Q_n^{\otimes n})$ Le Cam (1986).

Third, the pointwise convergence of trade-off functions for the sequence $(U^{\otimes n}, Q_n^{\otimes n})$ to the trade-off function for the pair (P_∞, Q_∞) is equivalent to weak convergence of the sequence on $\mathbb{R}$ [21]

$$U^{\otimes n} \circ \left(\log \frac{dQ_n^{\otimes n}}{dU^{\otimes n}} \right)^{-1} \xrightarrow{d} P_\infty \circ \left(\log \frac{dQ_\infty}{dP_\infty} \right)^{-1} \overset{d}{=} L(P_\infty) \tag{24}$$

Now, we observe that on the canonical probability space $\left([0,1]^n, \mathcal{B}_{[0,1]^n}, U^{\otimes n}\right)$ the random variable $\log\left(\frac{dQ_n^{\otimes n}}{dU^{\otimes n}}(x_1, \cdots, x_n) \right) = \sum_{i=1}^n \log\left(\frac{dQ_n}{dU}(x_i) \right)$ is a sum of independent and identically distributed random variables. Moreover, by assumption of the theorem, this IID sum converges in distribution to a probability distribution $L(P_\infty)$. Now, by the fundamental theorem of infinitely divisible distributions[22] $L(P_\infty)$ is infinitely divisible, and moreover $T(P_\infty, Q_\infty) = T(L(P_\infty), L(Q_\infty))$ where $dL(Q_\infty)(x) = e^x dL(P_\infty)(x)$. Hence, $f_\infty = T(P_\infty, Q_\infty) \in \mathcal{I}_T$ is infinitely divisible. $\square$

A Gaussian example. Consider the case $f_n(\cdot) = T(N(0,1), N(\mu_n, 1))(\cdot) = \Phi(\Phi^{-1}(1 - \cdot) - \mu_n)$ for $\mu_n > 0$. Then $f_n(0) = 1$, and by the dimension-freeness of Gaussian trade-off functions[23], we have $f_n^{\otimes n}(\cdot) = T(N(0,1), N(\mu_n \sqrt{n}, 1)) = \Phi(\Phi^{-1}(1 - \cdot) - \mu_n \sqrt{n})$. Therefore, the pointwise convergence of $f_n^{\otimes n}$ to a limiting trade-off function $f \in \mathcal{T}$ immediately implies that the sequence $\{\mu_n \sqrt{n} : n \in \mathbb{Z}_{\geq 1}\}$ has to converge, and moreover, the $\lim_{n\to\infty} \mu_n \sqrt{n} = \mu_\infty$ has to be finite, because the trade-off function f_∞ has to have the form $f_\infty(\cdot) = \Phi(\Phi^{-1}(1 - \cdot) - \mu_\infty)$, and if $\mu_\infty = \infty$, then $f_\infty \equiv 0$[23], a contradiction. We compute non-Gaussian examples in section 3.1

The Contiguity assumption in Theorem 6. The assumption of $f_n(0) = 1$ for all $n \in \mathbb{Z}_{\geq 1}$ in Theorem 6 is maintained so that the canonical pair of probability measures $P = U$ and $Q_n := Q_{f_n}$ satisfy absolute continuity $Q_n \ll U$ for all $n \in \mathbb{Z}_{\geq 1}$. Consequently, our results are not directly applicable to the convergence of trade-off functions involving $f_{\varepsilon_n, \delta_n}$ with $\delta_n > 0$ for all $n \in \mathbb{Z}_{\geq 1}$. However, the tensor product decomposition $f_{\varepsilon, \delta} = f_{\varepsilon, 0} \otimes f_{0, \delta}$ as observed in Dong et al. (2022) immediately allows us to apply our results directly to the $f_n = f_{\varepsilon_n, 0}$ part and treat the

---

[17]This follows from the fact that by definition of $f_n \in \mathcal{T}$ (1) we have Q_n is absolutely continuous with respect to U on $[0,1)$ with Radon-Nikodym density $\frac{dQ_n}{dP}(x) = -f'_n(1-x)$, P almost everywhere on $[0,1)$, (f_n is convex and therefore Lebesgue almost differentiable Rockafellar (1970)). Now $f_n(0) = 1$ implies that this absolute continuity $Q_n \ll U$ extends to the entire closed interval $[0,1]$.

[18]$Q_n \ll P \rightarrow$ 'for measurable $(\mathcal{X}, \mathcal{F}_{\mathcal{X}}) \xrightarrow{\varphi} [0,1]$ we have $\mathbb{E}_P(\varphi) = 0 \rightarrow \mathbb{E}_{Q_n}(\varphi) = 0 \rightarrow f_n(0) = 1$.

[19]The intuition is that if a collection of monotone functions $f_n : [0,1] \to [0,1]$ converges pointwise to a continuous monotone function $f : [0,1] \to [0,1]$, then the convergence is uniform. Moreover, the domain (in this case $[0,1]$) need not be compact, only the range (in this case $[0,1]$) needs to be. The version, where trade-off functions $\{(f_n)_{n \in \mathbb{Z}_{\geq 1}}, f\}$ are replaced by cumulative distributive function $\{(F_n)_{n \in \mathbb{Z}_{\geq 1}}, F\}$ with F continuous, this is known as the Glivenko-Cantelli lemma or the fundamental theorem of statistics, whose finite sample concentration is known as the Dvoretzky–Kiefer–Wolfowitz inequality Kallenberg (2021).

[20]Observe that we are not taking the canonical choice for f, because we believe the infinite divisible parametrization $\mathcal{I}_T$ is more useful for expressing limiting trade-off functions and practically more informative in the sense that will be once we demonstrate the procedures on how to achieve such privacy requirements.

[21]This equivalence of weak convergence of 'Blackwell experiments' for the two point set $\Theta = \{0,1\}$ and different equivalences has been established in detail in Le Cam (1986), Torgersen (1991). The version with weak convergence of the likelihood ratio (under the null) to the likelihood ratio under the limiting null is ubiquitous in statistics, and the pointwise convergence of the trade-off functions to a limiting trade-off function can be exactly written as 'Le Cam distance' between $(U^{\otimes n}, Q_n^{\otimes n})$ and (P, Q) going to zero.

[22]The fundamental theorem of infinitely divisible distributions Durrett (2019) says that if a probability measure μ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ is such that there exists a sequence of probability measures $\{\mu_n\}_{n \in \mathbb{Z}_{\geq 1}}$ on $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ satisfying weak convergence under convolutions $\mu_n^{*n} \xrightarrow{d} \mu$, then μ is infinitely divisible.

[23]Observe that for any $\alpha > 0$ we have $f_\infty(\alpha) = \Phi(\Phi^{-1}(1 - \alpha) - \mu_\infty) = 0$ whenever $\mu_\infty = \infty$, and therefore the continuity requirement on f throughout the closed interval on $[0,1]$ implies that $f_\infty(0) = 0$.

$f_n = f_{0,\delta_n}$ part separately, as has been done in Dong et al. (2022) using the identity $f_{0,\delta_1} \otimes f_{0,\delta_2} = f_{0,1-(1-\delta_1)(1-\delta_2)}$. Moreover, as Dong et al. (2022) established the limits of such tensor products

$$f_{0,\frac{\delta}{n}}^{\otimes n} \to f_\infty = f_{0,1-e^{-\delta}} \text{ where } f_{0,1-e^{-\delta}}(0) = 1 - \delta \neq 1 \text{ whenever } \delta > 0. \tag{25}$$

So $f_\infty = f_{0,1-e^{-\delta}}$ is not infinitely divisible according to our definition of $\mathcal{I}_\mathcal{T}$. On a different note, we believe that our conclusions of Theorem 6 will continue to hold under the weaker assumption that $f_n(0)$ need not to be equal to 1, but we require $f_n^{\otimes n}(0) \to 1$ as $n \to \infty$.

The Convergence assumption in Theorem 6 . The assumption $g_n(\alpha) := f_n^{\otimes n}(\alpha) \to f_\infty(\alpha)$ was made pointwise on $[0,1]$ for some trade-off function f (1) in the statement of Theorem 6. This is natural since our primary goal was to enlarge the class of Gaussian trade-off functions obtained in Dong et al. (2022) and to see what class of trade-off functions can be obtained in general under the composition of a large number of nearly perfect differentially private operations, under the assumption that these degradation curves $g_n := f_n^{\otimes n}$ do converge to a limiting trade-off curve. However, in practice of differential privacy through the f-differential privacy framework, one has to equivalently show (under the assumptions of our Theorem 6) that if $f_n = T(P_n, Q_n)$ with $f_n(0) = 1$, then $f_n^{\otimes n} \to f_\infty$ for some limiting trade-off function, $f_\infty = T(P_\infty, Q_\infty)$ is the same as the following (possibly more tractable) distributional convergence of the likelihood ratio under the null

$$\text{under } P_n \quad \log \frac{dQ_n^{\otimes n}}{dP_n^{\otimes n}} \xrightarrow{d} X \overset{d}{=} \log \frac{dQ_\infty}{dP_\infty} \text{ under } P_\infty. \tag{26}$$

In section 3.1, we show the distributional convergence of $\log \frac{dQ_n^{\otimes n}}{dP_n^{\otimes n}}$ under P_n for various statistical models, including the locally asymptotically normal, Poisson and Compound Poisson families.

Blackwell-Le-Cam equivalence. Observe that our conditions of Theorem 6 only involve trade-off functions $\{f_n : n \in \mathbb{Z}_{\geq 1}\}, f$, and it is deliberate since the definition of f-differential privacy of a mechanism 2 does not involve any details of the binary experiment $\mathcal{E} = (\Omega, \mathcal{F}, P, Q)$, except that it works for any such equivalent representation of the baseline trade-off function f through $f = T(P, Q)$. This relates to the Blackwell-Le Cam equivalence of experiments, which we will briefly describe for binary experiments with a parameter set $\Theta = \{0, 1\}$ consisting of two elements.

Given two binary experiments $\mathcal{E}_1 = (\Omega_1, \mathcal{F}_1, P_1, Q_1), \mathcal{E}_2 = (\Omega_2, \mathcal{F}_2, P_2, Q_2)$, a notion of distance between $\mathcal{E}_1, \mathcal{E}_2$ (better known as *Le-Cam deficiency*) was defined in Le Cam (1986) that is given as

$$\textbf{Le-Cam distance: } \Delta(\mathcal{E}_1, \mathcal{E}_2) = \max(\delta(\mathcal{E}_1, \mathcal{E}_2), \delta(\mathcal{E}_2, \mathcal{E}_1)), \text{ where} \tag{27}$$

$$\textbf{Le-Cam deficiency: } \delta(\mathcal{E}_1, \mathcal{E}_2) = \inf_{R:(\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)} \max(\|RP_1 - P_2\|_1, \|RQ_1 - Q_2\|_1), \tag{28}$$

where the infimum is taken over all (appropriate) Markov kernels R and for a pair of probability measures μ_1, μ_2 on $(\Omega, \mathcal{F}_2)$, $\|\mu_1 - \mu_2\|_1 := \sup_{A \in \mathcal{F}_2} |\mu_1(A) - \mu_2(A)|$ is the total variation distance.

One of the key results of Le Cam (1986), Torgersen (1991) is that $\Delta(\mathcal{E}_1, \mathcal{E}_2) = 0$ if and only if $f_1 := T(P_1, Q_1) \equiv T(P_2, Q_2) =: f_2$. Moreover, $f_2 \geq f_1$ if and only if $\delta(\mathcal{E}_1, \mathcal{E}_2) = 0$ if and only if there exists a Markov kernel $R : (\Omega_1, \mathcal{F}_1) \to (\Omega_2, \mathcal{F}_2)$ such that $RP_1 = P_2$ and $RQ_1 = Q_2$. Furthermore, the correspondence $\mathcal{E} = (\Omega, \mathcal{F}, P, \mathcal{Q}) \leftrightarrow f = T(P, Q)$ preserves the Le-Cam distance in the following sense and therefore is quantitative, since Torgersen (1991)[Corollary 9.3.27] states

$$\sqrt{2}\Delta(\mathcal{E}_1, \mathcal{E}_2) = \Lambda(Q^{(1)}, Q^{(2)}), \text{ where the} \tag{29}$$

cumulative distribution functions $F^{(1)}, F^{(2)}$ of probability measures $Q^{(1)}, Q^{(2)}$ on $[0,1]$ are given

$$\text{by } Q^{(i)}[0, \alpha] = 1 - f_i(\alpha) \text{ for } \alpha \in (0, 1] \text{ and } Q^{(i)}[\{0\}] = (1 - f_i(0)), \text{ for } i \in \{1, 2\} \tag{30}$$

and $\Lambda(Q^{(1)}, Q^{(2)})$ denotes the Levy distance between the probability measures $Q^{(1)}, Q^{(2)}$ through their cumulative distribution functions $F^{(1)}, F^{(2)}$ on $[0,1]$ as Kallenberg (2021)

$$\Lambda(F^{(1)}, F^{(2)}) = \inf\{\varepsilon > 0 : F^{(2)}(x - \varepsilon) - \varepsilon \leq F^{(1)}(x) \leq F^{(2)}(x + \varepsilon) + \varepsilon \text{ for all } x \in [\varepsilon, 1 - \varepsilon]\}. \tag{31}$$

The key observations about the Levy-distance $\Lambda(\cdot, \cdot)$ (hence Δ for binary Θ) is that it precisely captures convergence in distribution. More precisely, our assumption in Theorem 6 about the pointwise

convergence of $f_n^{\otimes n} \to f_\infty$ on $[0,1]$ [24] is the equivalent to the condition that $\Delta(1-f_n^{\otimes n}, 1-f_\infty) \to 0$, and therefore captures the meaning of convergence of experiments $\Delta(\mathcal{E}_n^{\otimes n}, \mathcal{E}_\infty) \to 0$ where the correspondence $\mathcal{E}_n \equiv (P_n, Q_n) \to f_n = T(P_n, Q_n)$ and $\mathcal{E}_\infty \equiv (P_n, Q_n) \to f_\infty \equiv T(P_\infty, Q_\infty)$.

From the definition of $\Delta(\mathcal{E}_1, \mathcal{E}_2)$, given a fixed parameter set $\Theta = \{0,1\}$ (in our case) one can (consistently) define a equivalence class of experiments $[\mathcal{E}] := \{\mathcal{E}_1 : \Delta(\mathcal{E}, \mathcal{E}_1) = 0\}$ with the same trade-off curve $f \leftrightarrow [\mathcal{E}]$. Now, in the f-differential privacy framework of Dong et al. (2022), it only depends on the equivalence classes of experiments and not the exact representation, such as $f = T(P, Q)$. However, this explicit representation $f = T(P, Q)$ is important in practice since, one has to implement a Markov kernel (a generalized noise-adding mechanism) such as adding Gaussian noise to achieve Gaussian differential privacy $f = T(N(0,1)N(\mu,1))$.

This is the reason our explicit parametrization of infinitely divisible trade-off functions $f_\infty = T(P_\infty, Q_\infty)$ of $\mathcal{I}_{\mathcal{T}}$ is important, and we will demonstrate in section 4 how to achieve infinitely divisible privacy for a large class of infinitely divisible trade-off functions $f_\infty = T(P_\infty, Q_\infty)$, including the practically relevant Poisson case $f_\infty = T(P(\lambda_1), P(\lambda_2))$.

Large deviation regime versus infinitely divisible regime. Instead of a sequence $\{f_n : n \in \mathbb{Z}_{\geq 1}\}$ of trade-off functions and considering their tensor product $f_n^{\otimes n}$, if one takes a single trade-off function f (independent of n) and considers $f^{\otimes n}$, one has a large deviation principle

$$\lim_{n \to \infty} \frac{1}{n} \log f^{\otimes n}(\alpha) = \int_0^{z_f} \log|f'(x)|dx, \text{ for all } \alpha \in (0,1) \tag{32}$$

where $z_f = f^{-1}(0)$ is the first zero of f, where f^{-1} is the generalized inverse of f (4) (seeDong et al. (2022)[Prop B.4], Polyanskiy & Wu (2025), Torgersen (1991)). Although our focus is on the infinitely divisible (central limit) regime, where there is some non-trivial privacy left even after a large number of nearly perfect private operations. The large deviation regime is also of interest and has been used in the theoretical development of watermark detection for LLMs Li et al. (2025).

Some consequences in practice. One can observe that, as a consequence of the continuous mapping theorem [25] the mean of the distribution $L(P)$ can be expressed in terms of the sequence $\{f_n\}_{n \in \mathbb{Z}_{\geq 1}}$.

$$\mathbb{E}_{L(P)}[x] = \mathbb{E}_P\big[\log\big(\frac{dQ}{dP}\big)\big] = \lim_{n \to \infty} \sum_{i=1}^n \mathbb{E}_U\big[\log\big(\frac{dQ_n}{dU}\big)\big] = \lim_{n \to \infty} n \int_0^1 (\log(-f_n'(1-x)))\,dx. \tag{33}$$

Moreover, the cumulant generating function (CGF) of the distribution $L(P)$ for $t \in [0,1]$ [26] can be expressed through the trade-off function sequence as well $\{f_n\}_{n \in \mathbb{Z}_{\geq 1}}$ [27].

$$\log \mathbb{E}_{L(P)}[\exp(tx)] = \log \mathbb{E}_P\big[\big(\frac{dQ}{dP}\big)^t\big] = \lim_{n \to \infty} n \log \big(\int_0^1 (-f_n'(1-x))^t\,dx\big)\big] \tag{34}$$

As a consequence, one can differentiate both sides with respect to t and derive an expression for moments of $L(P)$ in terms of the sequence of $\{f_n\}_{n \in \mathbb{Z}_{\geq 1}}$ Nica & Speicher (2006).

Triangular systems of differentially private procedures. Now, we rephrase the implications of the above statement in the f-differential privacy framework, where we consider triangular arrays of trade-off functions $\{f_{ni} : 1 \leq i \leq n\}_{n \in \mathbb{Z}_{\geq 1}}$ with $f_{ni}(0) = 1$ for all $1 \leq i \leq n \in \mathbb{Z}_{\geq 1}$. We consider the tensor product $f_{n1} \otimes \cdots \otimes f_{nn} = T(U^{\otimes n}, Q_{n1} \otimes \cdots \otimes Q_{nn})$, where U is the uniform probability measure on $([0,1], \mathcal{B}_{[0,1]})$ and $Q_{ni}[0,x] = f_{ni}(1-x)$ satisfying $Q_{ni} \ll U$. Now, if we assume that

---

[24] The assumption $f_n(0) = 1$ (hence $f_\infty(0) = 1$) is there to ensure that $Q_\infty \ll P_\infty$ for $f_\infty = T(P_\infty, Q_\infty)$.

[25] To be mathematically precise, we need the *uniform integrability* condition on the sequence of random variables $\{(\frac{dQ_n}{dU})\}_{n \in \mathbb{Z}_{\geq 1}}$ under $U^{\otimes n}$ Kallenberg (2021) so that we can use $X_n \xrightarrow{d} X \implies \mathbb{E}[X_n] \to \mathbb{E}[X]$.

[26] As pointed out in Torgersen (1991) the CGF might not be well defined outside $t \in [0,1]$.

[27] $(P, Q) \to \mathbb{E}_P\big[\big(\frac{dQ}{dP}\big)^t\big]$ is also known as the Renyi-divergence or Hellinger transform of order t Polyanskiy & Wu (2025). In Le-cam's theory of weak convergence of experiments for the two-point set $\Theta = \{0,1\}$, convergence of the Hellinger transforms of all orders $t \in [0,1]$ for the pair $(U^{\otimes n}, Q_n^{\otimes n})$ to that of the limiting pair (P, Q) is another equivalent conditions along with the weak convergence of the likelihood ratios under the null, and the Le-cam deficiency metric $\Delta(\mathcal{E}_n, \mathcal{E})$ between $\mathcal{E}_n = (U^{\otimes n}, Q_n^{\otimes n})$ and $\mathcal{E} = (P, Q)$ going to zero.

$f_{n1} \otimes \cdots \otimes f_{nn}$ converges pointwise to $f = T(P, Q)$ for some trade-off function f[28], by Le Cam (1986); Torgersen (1991), it becomes equivalent to the weak convergence of the likelihood ratio

$$U^{\otimes n} \circ \left(\log \frac{dQ_{n1} \otimes \cdots \otimes dQ_{nn}}{dU^{\otimes n}} \right)^{-1} \xrightarrow{d} P \circ \left(\log \frac{dQ}{dP} \right)^{-1} \overset{d}{=} L(P). \quad (35)$$

Now, the likelihood ratio $\log \left(\frac{dQ_{n1} \otimes \cdots \otimes dQ_{nn}}{dU \otimes \cdots \otimes dU}(x_1, \cdots, x_n) \right) = \sum_{i=1}^{n} \log \left(\frac{dQ}{dU}(x_i) \right)$ on the probability space $\left([0,1]^n, \mathcal{B}_{[0,1]^n}, U^{\otimes n} \right)$ is a random variable that is a sum of independent random variables $\{ \log \left(\frac{dQ_{ni}}{dU} \right) \}_{i=1}^{n}$ under $U^{\otimes n}$. This holds true for every $n \in \mathbb{Z}_{\geq 1}$. Now, under the technical assumption of a uniformly asymptotically negligible condition,

$$\lim_{n \to \infty} \max_{1 \leq i \leq n} U \left(x \in [0,1] : |\log \left(\frac{dQ_{ni}}{dU}(x) \right)| > \varepsilon \right) = 0 \text{ for all } \varepsilon > 0, \quad (36)$$

the limiting distribution $L(P)$ is still an IDD with its Esscher tilt distribution $dL(Q) = e^x dL(P)(x)$. One can observe that, as a consequence of the continuous mapping theorem[29], the mean $\mathbb{E}_{L(P)}[x]$ of the distribution $L(P)$ can be expressed in terms of the sequence $\{ f_{ni} : 1 \leq i \leq n \}_{n \in \mathbb{Z}_{\geq 1}}$.

$$\mathbb{E}_P \left[\log \left(\frac{dQ}{dP} \right) \right] = \lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{E}_U \left[\log \left(\frac{dQ_{ni}}{dU} \right) \right] = \lim_{n \to \infty} \sum_{i=1}^{n} \int_0^1 \left(\log \left(-f'_{ni}(1-x) \right) \right) dx. \quad (37)$$

Moreover, the cumulant generating function (CGF) of the distribution $L(P)$ for $t \in [0,1]$ can be expressed through the trade-off function sequence as well $\{ f_{ni} : 1 \leq i \leq n \}_{n \in \mathbb{Z}_{\geq 1}}$[30].

$$\log \mathbb{E}_{L(P)}[\exp(tx)] = \log \mathbb{E}_P \left[\left(\frac{dQ}{dP} \right)^t \right] = \lim_{n \to \infty} \sum_{i=1}^{n} \log \left(\int_0^1 \left(-f'_{ni}(1-x) \right)^t dx \right) \quad (38)$$

As a consequence, one can differentiate both sides with respect to t and derive an expression for moments (cumulants) of $L(P)$ in terms of the sequence of $\{ f_n \}_{n \in \mathbb{Z}_{\geq 1}}$ Nica & Speicher (2006). Moreover, in the above condition (36) if one inserts the maximum inside the probability, and demands $\lim_{n \to \infty} U \left(x \in [0,1] | \max_{1 \leq i \leq n} |\log \left(\frac{dQ_{ni}}{dU}(x) \right)| > \varepsilon \right) = 0$ for all $\varepsilon > 0$, then this condition is known as a version of the Lindeberg-type condition and causes the limit distribution $L(P)$ to be Gaussian, in our case, of the form $L(P) \overset{d}{=} -\frac{s^2}{2} + |s| Z$ for $Z \sim N(0,1)$ Le Cam & Yang (2000).

Connections to high dimensional hypothesis testing. Observe that in the f-differential privacy framework of Dong et al. (2022), while sequentially applying more than one differentially private operation: first M_1 (satisfying f_1 DP), then M_2 (satisfying f_2 DP)[31] it has been captured through tensor products of experiments $\mathcal{E}_1 = (\Omega_1, \mathcal{F}_1, P_1, Q_1)$ with $f_1 = T(P_1, Q_1)$ and $\mathcal{E}_2 = (\Omega_2, \mathcal{F}_2, P_2, Q_2)$ with $f_2 = T(P_2, Q_2)$ so that $M = (M_1, M_2)$ satisfy $f = f_1 \otimes f_2 = T(P_1 \otimes P_2, Q_2 \otimes Q_2)$ DP.

However, an inspection of the proof of the composition theorem Dong et al. (2022)[Lemma C.3] reveals that the condition on M_2 satisfying f_2 DP independent of the exact outcome of M_1 on the dataset S gives rise to the tensor product structure[32]. In principle, it is not unreasonable to consider a

---

[28]Pointwise convergence of the sequence of trade-off functions $\{ f_n \}$ with $f_n(0) = 1$ requires $f(0) = 1$ and therefore $Q \ll P$, a contiguity statement in disguise for the sequence $\left(U^{\otimes n}, Q_{n1} \otimes \cdots \otimes Q_{nn} \right)$. In our case, $f_n = f_{n1} \otimes \cdots f_{nn} = T(U^{\otimes n}, Q_{n1} \otimes \cdots \otimes Q_{nn})$, and because $f_{ni}(0) = 1 \leftrightarrow Q_{ni} \ll U \to Q_{n1} \otimes \cdots \otimes Q_{nn} \ll U^{\otimes n}$ (absolute continuity is preserved under tensor products Folland (1999)) $\leftrightarrow f_n(0) = 1$.

[29]To be mathematically precise we need the *uniform integrability* conditions analogous to the IID case.

[30]$(P, Q) \to \mathbb{E}_P \left[\left(\frac{dQ}{dP} \right)^t \right]$ is also known as the Renyi-divergence or Hellinger transform of order t Polyanskiy & Wu (2025). In Le-cam's theory of weak convergence of experiments for the two-point set $\Theta = \{0, 1\}$, convergence of the Hellinger transforms of all orders $t \in [0,1]$ for the pair $(U^{\otimes n}, Q_n^{\otimes n})$ to that of the limiting pair (P, Q) is another equivalent conditions along with the weak convergence of the likelihood ratios under the null, and the Le-cam deficiency metric $\Delta(\mathcal{E}_n, \mathcal{E})$ between $\mathcal{E}_n = (U^{\otimes n}, Q_n^{\otimes n})$ and $\mathcal{E} = (P, Q)$ going to zero.

[31]This includes adaptive operations, meaning the action of M_2 not only depends on the original dataset S, but also on the output of $M_1(S)$. So, the joint output is $M(S) = (M_1(S), M_2(S, M_1(S)))$ on a dataset S.

[32]Since this essentially assumes the worst case privacy loss for the second mechanism M_2 as f_2 uniformly over all outputs of the first mechanism M_1, it does not allow f_2 to depend on the exact output $M_1(S)$.

more flexible situation so that the degradation of privacy upon applying multiple procedures $M^{(n)} = (M_1, M_2, \cdots, M_n)$ (on the same dataset S) might occur in the more general form where each $M^{(n)}$ satisfy $g^{(n)}$-DP where $g^{(n)} = T(P^{(n)}, Q^{(n)})$ with the experiment $\mathcal{E}^{(n)} = (\Omega^{(n)}, \mathcal{F}^{(n)}, P^{(n)}, Q^{(n)})$. Observe that, in the previous worst-case analysis, these experiments had the product structure

$$\mathcal{E}^{(n)} = \mathcal{E}_1 \times \cdots \times \mathcal{E}_n = (\Omega_1 \times \cdots \times \Omega_n, \mathcal{F}_1 \otimes \cdots \otimes \mathcal{F}_n, P_1 \otimes \cdots \otimes P_n, Q_1 \otimes \cdots \otimes Q_n). \quad (39)$$

This product structure is precisely what led us to the infinitely divisible structure of our limiting trade-off functions $\mathcal{I}_{\mathcal{T}}$. However, for a general such experiment $\mathcal{E}^{(n)} = (\Omega^{(n)}, \mathcal{F}^{(n)}, P^{(n)}, Q^{(n)})$ capturing privacy degradation for $M^{(n)} = (M_1, \cdots, M_n)$ more adaptively than the tensor product framework of Dong et al. (2022), it is difficult in practice to establish the pointwise convergence of the sequence of trade-off functions $g^{(n)} = T(P^{(n)}, Q^{(n)}) \to g_\infty = T(P_\infty, Q_\infty)$ and to be able to guess the structure of the class of such limiting trade-off functions $g_\infty = T(P_\infty, Q_\infty)$ one might expect in these more general situations. Equivalently, it is indeed more difficult to prove distributional convergence for the likelihood ratio under the null hypothesis.

$$\text{under } P^{(n)} \quad \log \frac{dQ^{(n)}}{dP^{(n)}} \xrightarrow{d} X \overset{d}{=} \log \frac{dQ_\infty}{dP_\infty} \text{ under } P_\infty. \quad (40)$$

Proving the above distributional convergence as $n \to \infty$ is a central task in the problem of high dimensional hypothesis testing, which has turned out to be of increasing interest in recent times Han et al. (2023), Johnstone & Onatski (2020), Onatski et al. (2013). Among other results, distributional convergence of likelihood ratio under the null to a contiguous pair of Gaussians has been shown in some restricted cases, such as below the celebrated BBP threshold Baik et al. (2005) for spiked covariance models. However, it will be extremely interesting to analyze what kind of limiting binary experiments or trade-off curves one obtains, as we believe the limiting trade-off functions might belong to a different universality class beyond the infinitely divisible class of trade-off function, with the arise of new universality classes such as the Tracy-Widom distributions Karoui (2007)[33].

3.1 SOME CONSEQUENCES WITHIN AND BEYOND INFINITE DIVISIBILITY

In our Theorem 6, we used the results of Le Cam (1986), Torgersen (1991) to show if for a sequence of functions $f_n \in \mathcal{T}, f_n(0) = 1$, we have $f_n^{\otimes n} \to f_\infty$ converge pointwise on $[0, 1]$ for some function $f_\infty \in \mathcal{T}$, then the corresponding canonical likelihood ratio $\log \frac{dQ_n^{\otimes n}}{dU^{\otimes n}}$ under $U^{\otimes n}$ converges in distribution to a limiting ratio $\log \frac{dQ_\infty}{dP_\infty}$ under P_∞ for some $Q_\infty \ll P_\infty$ so that $f_\infty = T(P_\infty, Q_\infty)$. By the fundamental theorem of infinite divisibility, the ratio under P_∞ is infinitely divisible.

Application for statistical models. Consider a statistical model $\{P_\theta : \theta \in \Theta\}$, a collection of probability measures $\{P_\theta : \theta \in \Theta\}$ on some measurable space $(\Omega, \mathcal{F})$ with $f_n = T(P_{\theta_n}, P_{\theta_n + \delta_n})$[34] with $f_n(0) = 1 \leftrightarrow P_{\theta_n + \delta_n} \ll P_{\theta_n}$, and consider $f_n^{\otimes n} = T(P_{\theta_n}^{\otimes n}, P_{\theta_n + \delta_n}^{\otimes n})$. Now, again, by Le Cam (1986), Torgersen (1991) to establish the pointwise convergence of $f_n^{\otimes n} \to f_\infty \in \mathcal{T}$ it is enough to establish that a distributional limit of the likelihood ratio under the null $P_{\theta_n}^n$ (on $\Omega^{\times n}, \mathcal{F}^{\otimes n}$)

$$\log \frac{dP_{\theta_0 + \delta_n}^{\otimes n}}{dP_{\theta_n}}(w_1, \cdots, w_n) = \sum_{i=1}^{n} \log \frac{dP_{\theta_n + \delta_n}}{dP_{\theta_n}}(w_i) \xrightarrow{d} X \text{ satisfying } \mathbb{E}[\exp(X)] = 1 \quad (41)$$

for some random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$. Then one defines on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ the distributions $P_\infty \overset{d}{=} X$[35] and $Q_\infty(h) = \mathbb{E}[h \exp(X)]$ for bounded measurable $h : (\mathbb{R}, \mathcal{B}_\mathbb{R}) \to (\mathbb{R}, \mathcal{B}_\mathbb{R})$ or equivalently $dQ_\infty(x) = e^x dP_\infty(x)$. Observe that the distributional limit assumption along with its normalization is precisely Le-Cam's first contiguity lemma for $P_{\theta_n + \delta_n}^{\otimes n} \lhd P_{\theta_n}^{\otimes n}$ van der Vaart (1998)[Lemma 6.4].

The fact that $f_\infty = T(P_\infty, Q_\infty)$ follows from the following version of Le-Cam's result, combined with the fact that for binary experiments, the convergence of Le-Cam distance for binary experiments is the same as the pointwise convergence of their corresponding trade-off functions (29).

---

[33]More precisely, Tracy-Widom distributions appear as distributional limits of the largest eigenvalues (after appropriate translation and scaling) of a spiked covariance model. as a test statistic, but not as the limit of the likelihood ratio, so it still does not rule out an infinitely divisible limit for the likelihood ratio under the null.

[34]Assume the base points $\theta_n \in \Theta$ is such that the local perturbation $\theta_n + \delta_n \in \Theta$ for all $n \in \mathbb{Z}_{\geq 1}$.

[35]Since X is a distributional limit of IID sums of random variables $\{\log \frac{dP_{\theta_n + \delta_n}}{dP_{\theta_n}}\}$, it is infinitely divisible.

Lemma 7 (Pollard (2011)[Lemma 1]). *Consider binary experiments $\mathcal{E}_n = (\Omega^{(n)}, \mathcal{F}^{(n)}, P^{(n)}, Q^{(n)})$ with $Q^{(n)} \ll P^{(n)}$ and likelihood ratio $X_n = dQ^{(n)}/dP^{(n)}$, and let $\mathcal{E}_\infty = (\Omega_\infty, \mathcal{F}_\infty, P_\infty, Q_\infty)$ with $Q_\infty \ll P_\infty$ and likelihood ratio $X_\infty = dQ_\infty/dP_\infty$ be another. Suppose the random variable*

$$(\text{under } P^{(n)}) \quad X_n = dQ^{(n)}/dP^{(n)} \xrightarrow{d} Y = dQ_\infty/dP_\infty (\text{ under } P_\infty). \text{ Then} \tag{42}$$

$\lim_{n \to \infty} \Delta(\mathcal{E}^{(n)}, \mathcal{E}_\infty) = 0$ *or equivalently* $f_n := T(P^{(n)}, Q^{(n)}) \to f_\infty = T(P_\infty, Q_\infty)$ *pointwise.*

A dominating extension. The conclusion $f_n = T(P^{(n)}, Q^{(n)}) \to T(P_\infty, Q_\infty)$ in the lemma above continues to hold when we keep the assumption $Q^{(n)} \ll P^{(n)}$, and there exists a σ-finite[36] measure $\mu^{(n)}$ on $(\Omega^{(n)}, \mathcal{F}^{(n)})$ dominating both $P^{(n)}, Q^{(n)} \ll \mu^{(n)}$ with densities $p^{(n)}, q^{(n)}$ respectively, and a σ-finite measure μ_∞ on $(\Omega_\infty, \mathcal{F}_\infty)$ dominating $P_\infty, Q_\infty \ll \mu_\infty$ with density p_∞, q_∞ with $Q_\infty \ll P_\infty$, but instead of (42) we suppose the following distributional convergence of the densities

$$(\text{under } \mu^{(n)}) \quad (p^{(n)}, q^{(n)}) \xrightarrow{d} (p_\infty, q_\infty) \quad (\text{ under } \mu_\infty). \tag{43}$$

As a consequence of the above result, we immediately obtain results for the celebrated Locally Asymptotically Normal family, the proofs of which can be found in van der Vaart (1998).

Definition 6 (van der Vaart (1998)). *A statistical model $(P_\theta : \theta \in \Theta \subset \mathbb{R}^k)$ is called differentiable in quadratic mean at θ_0 if there exists a measurable vector-valued function $\dot{\ell}_{\theta_0}$ such that, as $\theta \to \theta_0$,*

$$\int \left[\sqrt{p_\theta} - \sqrt{p_{\theta_0}} - \tfrac{1}{2}(\theta - \theta_0)^\top \dot{\ell}_{\theta_0} \sqrt{p_{\theta_0}} \right]^2 d\mu = o(\|\theta - \theta_0\|^2), \tag{44}$$

where we assume $p_\theta = \frac{dP_\theta}{d\mu}$ are the densities with-respect to a dominating measure μ on the $(\Omega, \mathcal{F})$.

Theorem 8. *van der Vaart (1998)[Theorem 7.2] Suppose that $\Theta \subset \mathbb{R}^k$ open and the model $(P_\theta : \theta \in \Theta)$ is differentiable in quadratic mean at θ. Then $P_\theta \dot{\ell}_\theta = 0$ and the Fisher information matrix $I_\theta = P_\theta \dot{\ell}_\theta \dot{\ell}_\theta^\top$ exists. Furthermore, for every converging sequence $h_n \to h$[37], as $n \to \infty$,*

$$\text{under } P_\theta^{\otimes n} \quad \log \prod_{i=1}^{n} \frac{p_{\theta + h_n/\sqrt{n}}(X_i)}{p_\theta(X_i)} \xrightarrow{d} X \overset{d}{=} N\left(-\frac{1}{2} h^\top I_\theta h, h^\top I_\theta h \right) = P_\infty \tag{45}$$

Consequently, $f_n^{\otimes n} := T(P_\theta^{\otimes n}, P_{\theta + h_n/\sqrt{n}}^{\otimes n}) \to f_\infty = T(P_\infty, Q_\infty)$, where $dQ_\infty(x) = e^x dP_\infty(x)$.

A large class of statistical models van der Vaart (1998), including most exponential families with density $p_\theta(x) = d(\theta) h(x) \exp\left(Q(\theta)^T t(x) \right)$ for appropriate choices of $d(\cdot), h(\cdot), Q(\cdot), t(\cdot)$, location models with density $p_\theta(x) = f(x - \theta)$ for appropriate $f(\cdot)$ satisfies the conclusion of the theorem.

Now, we describe the Poisson case, which is beyond Gaussian, but still belongs to the infinitely divisible class $\mathcal{I}_\mathcal{T}$. It is one of the key motivating examples for unifying with the Gaussian theory.

Corollary 9 (Poisson limits). *Consider a Bernoulli model $(P_\theta \equiv \text{Ber}(\theta) : 0 \le \theta \le 1)$ on $(\Omega, \mathcal{F}) = (\{0,1\}, 2^{\{0,1\}})$ so that $P_\theta(1) = \theta$ and $P_\theta(0) = 1 - \theta$. Now, consider $P_n = P_{\frac{\lambda_1}{n}}$ and $Q_n = P_{\frac{\lambda_2}{n}}$ for some $\lambda_1, \lambda_2 > 0$[38] and denote $f_n = T(P_n, Q_n)$. Then $f_n^{\otimes n} \to f_\infty = T(P(\lambda_1), P(\lambda_2))$, where $P(\lambda)$ denotes the Poisson distribution on $(\mathbb{Z}_{\ge 0}, 2^{\mathbb{Z}_{\ge 0}})$, $P(\lambda)(n) := e^{-\lambda} \frac{\lambda^n}{n!}$ for all $n \in \mathbb{Z}_{\ge 0}$.*

Proof. The proof follows from a *sufficiency* argument. First, observe from the likelihood ratio test

$$f_n^{\otimes n} = T(P_n^{\otimes n}, Q_n^{\otimes n}) = T(B(n, \lambda_1/n), B(n, \lambda_2/n)), \text{ where } B(n, \theta) \text{ for } \theta \in [0,1] \tag{46}$$

is the binomial distribution on $(\{0, \cdots, n\}, 2^{\{0, \cdots, n\}})$, $B(n, \theta)(k) = \binom{n}{k} \theta^k (1 - \theta)^k$ for $k \in \{0, \cdots, n\}$. This equality follows from lemma 4 $T(c_1 X + c_2, c_1 Y + c_2) = T(X, Y)$ and

$$T(P, Q) = T(L_P, L_Q) \text{ where } L_P = P \circ \left(\log \frac{dQ}{dP} \right)^{-1} \text{ and } L_Q = Q \circ \left(\log \frac{dQ}{dP} \right)^{-1}. \tag{47}$$

---

[36]A positive measure μ on $(\Omega, \mathcal{F})$ is called σ-finite if there is a sequence $\{S_k : k \in \mathbb{Z}_{\ge 1}\}$ of increasing $S_k \subset S_{k+1}$ measurable sets $S_k \in \mathcal{F}$ of finite measure $\mu(S_k) < \infty$ for all $k \in \mathbb{Z}_{\ge 1}$ so that $\cup_{k \in \mathbb{Z}_{\ge 1}} S_k = \Omega$ Our examples of μ will be counting measures on at most countable set Kallenberg (2021)

[37]We are specializing to the case of our more general model (41) with $\theta_n = \theta$ and $\delta_n = h_n/\sqrt{n}$.

[38]We are specializing to the case of our more general model (41) with $\theta_n = \lambda_1/n$ and $\delta_n = (\lambda_2 - \lambda_1)/n$.

$$P_n^{\otimes n} \circ \left(\log(\frac{dQ_n^{\otimes n}}{dP_n^{\otimes n}})\right)^{-1} = n\log\left(\frac{1-\lambda_2/n}{1-\lambda_1/n}\right) + B(n,\lambda_1/n)\log\left(\frac{\lambda_2(1-\lambda_1/n)}{(1-\lambda_2/n)\lambda_1}\right) \quad (48)$$

$$Q_n^{\otimes n} \circ \left(\log(\frac{dQ_n^{\otimes n}}{dP_n^{\otimes n}})\right)^{-1} = n\log\left(\frac{1-\lambda_2/n}{1-\lambda_1/n}\right) + B(n,\lambda_2/n)\log\left(\frac{\lambda_2(1-\lambda_1/n)}{(1-\lambda_2/n)\lambda_1}\right) \quad (49)$$

From $B(n,\lambda/n) \xrightarrow{d} P(\lambda)$ Kallenberg (2021) and lemma 7 we have for $N_1 \overset{d}{=} P(\lambda_1), N_2 \overset{d}{=} P(\lambda_2)$

$$T(P_n^{\otimes n}, Q_n^{\otimes n}) \to T\left(\lambda_1 - \lambda_2 + N_1 \log\left(\frac{\lambda_2}{\lambda_1}\right), \lambda_1 - \lambda_2 + N_2 \log\left(\frac{\lambda_2}{\lambda_1}\right)\right) = T(N_1, N_2) \quad (50)$$

$\square$

The triangular case. *Consider the inhomogeneous products of Bernoulli model $P^{(n)} = \otimes_{i=1}^n P_{n,i}$, with $P_{n,i} = \mathrm{Ber}(p_{n,i})$ and $Q^{(n)} = \otimes_{i=1}^n Q_{n,i}$ with $Q_{n,i} = \mathrm{Ber}(q_{n,i})$ where*

$$0 \leq p_{n,i}, q_{n,i} \leq 1, \quad \max_{1\leq i\leq n}(p_{n,i} \vee q_{n,i}) \to 0, \text{ and } \sum_{i=1}^{k_n} p_{n,i} \to \lambda_1, \quad \sum_{i=1}^{k_n} q_{n,i} \to \lambda_2. \text{ Then} \quad (51)$$

a similar computation yields $T(P^{(n)}, Q^{(n)}) \to T(P(\lambda_1), P(\lambda_2))$ Durrett (2019)[Theorem 3.6.1].

Proposition 1 (More infinitely divisible limits)**.** *Given trade-off functions $\{f^{(n)}, g^{(n)} : n \in \mathbb{Z}_{\geq 1}\}$*

$$\text{If } f^{(n)} \to f_\infty \in \mathcal{T} \text{ and } g^{(n)} \to g_\infty \in \mathcal{T} \text{ poinwise then } f^{(n)} \otimes g^{(n)} \to f_\infty \otimes g_\infty \in \mathcal{T}. \quad (52)$$

One can specialize to the product case $f^{(n)} := f_n^{\otimes n}, g^{(n)} = g_n^{\otimes n}$ for some $f_n, g_n \in \mathcal{T}$.

Proof. First, recall that the pointwise convergence of monotone functions $\{f^{(n)}\}_n(\{g^{(n)}\}_n)$ mapping $[0,1] \to [0,1]$ to a limiting monotone continuous function $f_\infty(g_\infty)$ mapping $[0,1] \to [0,1]$ implies uniform convergence on $[0,1]$, and so we analyze under $\|h_1 - h_2 \cdot\|_\infty^{[0,1]} = \sup\{h_1(x) - h_2(x) : x \in [0,1]\}$. Now, the proof follows from a simple *Lipschitz* argument, where we observe that

$$\|f^{(n)} \otimes g^{(n)} - f_\infty \otimes g_\infty\|_\infty^{[0,1]} \leq \|f^{(n)} - f_\infty\|_\infty^{[0,1]} + \|g^{(n)} - g_\infty\|_\infty^{[0,1]} \to 0 \quad (53)$$

This follows from the 'trick' $|G(a_n, b_n) - G(a,b)| = |G(a_n,b_n) - G(a_n,b) + G(a_n,b) - G(a,b)|$

applied to $\|g_1 \otimes f_1 - g_2 \otimes f_2\|_\infty^{[0,1]} = \|g_1 \otimes f_1 - g_1 \otimes f_2 + g_1 \otimes f_2 - g_2 \otimes f_2\|_\infty^{[0,1]} \quad (54)$

$$\leq \|g_1 \otimes f_1 - g_1 \otimes f_2\|_\infty^{[0,1]} + \|g_1 \otimes f_2 - g_2 \otimes f_2\|_\infty^{[0,1]} \overset{(a)}{\leq} \|f_1 - f_2\|_\infty^{[0,1]} + \|g_1 - g_2\|_\infty^{[0,1]} \quad (55)$$

where the inequality (a) $\|g \otimes f_1 - g \otimes f_2\|_\infty^{[0,1]} \leq \|f_1 - f_2\|_\infty^{[0,1]}$ is proven below. $\square$

Lemma 10 (Le-Cam distance under product)**.** *Consider trade off functions $g, f_1, f_2 \in \mathcal{T}$. Then,*

$$\|g \otimes f_1 - g \otimes f_2\|_\infty^{[0,1]} \leq \|f_1 - f_2\|_\infty^{[0,1]}. \quad (56)$$

Proof. We need to show that for any given $\varepsilon > 0$ if for all $x \in [0,1], |f_1(x) - f_2(x)| \leq \varepsilon$, then

$$-\varepsilon \leq g \otimes f_1(\alpha) - g \otimes f_2(\alpha) \leq \varepsilon \text{ for all } \alpha \in [0,1] \quad (57)$$

Now, denote $g = T(P,Q), f_1 = T(P_1, Q_1), f_2 = T(P_2, Q_2)$, corresponding to binary Blackwell-Le-Cam experiments $(\Omega, \mathcal{F}, P, Q) \leftrightarrow g, (\Omega_1, \mathcal{F}_1, P_1, Q_1) \leftrightarrow f_1, (\Omega_2, \mathcal{F}_2, P_2, Q_2) \leftrightarrow f_2$[39].

So, we need to show $\|T(P \otimes P_1, Q \otimes Q_1) - T(P \otimes P_2, Q \otimes Q_2)\| \leq \|T(P_1, Q_1) - T(P_2, Q_2)\|$

Now, consider $\phi_1' : \Omega \times \Omega_1 \to [0,1]$ as the (Borel measurable) Neyman-Pearson optimizer at level $\alpha = P \otimes P_1(\phi_1')$, with the value of the minimum type II error $g \otimes f_1(\alpha) = 1 - Q \otimes Q_1(\phi_1')$.

---

[39]First, the exact representation of the probability measures such as (P,Q) and the associated sample spaces $(\Omega, \mathcal{F})$ on which they are defined do not appear in the conclusion of the result. Second, this proof relies on the fact that one can construct a jointly measurable map $f(w, w_2) : \Omega \times \Omega_2 \to [0,1]$ from a collection $\{f^w(\cdot) : w \in \Omega\}$ of marginally measurable maps $f^w(w_2) : \Omega_2 \to [0,1]$, we restrict ourselves to the canonical choice of the standard Borel space $\Omega = \Omega_1 = \Omega_2 = [0,1]$ with $\mathcal{F} = \mathcal{F}_1 = \mathcal{F}_2 = \mathcal{B}_{[0,1]}$

It suffices to construct a Borel measurable $\phi_2' : \Omega \times \Omega_2 \to [0,1]$ such that $\alpha = P \otimes P_2(\phi_2')$ and $g \otimes f_2(\alpha) \leq 1 - Q \otimes Q_2(\phi_2') \leq g \otimes f_1(\alpha) + \varepsilon$. For a fixed $w \in \Omega$ define $\phi_1^{'w} : \Omega_1 \to [0,1]$ be the slice of ϕ_1' at $w \in \Omega$ as $\phi_1^{'w}(w_1) = \phi_1'(w, w_1)$, and therefore Borel measurable. Then $\phi_1^{'w}$ is a suboptimal test between P_1 and Q_1 with type I error $\alpha^w = P_1(\phi_1^{'w})$ and type II error $Q_1(1 - \phi_1^{'w}) \geq f_1(\alpha^w)$.

Now, consider the optimal (Borel measurable) Neyman-Pearson test $\phi_2^{'w} : \Omega_2 \to [0,1]$ between P_2 and Q_2 at level $\alpha^w = P_2(\phi_2^{'w})$ so that $f_2(\alpha^w) = Q_2(1 - \phi_2^{'w})$. Moreover, by patching these marginally measurable maps[40], we obtain a jointly measurable mapping $\phi_2' : \Omega \times \Omega_2 \to [0,1]$ defined as $\phi_2'(w, w_2) = \phi_2^{'w}(w_2)$. Observe that the type I error of the map ϕ_2' is given by

$$P \otimes P_2(\phi_2') = P(\alpha^w) = P \otimes P_1(\phi_1') = \alpha. \tag{58}$$

Now, consider the type II error $g \otimes f_2(\alpha) \leq Q \otimes Q_2(1 - \phi_2') = Q(f_2(\alpha^w)) \leq Q(f_1(\alpha^w)) + \varepsilon) \leq Q \otimes Q_1(1 - \phi_1') + \varepsilon = g \otimes f_1(\alpha) + \varepsilon$. As a consequence we have $g \otimes f_2(\alpha) \leq g \otimes f_1(\alpha) + \varepsilon$, and since $\alpha \in [0,1]$ was arbitrary, and moreover one can switch f_1 and f_2 in the conclusion above, since the assumption $\|f_1 - f_2\|_\infty \leq \varepsilon$ is symmetric, and therefore the result holds immediately. $\square$

Blackwell Le-Cam perspective. The proof of Lemma 10 is inspired from the proof of Dong et al. (2022)[Lemma C.1] which proves the comparison version: for trade-off functions $g, f_1, f_2 \in \mathcal{T}$ if $f_1 \leq f_2$ then $g \otimes f_1 \leq g \otimes f_2$. However, we comment that this quantitative ε generalization $\|f_1 - f_2\|_\infty^{[0,1]} \leq \varepsilon$ implies $\|g \otimes f_1 - g \otimes f_2\|_\infty^{[0,1]} \leq \varepsilon$ closely matches with the generalization of Le-Cam Le Cam (1986) going from Blackwell-comparison (0-deficiency) to ε-deficiency. More precisely, a similar result holds for the Le-Cam distance as well $\Delta(\mathcal{E} \otimes \mathcal{E}_1, \mathcal{E} \otimes \mathcal{E}_2) \leq \Delta(\mathcal{E}_1, \mathcal{E}_2)$ for binary experiments $\mathcal{E}, \mathcal{E}_1, \mathcal{E}_2$ Torgersen (1991), Le Cam (1986).

A different proof of Lemma 28 Observe that the proof of Lemma 28 was conducted by examining various Neyman-Pearson optimizers. This is because no explicit description of $g \otimes f$ was given purely in terms of trade off functions g and f. Now, by a reformulation of Torgersen (1991)[Complement 2, Chapter 10] we observe that with $Q_g'[0, \alpha] = 1 - g(\alpha)$ for $\alpha \in (0,1)$ and $Q_f'[\{0\}] = 1 - g(0)$

$$g \otimes f(\alpha) = \inf_\varphi \{Q_g'(f(\varphi(\cdot))) : ([0,1], \mathcal{B}_{[0,1]}) \xrightarrow{\varphi} ([0,1], \mathcal{B}_{[0,1]}) \text{ decreasing}, \int_0^1 \varphi(x) = \alpha\} \tag{59}$$

Now, using the variational representation above, the proof of the Lemma 28 is immediate, since $\|f_1 - f_2\|_\infty \leq \varepsilon$ implies $Q_g'(f_2(\varphi(\cdot))) \leq Q_g'(f_1(\varphi(\cdot))) + \varepsilon$ for any given (appropriate) φ, and therefore $g \otimes f_1(\alpha) \leq g \otimes f_2(\alpha) + \varepsilon$. Now, $\alpha \in [0,1]$ is arbitrary, and one can switch f_1 and f_2 as the assumption $\|f_1 - f_2\|_\infty \leq \varepsilon$ is symmetric. Hence, the result of the lemma 28 holds.

More infinitely divisible limits. A consequence of Proposition 1, the closure of the infinitely divisible trade-off functions $\mathcal{I}_\mathcal{T}$ under tensor products 5 and weak convergence[41], and given that one can explicitly derive Gaussian 8 and Poisson limits 9, one can allow $P_\infty \circ \log\left(\frac{dQ_\infty}{dP_\infty}\right)^{-1} = G + P$ to have a Gaussian G and a Poisson P component. Theoretically, given any infinitely divisible random variable X on $\mathbb{R}$ with $\mathbb{E}[\exp(\mathbb{X})] = 1$, one can construct $P_\infty \overset{d}{=} X$ and define $dQ_\infty(x) = e^x dP_\infty(x)$. Then, since X is infinitely divisible, by lemma 5 there exists $f_n = T(P_n, Q_n) \in \mathcal{I}_\mathcal{T}$ with $dQ_n = e^x P_n$ such that $T(P_n^{\otimes n}, Q_n^{\otimes n}) = T(P_\infty, Q_\infty)$. Now, as $n \to \infty$, $f_n^{\otimes n} \to f_\infty = T(P_\infty, Q_\infty)$.

Beyond infinite divisibility (mixtures) Here, we demonstrate how to go beyond the infinitely divisible privacy framework $\mathcal{I}_\mathcal{T}$. We achieve this by working under a random (but growing) number of differentially private operations[42]. Consider the canonical probability space $([0,1]^\infty, \mathcal{B}_\infty, U^\infty)$ as the countable product of the standard uniform distribution U on $[0,1]$. Now consider a sequence

---

[40]Although this is outside the scope of this paper, we believe some abstract results of the kind required to define joint Radon-Nikodym derivatives out of marginal ones Polyanskiy & Wu (2025)[Theorem 2.12], Doob (1953), Çınlar (2011) would be helpful to conclude, in our context, that indeed such a jointly measurable map ϕ_2' can be constructed. given that all the intermediate choices we made in the proof are well behaved.

[41]Weak limits of infinitely divisible distributions are infinitely divisible Kallenberg (2021)[Theorem 7.7].

[42]The intuition behind this is that in practice, often one does not know based on the same dataset S, how many differentially private statistics $(M_1(S), M_2(S, M_1(S)), \cdots)$ one has to release over time. So, we capture it as a time change sequence of integer-valued random variables $\{N_n : n \in \mathbb{Z}\}$.

of trade-off functions $\{f_n : n \in \mathbb{Z}_{\geq 1}\} \subset \mathcal{T}$ with $f_n(0) = 1$ so that $f_n = T(U, Q_{f_n})$ with $Q_n := Q_{f_n} \ll U$. Define the likelihood ratio process $\{g_n(t) := L_n(\lfloor tn \rfloor) : t \geq 0\}$ under U^∞

$$L_n(k)(\underline{x}) := \log\left(\frac{dQ_n^{\otimes k}}{dU^{\otimes k}}\right)(\underline{x}) = \sum_{i=1}^{k} \log|f_n'(1 - x_i)| \text{ for } \underline{x} \in [0,1]^\infty. \tag{60}$$

Assume that there exists a Lévy process $\{L(t) : t \geq 0\}$ so that the entire likelihood ratio process $\{g_n(t) := L_n(\lfloor tn \rfloor) : t \geq 0\}$, under U^∞ converges in distribution to the process $\{L(t) : t \geq 0\}$

$$\text{under } U^\infty \quad \{g_n(t) : t \geq 0\} \xrightarrow{d} \{L(t) : t \geq 0\}, \text{ and } \mathbb{E}[\exp(L(t))] = 1 \text{ for all } t \geq 0. \tag{61}$$

Consider a sequence $\{N_n : n \in \mathbb{Z}_{\geq 1}\}$ be integer-valued random variables on some probability space $(\Omega', \mathcal{F}', \mathbb{P}')$, and consider the joint product space $(\Omega' \times [0,1]^\infty, \mathcal{F} \otimes \mathcal{B}_\infty, \mathbb{P}' \otimes U^\infty)$, so that[43]

$$\text{for all } n \in \mathbb{Z}_{\geq 1} \text{ we have independence } N_n \perp g_n. \tag{62}$$

We assume $\mathbb{P}'$ we have $N_n/n \xrightarrow{d} \tau$ for some probability measure τ on $([0,\infty), \mathcal{B}_{[0,0\infty})$[44]. Then

$$\text{under } \mathbb{P}' \otimes U^\infty \quad g_n\left(\frac{N_n}{n}\right) = L_n(N_n) \xrightarrow{d} L(\tau), \text{ and } \mathbb{E}[\exp L(\tau)] = 1 \tag{63}$$

Moreover, by 7 we have pointwise convergence of $f_n^{\otimes N_n} = T(U^{\otimes N_n}, Q_n^{N_n}) \to f_\infty$ on $[0,1]$ where

$$f_\infty = T(P_\infty, Q_\infty), \text{ with } P_\infty \stackrel{d}{=} \text{Law}(L(\tau)) \text{ and } dQ_\infty(x) = e^x dP_\infty(x). \tag{64}$$

Basics of Levy processes. A Levy process L with $L(0) \equiv 0$ is a stochastic process $\{L(t) : t \geq 0\}$, a collection of random variables on some probability space (assumed to be right-continuous function with left limits almost surely) with independent[45] and stationary[46] increments.

Since we are interested only in weak convergence[47], we recall that marginal distributions of a Levy process $\{\mu_t : t \geq 0\}$ are all infinitely divisible satisfying $\mu_{t+s} = \mu_t * \mu_s$ for all $s, t \geq 0$ with $\mu_0 = \delta_0$. Moreover, the joint distribution of the process $\{L(t) : t \geq 0\}$ is determined by the $\mu_1 \stackrel{d}{=} L(1)$ Kallenberg (2021)[Lemma 16.1] and can be written explicitly as a sum of drifted Brownian motion

$$\{mt + \sigma B_t : t \geq 0\} \text{ for a standard Brownian motion } \{B_t : t \geq 0\}, \text{ and} \tag{65}$$

a Poisson process (see Kallenberg (2021)[Theorem 16.2] for details on the Poisson process part).

Gaussian mixtures example. For demonstrations of our results of (64) going beyond the infinitely divisible privacy framework of $\mathcal{I}_\mathcal{T}$, we restrict ourselves to limiting Levy processes $\{L(t) : t \geq 0\}$ with almost surely continuous paths, and hence reduce to the case of drifted Brownian motion Kallenberg (2021)[Proposition 16.9] $\{L(t) = mt + \sigma B_t : t \geq 0\}$ for some real number $m = -\frac{\sigma^2}{2}$[48]

For simplicity, consider the case where $\tau = \lambda_1 \delta_{t_1} + \cdots \lambda_k \delta_{t_k}$ for some $0 < t_1 < \cdots < t_k < \infty$ with $\lambda_i \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$. Then with the notation $\mu_t = N(-\frac{\sigma^2}{2}t, \sigma^2 t)$ for $t \geq 0$ we have

$$P_\infty = \lambda_1 \mu_{t_1} + \cdots + \lambda_k \mu_{t_k}, \text{ and } dQ_\infty = \lambda_1 e^x d\mu_{t_1} + \cdots + \lambda_k e^x d\mu_{t_k}, \text{ so } f_\infty = T(P_\infty, Q_\infty). \tag{66}$$

It is known that nontrivial mixtures of Gaussians are not infinitely divisible Bose et al. (2002). So, such a trade-off function $f_\infty = T(P_\infty, Q_\infty)$ is not just a departure from the asymptotic Gaussian differential privacy Dong et al. (2022), but also from our enlarged infinitely divisible privacy $\mathcal{I}_\mathcal{T}$, which we established under a large (but deterministic) number of differentially private operations.

---

[43]This independence assumption precisely captures the intuition that we do not know in practice (and hence model it as random) how many rounds of differentially private operations one has to run on a given dataset S.

[44]Under independence assumption $\{X_n\}_n \perp \{Y_n\}_n$ marginal convergence $X_n \xrightarrow{d} X$, $Y_n \xrightarrow{d} Y$ implies joint convergence $(X_n, Y_n) \to (X, Y)$ van der Vaart (1998). We apply this to $(g_n, N_n/n) \xrightarrow{d} (g, \tau)$

[45]It means $(L(t_1), L(t_2) - L(t_1), \cdots, L(t_k) - L(t_{k-1}))$ are jointly independent for all $0 \leq t_1 < \cdots < t_k$.

[46]Stationarity $L(t) - L(s) \stackrel{d}{=} L(t - s)$ is dropped, when one analyses triangular systems $f_{n1} \otimes \cdots \otimes f_{nN_n}$.

[47]We mean weak convergence of the entire process, which implies convergence of finite dimensional distributions and much more. We refer to Johnstone & Onatski (2020)[Conclusions] for more details on this.

[48]This constraint comes from the restriction $\mathbb{E}[\exp L(t)] = \mathbb{E}[\exp(mt + \sigma B_t)] = 1$ for all $t \geq 0$.
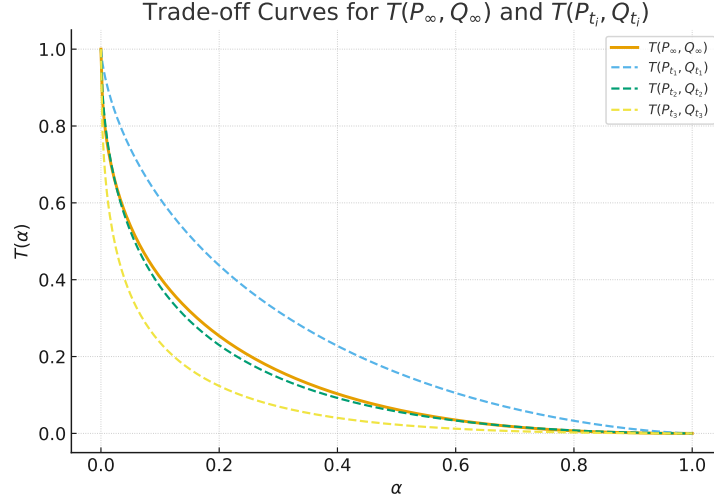
Figure 1: Trade-off curves for the Gaussian mixture experiment $T(P_\infty, Q_\infty)$ and for the component Gaussian experiments $T(P_{t_i}, Q_{t_i})$, $i = 1, 2, 3$. Here $P_\infty = \sum_{i=1}^{3} \lambda_i \mu_{t_i}$ and $dQ_\infty(x) = \sum_{i=1}^{3} \lambda_i e^x \, d\mu_{t_i}(x)$, with $\mu_{t_i} = \mathcal{N}(-\frac{\sigma^2}{2} t_i, \sigma^2 t_i)$. The solid curve shows $T(P_\infty, Q_\infty)$, while the dashed curves show the individual components $T(P_{t_1}, Q_{t_1})$, $T(P_{t_2}, Q_{t_2})$, and $T(P_{t_3}, Q_{t_3})$.

Blackwell-Le-Cam perspective. The Gaussian mixture example (P_∞, Q_∞) with the mixing distribution $\tau = \lambda_1 \delta_{t_1} + \cdots \lambda_k \delta_{t_k}$ can be seen as a *mixture* of binary experiments $\mathcal{E}_i = (P_{t_i}, Q_{t_i})$ where $P_{t_i} = \mu_{t_i} = N(-\frac{\sigma^2}{2} t_i, \frac{\sigma^2}{2} t_i)$ and $dQ_{t_i}(x) = e^x dP_{t_i}(x)$, all of which are on $\Omega_i = \mathbb{R}, \mathcal{F}_i = \mathcal{B}_\mathbb{R}$. P_∞ can also be seen to fit within the framework of locally asymptotic mixed normality ('Normal' distributions with random variances), as introduced in Jeganathan (1982), and has recently also appeared in the asymptotics of (local) differentially private statistical analysis Steinberger (2024).

Moreover, a reformulation of Torgersen (1991)[Complement 2, Chapter 10] is that for a binary mixture experiment (P_∞, Q_∞) by mixing experiments $\{(P_i, Q_i) : 1 \le i \le k\}$ so that $P_\infty = \sum_{i=1}^{k} \lambda_i P_i$, and $Q_\infty = \sum_{i=1}^{k} \lambda_i Q_i$. Then we have for $f_\infty = T(P_\infty, Q_\infty)$ and $f_i = T(P_i, Q_i)$

$$f_\infty(\alpha) = \inf_{0 \le \alpha_i \le 1} \{ \sum_{i=1}^{k} \lambda_i f_i(\alpha_i) : \sum_{i=1}^{k} \alpha_i = \alpha \} \text{ for all } \alpha \in [0, 1] \tag{67}$$

Wald-Wolfowitz extension. In this 'mixture' framework above, we assumed that the random (but growing) number $\{N_n\}_{n \ge 1}$ of differentially private operations are independent of the 'outcomes' of the differentially private operations $M^{(n,k)}(\cdot) = (M_{n1}(\cdot), \cdots, M_{nk}(\cdot))$ on a given dataset S.

Under the sequential analysis, one might consider stopping at time N_n based on the release of the previous outputs $M^{n, N_n - 1}(S)$, if enough information about the data has already been released by that time. Although this is beyond the scope of this paper, we think that, from a practical point of view, it is an extremely interesting direction to pursue, where we believe a different analysis, more in line with Wald-Wolfowitz's framework of the sequential probability ratio test Siegmund (1985), along with the techniques of de la Peña & Giné (1999), de la Peña et al. (2009), Silvestrov (2004).

We believe that the result 64 would still hold, where τ would be a stopping time with respect to the Filtration $\{\mathcal{F}_t = \sigma(L(s) : 0 \le s \le t) : t \ge 0\}$ generated by the Levy process $\{L(t) : t \ge 0\}$ [49].

---

[49]To be mathematically precise, in concluding, $\mathbb{E}[\exp(L(\tau))] = 1$ we need some uniform integrability conditions on the sequence $\{\exp(L_n(N_n)) : n \in \mathbb{Z}_{\ge 1}\}$ to apply Martingale optional stopping theorems.
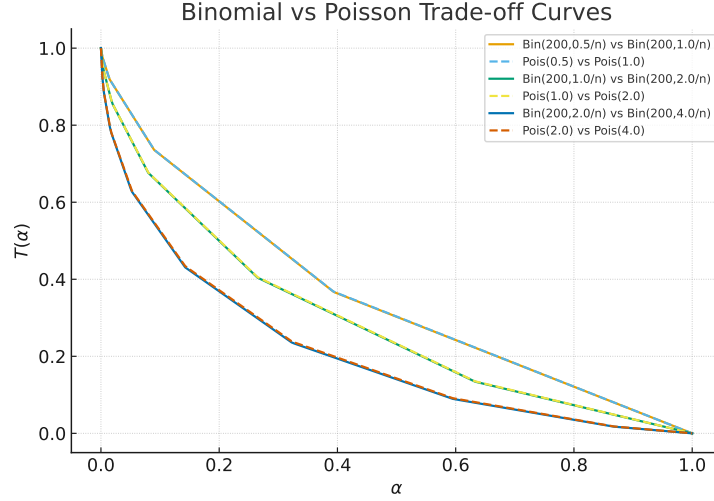
Figure 2: Trade-off curves for testing between $\mathrm{Bin}(n, \lambda_1/n)$ and $\mathrm{Bin}(n, \lambda_2/n)$ (solid lines, with $n = 200$) compared to the Poisson limits $T(\mathrm{Pois}(\lambda_1), \mathrm{Pois}(\lambda_2))$ (dashed lines), for several pairs (λ_1, λ_2). The curves are nearly indistinguishable, illustrating the Poisson approximation at the level.

4 HOW TO BE POISSON DIFFERENTIALLY PRIVATE IN PRACTICE

Now, we describe a mechanism for achieving a baseline notion of Poisson differential privacy $f_\mu = T(P(1), P(\mu))$ for Poissons $P(\cdot)$ with mean $\mu > 1$ 2. The purpose is similar to the noise-adding mechanism results for achieving Gaussian differential privacy $f_\mu = T(Z, Z+\mu)$ for $Z \sim N(0, 1)$ or, more generally, $f_\mu = T(X, X + \mu)$ for a symmetric log-concave random variable X 24. However, as we will demonstrate, for baseline privacy families such as $\{P(\lambda) : \lambda > 1\}$ with respect to $P(1)$, that are not inherently a shift family including $\{X + \mu : \mu > 0\}$ with respect to X (distribution of X is symmetric and log-concave), noise adding mechanisms is not the natural way to achieve privacy.

Consider a metric space $(\mathcal{X}, d)$[50] It consists of the allowable datasets S and a real-valued statistic $g : (\mathcal{X}, d) \to \mathbb{R}$. Denote the sensitivity or the modulus of continuity at 1 as[51]

$$w_g(1) := \sup_{d(x,y) \leq 1} |g(x) - g(y)| \quad \text{satisfying} \quad w_{cg}(1) = c w_g(1) \text{ for any } c > 0. \quad (68)$$

Lemma 11 (Modulus of continuity). *For a map $(X, d_X) \xrightarrow{g} (Y, d_Y)$ between metric spaces, define*

$$w_g : [0, \infty] \to [0, \infty] \quad w_g(\delta) = \sup_{d_X(x_1, x_2) \leq \delta} d_Y(g(x_1), g(x_2)). \quad (69)$$

Consider the composition $(X, d_X) \xrightarrow{g} (Y, d_Y) \xrightarrow{f} (Z, d_Z)$ then $w_{f \circ g}(\delta) \leq w_f(w_g(\delta))$ for all $\delta \geq 0$.

Proof. Define $L = w_{f \circ g}(\delta) = \sup_{d_X(x_1, x_2) \leq \delta} d_Z(f(g(x_1)), f(g(x_2)))$. Now for all $x_1, x_2 \in X$

$$d_X(x_1, x_2) \leq \delta \to d_Y(g(x_1), g(x_2)) \leq w_g(\delta). \quad (70)$$

Consequently, $\{(g(x_1), g(x_2)) : d_X(x_1, x_2) \leq \delta\} \subset \{(y_1, y_2) : d_Y(y_1, y_2) \leq w_g(\delta)\}$. Therefore

$$L \leq \sup_{d_Y(y_1, y_2) \leq w_g(\delta)} d_Z(f(y_1), f(y_2)) \leq w_f(w_g(\delta)) \quad (71)$$

$\square$

---

[50]The definition of a metric space includes the product space example $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$ under the Hamming metric $d_H(\underline{x}, \underline{y}) = \sum_{i=1}^{n} \mathbf{1}(x_i \neq y_i)$, capturing how many coordinates $\underline{x}$ and $\underline{y}$ differ.

[51]The modulus of continuity $w_g(\delta)$ is an important measure of continuity of g. In the literature on statistical estimation Donoho & Liu (1991) under differential privacy constraints Rohde & Steinberger (2020), Duchi & Ruan (2024), one usually considers the metric space $\mathcal{X}$ as an appropriate subset of Borel probability measures $\mathcal{P}(\mathbb{R}^d, \mathcal{B}_{\mathbb{R}^d})$ with an appropriate choice of the metric d, including total variation and Hellinger distance.

Theorem 12. *Dong et al. (2022)[Theorem 1, Proposition A.3] Given a metric space $(\mathcal{X}, d)$ of datasets and a real-valued statistic $g : (\mathcal{X}, d) \to \mathbb{R}$, consider the noise-adding mechanism (Markov Kernel) $M : (\mathbb{R}, \mathcal{B}_\mathbb{R}) \to (\mathbb{R}, \mathcal{B}_\mathbb{R})$ as $M(g(x)) = g(x) + X$ for a symmetric-log-concave X [52]. Then*

$$\inf_{d(x,y) \leq 1} T(M(g(x)), M(g(y))) \geq T(X, X + w_g(1)) \tag{72}$$

Consequently, the mechanism M satisfies $f^X_{w_g(1)}$-DP with $f^X_\mu = T(X, X + \mu)$ for any $\mu > 0$.

Noise level. Observe that having fixed the noise distribution X, the noise level $\mu > 0$ and the statistic g before implementing the mechanism M, one can scale $g \to cg$ so that $w_{cg}(1) = cw_g(1) = \mu$ (if the mechanism designer is allowed to scale the original statistic g) or equivalently scale the noise $X \to cX$ with c so that $T(cX, cX + w_g(1)) = T(X, X + c^{-1}w_g(1))$ with $c^{-1}w_g(1) = \mu$.

Now, we describe a mechanism that achieves an asymmetric version of Poisson differential privacy.

Theorem 13 (Poisson differential privacy). *Given a metric space $(\mathcal{X}, d)$ of datasets and a count statistic $g : (\mathcal{X}, d) \to \mathbb{Z}_{\geq 0}$ [53] with the baseline Poisson trade-off function $f_\infty = T(P(\mu_1), P(\mu_2))$ for some $\mu_2 > \mu_1 > 0$. Consider the mechanism $M : (\mathbb{Z}_{\geq 0}, 2^{\mathbb{Z}_{\geq 0}}) \to (\mathbb{Z}_{\geq 0}, 2^{\mathbb{Z}_{\geq 0}})$ defined as*

$$M(g(x)) \sim P(N_2 e^{N_1 g(x)}), \text{ with } N_1 = \frac{\log\left(\frac{\mu_2}{\mu_1}\right)}{w_g(1)} \text{ and } N_2 = \frac{|\mu_2 - \mu_1|}{w_{h \circ g}(1)}, \text{ where} \tag{73}$$

$$h : (\mathbb{R}, |\cdot|) \to (\mathbb{R}, |\cdot|) \quad h(y) = e^{N_1 y} = \left(\frac{\mu_2}{\mu_1}\right)^{\frac{y}{w_g(1)}}. \tag{74}$$

Then M satisfies the following (asymmetric) version of Poisson differential privacy [54].

$$\inf_{d_{\mathcal{X}}(x_1, x_2) \leq 1 : g(x_2) \geq g(x_1)} T(M(g(x_1)), M(g(x_2))) \geq T(P(\mu_1), P(\mu_2)) = f_\infty \tag{75}$$

$$\inf_{d_{\mathcal{X}}(x_1, x_2) \leq 1 : g(x_2) \leq g(x_1)} T(M(g(x_1)), M(g(x_2))) \geq T(P(\mu_2), P(\mu_1)) = f_\infty^{-1} \tag{76}$$

Consequently, M satisfy (symmetric) f_{μ_1, μ_2}-DP 2 where $f_{\mu_1, \mu_2} = \min(f_\infty, f_\infty^{-1})$. More precisely,

$$\inf_{d_{\mathcal{X}}(x_1, x_2) \leq 1} T(M(g(x_1)), M(g(x_2))) \geq f_{\mu_1, \mu_2} \tag{77}$$

Proof. For parameters $\lambda'_1, \lambda'_2, \mu'_1, \mu'_2 > 0$, if (λ'_1, λ'_2) and (μ'_1, μ'_2) are similarly ordered (this means either $\lambda'_1 < \lambda'_2$ and $\mu'_1 < \mu'_2$ simultaneously or $\lambda'_2 < \lambda'_1$ and $\mu'_2 < \mu'_1$ simultaneously) and satisfy

sufficient conditions: $\quad |\lambda'_2 - \lambda'_1| \leq |\mu'_2 - \mu'_1|$ and $\dfrac{\max(\lambda'_1, \lambda'_2)}{\min(\lambda'_1, \lambda'_2)} \leq \dfrac{\max(\mu'_1, \mu'_2)}{\min(\mu'_1, \mu'_2)}$ $\tag{78}$

Then by taking $0 < c := \frac{|\lambda'_2 - \lambda'_1|}{|\mu'_2 - \mu'_1|} \leq 1$ and a $\lambda \geq 0$ defined in the following way [55]

$$\lambda := \lambda'_1 - c\,\mu'_1 = \begin{cases} \dfrac{\lambda'_1 \mu'_2 - \lambda'_2 \mu'_1}{\mu'_2 - \mu'_1} & \text{when } \lambda'_2 > \lambda'_1 \text{ and } \mu'_2 > \mu'_1 \text{ satisfying (78)}, \\[2ex] \dfrac{\lambda'_2 \mu'_1 - \lambda'_1 \mu'_2}{\mu'_1 - \mu'_2} & \text{when } \lambda'_1 > \lambda'_2 \text{ and } \mu'_1 > \mu'_2 \text{ satisfying (78)}. \end{cases}$$

---

[52] With an abuse of notation, we denote the mechanism M and the corresponding Markov kernel $M(z, A) = \mathbb{P}(z + X \in A) = P_X(A - z)$ for $A \subset \mathcal{B}_\mathbb{R}$ and $z \in \mathbb{R}$, where $A - z$ is the Minkowski translate $\{A - z\} := \{a - z : a \in A\}$, and P_X is a symmetric ($P_X(A) = P_X(-A)$, with $-A := \{-a : a \in A\}$), log-concave ($x \to \log \frac{dP_X}{d\lambda}(x)$ is a concave function on the convex support (Range(X))) measure, the distribution of X.

[53] We are trying to achieve Poisson differential privacy, which is suitable when we expect the released outputs to be count statistics (elements of $\mathbb{Z}_{\geq 0}$). So, its natural to consider the original statistic g to also take values in $\mathbb{Z}_{\geq 0}$. However, a quick inspection of the proof reveals that our results hold for any real valued statistic g, since nowhere in the proof do we use the fact that g takes only integer values or non-negative values.

[54] For $f = T(P, Q)$ implies the generalized inverse $f^{-1} = T(Q, P)$ Dong et al. (2022)[Lemma A.2].

[55] (c, d) captures the (slope, intercept) pair of the line $y = cx + d$ joining the points (μ'_1, λ'_1) and (μ'_2, λ'_2).

we have $\lambda_1' = c\mu_1' + \lambda$ (by definition of λ), and $\lambda_2' = c\mu_2' + \lambda$ holds because of the following.

$$
\lambda_2' - c\,\mu_2' = \begin{cases} \dfrac{\lambda_1'\mu_2' - \lambda_2'\mu_1'}{\mu_2' - \mu_1'} = \lambda & \text{when } \lambda_2' > \lambda_1' \text{ and } \mu_2' > \mu_1' \text{ satisfying (78),} \\[3mm] \dfrac{\lambda_2'\mu_1' - \lambda_1'\mu_2'}{\mu_1' - \mu_2'} = \lambda & \text{when } \lambda_1' > \lambda_2' \text{ and } \mu_1' > \mu_2' \text{ satisfying (78).} \end{cases}
$$

Then applying (97) (Thinning) and (98) (Superposition) of Lemma 21 we have the following chain of inequalities for any similarly ordered (all of them > 0) (λ_1', λ_2') and (μ_1', μ_2') satisfying (78)

$$
T(P(\mu_1'), P(\mu_2')) \overset{(97)}{\leq} T(P(c\mu_1'), P(c\mu_2')) \overset{(98)}{\leq} T(P(c\mu_1' + \lambda), P(c\mu_2' + \lambda)) = T(P(\lambda_1'), P(\lambda_2')). \tag{79}
$$

Now, it remains to show with the choice of N_1, N_2 as mentioned in the statement of Theorem 13

$$
N_1 = \frac{\log\left(\frac{\mu_2}{\mu_1}\right)}{w_g(1)} \text{ and } N_2 = \frac{|\mu_2 - \mu_1|}{w_{h\circ g}(1))} \text{ with } h(y) = e^{N_1 y} = \left(\frac{\mu_2}{\mu_1}\right)^{\frac{y}{w_g(1)}} \text{ for a given statistic} \tag{80}
$$

$g : (\mathcal{X}, d) \to \mathbb{Z}_{\geq 0}$ and a baseline trade-off function $f_\infty = T(P(\mu_1), P(\mu_2))$ for some $\mu_2 > \mu_1 > 0$ that (78) holds for any $(\lambda_1', \lambda_2') = (N_2 e^{N_1 g(x_1)}, N_2 e^{N_1 g(x_2)})$ against $(\mu_1', \mu_2') = (\mu_1, \mu_2)$ as

$$
\sup_{d_{\mathcal{X}}(x_1, x_2) \leq 1 : g(x_2) \geq g(x_1)} \frac{\exp(N_1 g(x_2))}{\exp(N_1 g(x_1))} \leq \frac{\mu_2}{\mu_1} = \frac{\max(\mu_1, \mu_2)}{\min(\mu_1, \mu_2)} \leftrightarrow N_1 w_g(1) \leq \log\left(\frac{\mu_2}{\mu_1}\right) \text{ and} \tag{81}
$$

$$
\sup_{d_{\mathcal{X}}(x_1, x_2) \leq 1 : g(x_2) \geq g(x_1)} |N_2 e^{N_1 g(x_1)} - N_2 e^{N_1 g(x_2)}| \leq |\mu_2 - \mu_1| \leftrightarrow N_2 w_{h\circ g}(1)) \leq |\mu_2 - \mu_1| \tag{82}
$$

Now, having verified the sufficient conditions for every pair (x_1, x_2) such that $d_{\mathcal{X}}(x_1, x_2) \leq 1$ and $g(x_2) \geq g(x_1)$, we have the following by applying the conclusion of (79)

$$
\inf_{d_{\mathcal{X}}(x_1, x_2) \leq 1 : g(x_2) \geq g(x_1)} T\left(P(N_2 e^{N_1 g(x_1)}), P(N_2 e^{N_1 g(x_2)})\right) \geq T(P(\mu_1), P(\mu_2)) = f_\infty. \tag{83}
$$

This concludes the proof of (75), and by symmetry[56] we have (76) (same choice of N_1, N_2 as above)

$$
\inf_{d_{\mathcal{X}}(x_1, x_2) \leq 1 : g(x_2) \leq g(x_1)} T\left(P(N_2 e^{N_1 g(x_1)}), P(N_2 e^{N_1 g(x_2)})\right) \geq T(P(\mu_2), P(\mu_1)) = f_\infty^{-1}. \tag{84}
$$

Now, combining the two above (83), (84) we have the following (77).

$$
\inf_{d_{\mathcal{X}}(x_1, x_2) \leq 1} T(M(g(x_1)), M(g(x_2))) \geq \min(f_\infty, f_\infty^{-1}) =: f_{\mu_1, \mu_2}. \tag{85}
$$

$\square$

Now, we remark one several aspects of this mechanism M along with extensions of the result.

Remark 14 (Double conjugate). *$f_{\mu_1, \mu_2} = \min(T(P(\mu_1), P(\mu_2)), T(P(\mu_2), P(\mu_1)))$ for $\mu_1, \mu_2 > 0$ is not a trade-off function, since the minimum of two convex functions is not a convex function in general, and it does not hold in the Poisson trade-off functions (see figure 4 for a counterexample).*

Therefore, to fit with the f-DP framework of Dong et al. (2022) for a trade-off function $f \in \mathcal{T}$ 2, we replace f_{μ_1, μ_2} with its double conjugate $g = f_{\mu_1, \mu_2}^{**} : [0, 1] \to [0, 1]$, which is the (convex envelope from below) largest lower-semi-continuous convex function g lying below $f_{\mu_1, \mu_2} = \min(f_\infty, f_\infty^{-1})$ Rockafellar (1970). However, for a trade-off function $f \in \mathcal{T}$ we have an explicit expression for the double conjugate $g = \min(f, f^{-1})^{**} \in \mathcal{T}$, following Dong et al. (2022)[Prop E.1].

---

[56]The required sufficient conditions (78) to establish for the choice of N_1, N_2 as well as its conclusions (79) are symmetric with respect to order reversing map $(\lambda_1', \lambda_2'), (\mu_1', \mu_2') \to (\lambda_2', \lambda_1'), (\mu_2', \mu_1')$

For $f \in \mathcal{T}$, define the first time the slope (derivative) of f matches that of the function $y(x) = 1 - x$ (ideal privacy curve) as $\bar{x} := \bar{x}_f = \inf\{x \in [0, 1] | -1 \in \partial f(x)\}$, where $\partial f(x)$ is the sub-gradient of the trade-off function f at x Rockafellar (1970). Then in the case $\bar{x} \leq f(\bar{x})$ we have

$$g = \min\{f, f^{-1}\}^{**}(x) = \begin{cases} f(x), & x \in [0, \bar{x}], \\ \bar{x} + f(\bar{x}) - x, & x \in [\bar{x}, f(\bar{x})], \\ f^{-1}(x), & x \in [f(\bar{x}), 1]. \end{cases} \tag{86}$$

When $\bar{x} > f(\bar{x})$ below, see figure 3 (see Dong et al. (2022)[Figure 6, Appendix] for $\bar{x} \leq f(\bar{x})$ case)

$$g = \min\{f, f^{-1}\}^{**}(x) = \begin{cases} f^{-1}(x), & x \in [0, f(\bar{x})], \\ \bar{x} + f(\bar{x}) - x, & x \in [f(\bar{x}), \bar{x}], \\ f(x), & x \in [\bar{x}, 1]. \end{cases} \tag{87}$$
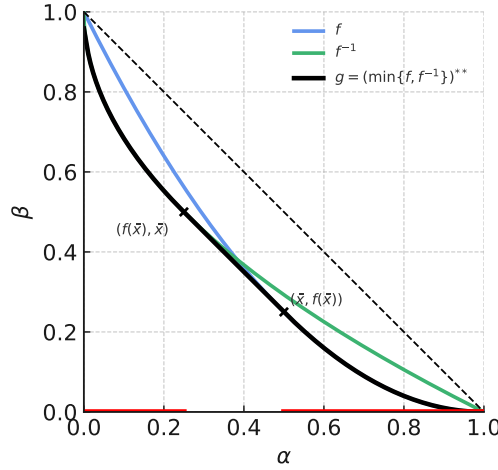

Figure 3: Double convex conjugate. Here, the base trade-off function is $f : [0, 1] \to [0, 1]$ given by $f(x) = 1 - 2x + x^2$, for $x \in [0, 1]$ and and its functional inverse on $[0, 1]$ is $f^{-1}(y) = 1 - \sqrt{y}$, for $y \in [0, 1]$. The black curve shows $g = \left(\min\{f, f^{-1}\}\right)^{**}$, the double convex conjugate of $\min\{f, f^{-1}\}$. The two marked points are $(f(\bar{x}), \bar{x})$ and $(\bar{x}, f(\bar{x}))$, with $\bar{x} = 1/2$ and $f(\bar{x}) = 1/4$.

Remark 15 (Tightness of our mechanism). *It has been proved in Dong et al. (2022)[Theorem 1] (see also Theorem 12), that for a Gaussian baseline trade-off function $f_\mu = T(N(0, 1), N(\mu, 1))$[57], Gaussian noise mechanism $M(g(x)) = g(x) + cZ$ with $c^{-1}w_g(1) = \mu$ tightly captures f_μ-DP.*

This means that the inequality in Equation (72) is an equality and it happens as soon as there exist two datapoints x_1, x_2 with $d_{\mathcal{X}}(x_1, x_2) \leq 1$ achieve the supremum $w_g(1) = |g(x_2) - g(x_1)|$.

Similarly, if there exist two data points x_1, x_2 with $d_{\mathcal{X}}(x_1, x_2) \leq 1$ achieve the supremum jointly,

$$w_g(1) = |g(x_2) - g(x_1)| \text{ and } w_{h \circ g}(1) = |h(g(x_2)) - h(g(x_1))|. \tag{88}$$

Then we have simultaneous equality in Equations (75) and (76)

Therefore, the mechanism $M(g(x)) \sim P(N_2 e^{N_1 g(x)})$ achieves tightness (75) and (76) together (or individually based on what we require on Equation (88)). However, because of the inherent asymmetry of the Poisson trade-off curves $T(P(\lambda_1), P(\lambda_2)) \neq T(P(\lambda_2), P(\lambda_1))$, and the fact that Blackwell ordering (5) is a partial order (see figure 5 for a proof of this incomparability)

$$T(P(\lambda_1), P(\lambda_2)) \not\succeq T(P(\lambda_2), P(\lambda_1)) \text{ and } T(P(\lambda_2), P(\lambda_1)) \not\succeq T(P(\lambda_1), P(\lambda_2)). \tag{89}$$

We have that the mechanism $M(g(x)) = P(N_2 e^{N_1 g(x)})$ proposed in Theorem 13 overall achieves a $\min(f_\infty, f_\infty^{-1})^{**}$- differential privacy 2 , where $f_\infty = T(P(\lambda_1), P(\lambda_2))$.
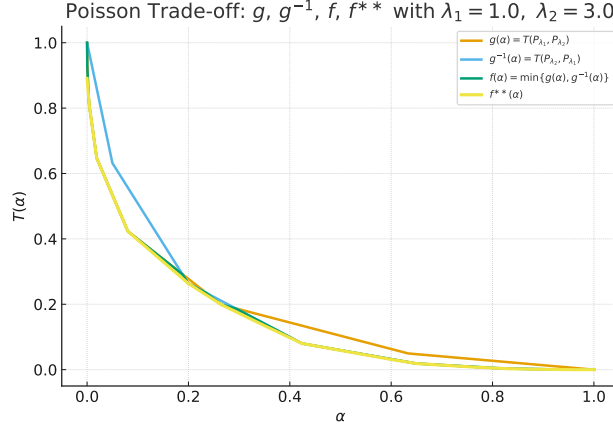
Figure 4: Trade-off functions for a Poisson experiment. Let $P_{\lambda_1} = \mathrm{Pois}(\lambda_1)$ and $P_{\lambda_2} = \mathrm{Pois}(\lambda_2)$ with $\lambda_1 = 1$ and $\lambda_2 = 3$. The blue curve shows $g(\alpha) = T(P_{\lambda_1}, P_{\lambda_2})(\alpha)$, and the orange curve shows $g^{-1}(\alpha) = T(P_{\lambda_2}, P_{\lambda_1})(\alpha)$. The green curve corresponds to $f(\alpha) = \min\{g(\alpha), g^{-1}(\alpha)\}$, and the black curve shows the double convex conjugate $f^{**}(\alpha) = \left(\min\{g, g^{-1}\}\right)^{**}(\alpha)$.
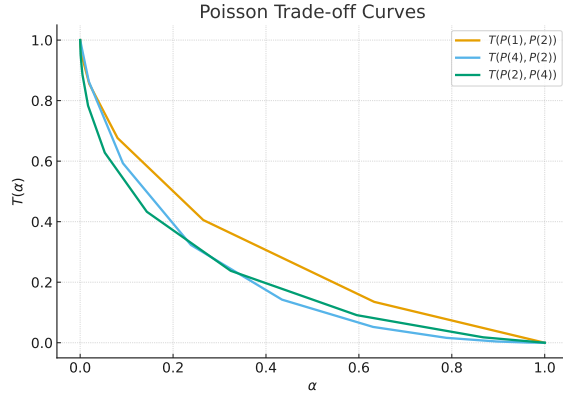

Figure 5: Trade-off curves for Poissons: $T(P(1), P(2))$, $T(P(4), P(2))$, and $T(P(2), P(4))$.

Remark 16 (Asymmetry in Poisson). *Observe that $T(P(\lambda_1), P(\lambda_2)) \neq T(P(\lambda_2), P(\lambda_1))$ (see figure 4), as opposed to the symmetry of trade-off curves corresponding to the shift (location) family $T(X, X + \mu) = T(X + \mu, X)$ for symmetric-log concave X and $\mu > 0$. This asymmetry also happens because of the one-sided nature of the support of $P(\lambda)$ being $\mathbb{Z}_{\geq 0}$. Therefore, we believe it is not natural to consider a symmetric (around zero) noise-adding mechanism (discrete or continuous) for privacy concerning count statistics g that are inherently non-negative integer valued.*

Remark 17 (Utility of the framework). *The utility of the f-differential privacy framework of Dong et al. (2022) lies in the ability to choose the baseline f based on the class of problems at hand. For count statistics, it is natural to have a comparable Poisson trade-off function as the baseline rather than two Gaussians, which are inherently continuous. In the Poisson case (see Theorem 13), instead of adding Poisson noise, having received the value of the statistic $g(x)$, the mechanism M outputs a sample from the Poisson distribution with parameter $N_2 e^{N_1 g(x)}$.*

Remark 18 (Almosus sure Implementation in practice). *First, given the privacy level $\mu_2 > \mu_1 > 0$, and the statistic g the computation of the constants N_1, N_2 (see Equation (73)) is immediate, once the modulus of continuity values $w_g(1), w_{h \circ g}(1)$ have been computed[58] (see Equation (74) for the*

---

[57]This holds more generally for a shift family $f_\mu = T(X, X + \mu)$ with symmetric log concave X.

[58]The values $w_g(1), w_{h \circ g}(1)$ can be unbounded, unless the range of g is bounded, which it will be in many use case scenarios of count statistics, such as degree of a vertex in a sparse network Karwa & Slavković (2016).

definition of h). In Theorem 13, we have given only the distributional information that for all $x \in \mathcal{X}$ we have $M(g(x)) \sim P(\lambda(g(x)))$, where $\lambda(n) = N_2 e^{N_1 n}$, but one can prescribe a sample by sample (almost sure) implementation using the standard conversion between the cumulative distribution function F and the quantile function F^{-1}. In the Poisson case, more precisely, let

$$M(n) \sim P(\lambda(n)), \text{ and } p_n(k) := P(\lambda(n))(\{k\}) = e^{-\lambda(n)} \frac{\lambda(n)^k}{k!}, \text{ for } k \in \mathbb{Z}_{\geq 0}, \quad (90)$$

Denote the Cumulative distribution function $F_n : \mathbb{R} \to [0,1]$ for $P(\lambda(n))$, for $k \in \mathbb{Z}_{\geq 0}$ as[59]

$$F_n(k) := P(\lambda(n))(\{0, \cdots, k\}) = \sum_{j=0}^{k} p_j(n), \text{ with } F_n(-1) := 0, \text{ and} \quad (91)$$

$$0 = F_n(-1) < F_n(0) < F_n(1) < \cdots \leq 1, \text{ and} \quad \lim_{k \to \infty} F_n(k) = 1. \quad (92)$$

Now, almost surely set $M(g(x)) = k$ if and only if $F_n(k-1) \leq U < F_n(k)$. for $n = g(x)$, and U is a uniformly random sample from $[0,1]$ (independent of $g(x))$[60]. Moreover, for each $x \in \mathcal{X}$

$$\mathbb{P}\big(M(g(x)) = k\big) = U(\{u \in [0,1] : F_n(k-1) \leq u < F_n(k)\}) = P(\lambda(n))(\{k\}) \quad k \in \mathbb{Z}_{\geq 0}. \quad (93)$$

So $M(g(x)) \sim P(\lambda(g(x)))$ satisfies the distributional requirement[61]. As is immediate, any Markov kernel M with one-dimensional output $M(g(x))$ and a different sampling distribution $F_{g(x)}$ from $P(\lambda(g(x)))$ can also be implemented for a released statistic $g(x)$ in a similar way, almost surely as $F_{g(x)}^{-1}(U)$[62] where $F_{g(x)}^{-1} : [0,1] \to \mathbb{R}$ is the quantile function defined as

$$F_{g(x)}^{-1}(u) = \inf\{t \in \mathbb{R} : F_{g(x)}(t) \geq u\} \quad (94)$$

Remark 19 (Infinitely divisible extension). *It is an interesting future direction to generalize the result of Theorem 13 to describe differentially private mechanisms covering the cases of other baseline trade-off functions $f = T(P, Q)$ with infinitely divisible distributions, such as Geometric, discrete Laplace, and Negative-binomial, more generally for a exponential family of the following form*

$$f_{\theta_1, \theta_2} = T(P_{\theta_1}, P_{\theta_2}) \text{ where } \{P_\theta : dP_\theta(x) = \frac{e^{\theta x} dP(x)}{P(e^{\theta x})}\}_{\theta \in [-1,1]} \text{ with } T(P_0, P_1) \in \mathcal{I}_{\mathcal{T}}. \quad (95)$$

It would also be interesting to analyse how distribution-specific $(T(P, Q))$ mechanisms [63] would compare with noise addition mechanisms such as Skellam Agarwal et al. (2021), discrete Gaussian Canonne et al. (2020)[64], and canonical noise-distribution framework Awan & Vadhan (2023) .

Remark 20 (Natural apperance of Poisson trade-off functions). *We demonstrate how the Poisson differentially private framework with a baseline trade-off function $f_\infty = T(P(\lambda_1), P(\lambda_2))$ would be an appropriate benchmark for privacy on a graph-based relational private dataset.*

More precisely, consider a graph-based private dataset $G_n = ([n], E_n)$, where we would like to protect the community label vector $\beta \in \{0,1\}^n$ of the vertices $\{1, \cdots, n\}$ upon sequential queries on the labels of the vertices. Since a single coordinate of the entire mechanism $M = (M_{n1}, \cdots, M_{nn})$, where M_{ni} releases a noisy label for vertex i, is binary, it is natural to consider that for each M_{ni} (for each $1 \leq i \leq n$) satisfies f_n-DP 2 for a baseline trade-off function $f_n = T(Ber(\lambda_1/n), Ber(\lambda_2/n))$. Then, as $n \to \infty$ or, equivalently, as the number of vertices within the network grows, upon releasing the (noisy) labels of the vertices, the joint privacy level is given by $f_n^{\otimes n}$ and converges to a Poisson trade-off curve $f_\infty = T(P(\lambda_1), P(\lambda_2))$ 9.

---

[59]F_n remains constant between points of jump $F_n(x) = F_n(k)$ for all $k \leq x < k+1$ for all $k \in \mathbb{Z}_{\geq -1}$.

[60]This is the step where we are using external randomness. This is very similar to how one injects independent centered Gaussian noise G (of appropriate variance) into a released statistic $g(x)$ to generate a sample from a Gaussian with mean $g(x)$ to achieve Gaussian differential privacy 12.

[61]Observe that, since in practice, one only outputs one noisy output $M(g(x))$ given a query request $g(x)$, and hence there is no requirement on the joint distributions of the output $\{M(g(x)) : x \in \mathcal{X}\}$.

[62]Assume for simplicity, $F_{g(x)}$ has continuous distribution, otherwise, one has to randomize .

[63]We believe that except for shift baseline trade-off functions $f = T(X, X+\mu)$ for some $\mu > 0$ mechanisms of the noise adding kind would not be tight in general such as in the Poisson case 13.

[64]Although the Gaussian distribution is infinitely divisible, but discrete Gaussian is not Bose et al. (2002).

Now, observe that the natural limits obtained here deviate from Gaussian trade-off curves Dong et al. (2022) (see figure 2), and the sequential queries for coordinates of a high-dimensional vector $\beta \in \{0,1\}^n$ can be asked without its connection to a graph-based dataset. For example, the coordinates of β might correspond to the characteristics of a single person on different topics.

We end by referring to Karwa & Slavković (2016), which discusses the $(\varepsilon, \delta = 0)$ differentially private framework on an exponential random graph model G and achieves this by adding discrete Laplace noise on the degree sequence $d = (d_1, \cdots, d_n)$ of the graph G. We believe that it would be interesting to work under the Poisson differential privacy framework 13 and analyze how the results vary when we apply our Poisson optimal mechanism M instead.

Now, we prove a Poisson Thinning and superposition result required in the proof of Theorem 13.

Lemma 21 (Montonicity properties of Poisson trade-off functions under translation and scaling)**.**

$$T(P(\lambda_1), P(\lambda_2)) \geq T(P(c\lambda_1), P(c\lambda_2)) \text{ if } 1 \leq c < \infty, \lambda_1, \lambda_2 > 0, \quad (96)$$
$$T(P(\lambda_1), P(\lambda_2)) \leq T(P(c\lambda_1), P(c\lambda_2)) \text{ if } 0 < c \leq 1, \lambda_1, \lambda_2 > 0, \quad (97)$$
$$T(P(\lambda_1), P(\lambda_2)) \leq T(P(\lambda_1 + \lambda), P(\lambda_1 + \lambda)) \text{ if } \lambda, \lambda_1, \lambda_2 > 0, \quad (98)$$
$$T(P(\lambda_1), P(\lambda_2)) \geq T(P(\lambda_1 - \lambda), P(\lambda_1 - \lambda)) \text{ if } \lambda, \lambda_1 - \lambda, \lambda_2 - \lambda > 0. \quad (99)$$

Proof. From Blackwell's theorem Dong et al. (2022)[Theorem 2] $T(P, Q) \leq T(KP, KQ)$ for any pairs of probability measures P, Q and Markov kernel K. So, its enough to establish that there exists Markov kernels K_c, K_λ defined from $(\Omega, \mathcal{F}) = (\mathbb{Z}_{\geq 0}, 2^{\mathbb{Z}_{\geq 0}})$ to itself so that for $0 < c < 1$ and $\lambda, \lambda_1, \lambda_2 > 0$ we have (by symmetry $(c \to c^{-1}, \lambda \to -\lambda)$, its enough to establish (97) and (98))

$$K_c P(\lambda_1) = P(c\lambda_1) \text{ and } K_c P(\lambda_2) = P(c\lambda_2) \quad (100)$$
$$K_\lambda P(\lambda_1) = P(\lambda_1 + \lambda) \text{ and } K_\lambda P(\lambda_2) = P(\lambda_2 + \lambda) \quad (101)$$

The construction of K_λ captures the intuition that we are adding the same independent Poisson noise of mean λ to both sides. More precisely, define $K_\lambda : \mathbb{Z}_{\geq 0} \times 2^{\mathbb{Z}_{\geq 0}} \to [0,1]$ as[65]

$$K_\lambda(n, A) := \mathbb{P}(n + P \in A), \text{ for } n \in \mathbb{Z}_{\geq 0}, \ A \subseteq \mathbb{Z}_{\geq 0}, P \sim P(\lambda). \quad (102)$$

Then by the definition of Markov kernel $K : (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F}')$ we have from Kallenberg (2021)

$$(KP)(A) := \int_\Omega K(x, A) dP(dx), \text{ for } A \in \mathcal{F} \text{ and a probability measure } P \text{ on } (\Omega, \mathcal{F}). \quad (103)$$

So, for all $k \in \mathbb{Z}_{\geq 0}$ with $P(\lambda)(\{k\}) = e^{-\lambda} \frac{\lambda^k}{k!}$ supported on $\mathbb{Z}_{\geq 0}$ we have the following

$$(K_\lambda P(\lambda_1))(\{k\}) = \sum_{0 \leq i \leq k} e^{-\lambda} \frac{\lambda^{k-i}}{(k-i)!} e^{-\lambda_1} \frac{\lambda_1^i}{i!} = e^{-(\lambda + \lambda_1)} \frac{(\lambda_1 + \lambda)^k}{k!} = P(\lambda + \lambda_1)(\{k\}) \quad (104)$$

Therefore, $K_\lambda P(\lambda_1) = P(\lambda + \lambda_1)$, and by a similar computation, we have $K_\lambda P(\lambda_2) \equiv P(\lambda + \lambda_2)$.

The construction of K_c is based on Poisson thinning. For $c \in (0, 1]$, define $K_c : \mathbb{Z}_{\geq 0} \times 2^{\mathbb{Z}_{\geq 0}} \to [0,1]$

$$K_c(n, \{k\}) = \mathbb{P}\big(\text{Bin}(n, c) := k\big) = \binom{n}{k} c^k (1-c)^{n-k} \mathbf{1}(0 \leq k \leq n) \text{ for } n \in \mathbb{Z}_{\geq 0}.$$

Then K_c is a Markov kernel on $(\mathbb{Z}_{\geq 0}, 2^{\mathbb{Z}_{\geq 0}})$, since $\sum_{k \in \mathbb{Z}_{\geq 0}} K_c(n, \{k\}) = 1$ for all $n \in \mathbb{Z}_{\geq 0}$, and for every $A \subset \mathbb{Z}_{\geq 0}$, the map $\mathbb{Z}_{\geq 0} \ni n \to K_c(n, A) \in [0, 1]$ is $2^{\mathbb{Z}_{\geq 0}}$ to $\mathcal{B}_{[0,1]}$ measurable. Now,

$$L = (K_1 P(\lambda_1))(\{k\}) = \sum_{n=k}^{\infty} \binom{n}{k} c^k (1-c)^{n-k} e^{-\lambda_1} \frac{\lambda_1^n}{n!}. \quad (105)$$

Now, applying the change of variable $n \to k + n$ to the sum above, we have that

$$L = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!k!} c^k (1-c)^n e^{-\lambda_1} \frac{\lambda_1^{n+k}}{(n+k)!} = e^{-\lambda_1} \frac{(c\lambda_1)^k}{k!} \sum_{n=0}^{\infty} \frac{\big((1-c)\lambda_1\big)^n}{n!} \quad (106)$$

which equals $e^{-c\lambda_1} \frac{(c\lambda_1)^k}{k!} = P(c\lambda_1)(\{k\})$. Since this is true for any $k \in \mathbb{Z}_{\geq 0}$, we have $K_c P(\lambda_1) = P(c\lambda_1)$. A similar computation shows $K_c P(\lambda_2) = P(c\lambda_2)$, thereby establishing (100). $\square$

---

[65]By definition of $K_\lambda(n, A)$, we have $A \to K_\lambda(n, A)$ is a probability measure on $(\mathbb{Z}_{\geq 0}, 2^{\mathbb{Z}_{\geq 0}})$ for every $n \in \mathbb{Z}_{\geq 0}$, and for every $A \subset \mathbb{Z}_{\geq 0}$, the map $\mathbb{Z}_{\geq 0} \ni n \to K_\lambda(n, A) \in [0, 1]$ is $2^{\mathbb{Z}_{\geq 0}}$ to $\mathcal{B}_{[0,1]}$ measurable .

5 NEYMAN-PEARSON LEMMA UNDER A COARSER σ-ALGEBRA

The Neyman-Pearson optimizer φ^* of the trade-off function $T(P,Q)$ for probability measures (P,Q) on a measurable space $(\Omega, \mathcal{F})$ involves the likelihood ratio (Radon-Nikodym derivative) random variable $F = \frac{dQ}{dP} : (\Omega, \mathcal{F}) \to [0, \infty]$ (2), but we pose a natural question: what happens when we shrink the original σ-algebra $\mathcal{F}$ to a smaller one $\mathcal{G}$[66], so that the Radon-Nikodym derivative $H = \frac{dQ}{dP}$ is not a $\mathcal{G}$ measurable random variable; therefore, φ^* cannot be the Neyman-Pearson optimizer for (P,Q) on $(\Omega, \mathcal{G})$. In this section, we answer this question in full generality. Our motivation comes from a Bayesian perspective of differential privacy as proposed in Strack & Yang (2024), whose connections with the f-DP framework will also be an interesting future direction.

Theorem 22. *Consider a measurable space $(\Omega, \mathcal{F})$, and let P, Q be probability measures on $(\Omega, \mathcal{F})$[67]. Denote the likelihood ratio $F := \frac{dQ}{dP} : (\Omega, \mathcal{F}) \to [0, \infty]$ (defined only P-almost surely). Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra (a coarsening), and $P^{\mathcal{G}}$ and $Q^{\mathcal{G}}$ are the restrictions of P, Q to $\mathcal{G}$. Define the* conditional likelihood ratio $G := \mathbb{E}_P[F \mid \mathcal{G}] : (\Omega, \mathcal{G}) \to [0, \infty]$. *Then,*

- *G is (a $P^{\mathcal{G}}$-almost sure version of) the Radon–Nikodym derivative $\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}}$ on $(\Omega, \mathcal{G})$.*

Now, consider the coarsened trade-off function $T_{\mathcal{G}}(P,Q) \equiv T(P^{\mathcal{G}}, Q^{\mathcal{G}}) : [0,1] \to [0,1]$ defined as

$$T(P^{\mathcal{G}}, Q^{\mathcal{G}})(\alpha) = \inf_{\varphi} \left\{ \mathbb{E}_Q[1 - \varphi] \;\middle|\; \mathbb{E}_P[\varphi] \le \alpha, \; \varphi : (\Omega, \mathcal{G}) \to [0,1] \text{ measurable} \right\}. \tag{107}$$

- *The Neyman-Pearson optimizer $\varphi^*_{\mathcal{G}}$ achieving $T_{\mathcal{G}}(P,Q)$ at level $\alpha \in [0,1]$ is given by (2)*

$$\varphi^*(w) := \mathbf{1}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} > \tau^* \right) + \lambda^* \mathbf{1}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} = \tau^* \right), \quad \frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} := \frac{dQ^{\mathcal{G}}}{d\mu^G} \Big/ \frac{dP^{\mathcal{G}}}{d\mu^{\mathcal{G}}} : (\Omega, \mathcal{G}) \to [0, \infty] \tag{108}$$

for $2\mu^{\mathcal{G}} = P^{\mathcal{G}} + Q^{\mathcal{G}}$, the ratios of Radon–Nikodym derivatives, and τ^*, λ^* are determined by choosing $\tau^* \in [0, \infty]$ as the unique number

such that $P^{\mathcal{G}}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} \ge \tau^* \right) \ge \alpha \ge P^{\mathcal{G}}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} > \tau^* \right)$,

$$\lambda^* := \frac{\alpha - P^{\mathcal{G}}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} > \tau^* \right)}{P^{\mathcal{G}}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} \ge \tau^* \right) - P^{\mathcal{G}}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} > \tau^* \right)} \mathbf{1}\left(\alpha - P^{\mathcal{G}}\left(\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}} > \tau^* \right) > 0 \right) \text{ so } \mathbb{E}_P[\varphi^*_{\mathcal{G}}] = \alpha.$$

- *Moreover, the trade-off function satisfies data-processing inequality (Blackwell ordering)*

$$T(P^{\mathcal{G}}, Q^{\mathcal{G}})(\alpha) \ge T(P,Q)(\alpha) \text{ for all } \alpha \in [0,1], \text{ with equality for all } \alpha \text{ if and only if} \tag{109}$$

F itself is $\mathcal{G}$-measurable P almost surely or equivalently, $\mathbb{E}_P[F|\mathcal{G}] = F$, P almost surely[68].

Proof. First, we show that G is the Radon-Nikodym derivative $\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}}$. Now, observe that for any measurable function $(\Omega, \mathcal{G}) \xrightarrow{\varphi} [0,1]$, we have $\mathbb{E}_Q[\varphi] = \mathbb{E}_P[F\varphi] = \mathbb{E}_P[\mathbb{E}_P(F \mid \mathcal{G})\varphi] = \mathbb{E}_P[G\varphi]$.

Second, the constrained minimization of $\mathbb{E}_Q[1 - \varphi]$ over $\mathcal{G}$-measurable φ with $\mathbb{E}_P[\varphi] \le \alpha$ reduces to the standard Neyman-Pearson problem for the reduced experiment $(\Omega, \mathcal{G}, P^{\mathcal{G}}, Q^{\mathcal{G}})$ with likelihood ratio $\frac{dQ^{\mathcal{G}}}{dP^{\mathcal{G}}}$. The Neyman-Pearson lemma (108) gives the threshold form of $\varphi^*_{\mathcal{G}}$.

---

[66]Shrinking the σ-algebra to a smaller σ algebra captures the intuition of blurring the information in subsets of the sample space Ω by reducing the number of allowed subsets from $\mathcal{F}$ to $\mathcal{G}$. A complementary way would be to say that coarsening information is the same as testing with fewer allowed observables φ, which are available at our disposal. See Revuz & Yor (1999) for more details on the intuition behind *filtration* of σ-algebras.

[67]We do not need this assumption of absolute continuity of $Q \ll P$, and the result holds more generally when we extend our definition of the Radon-Nikodym derivative $\frac{dQ}{dP} := \frac{dQ}{d\mu} / \frac{dP}{d\mu}$ for a dominating $2\mu = P + Q$.

[68]A complementary way is to say that the information $\mathcal{G}$ is sufficient for the binary experiment $(\Omega, \mathcal{F}, P, Q)$.

Finally, the inequality $T(P^{\mathcal{G}}, Q^{\mathcal{G}}) \geq T(P, Q)$ follows immediately because $\varphi : (\Omega, \mathcal{G}) \to [0, 1]$ is measurable implies $\varphi : (\Omega, \mathcal{F}) \to [0, 1]$ is measurable. Moreover, if $\mathbb{E}_P[F|\mathcal{G}] = F$, then $\varphi_{\mathcal{G}}^* = \varphi^*$ in (108) and hence equality holds for all α holds. Conversely, equality holds at all $\alpha \in [0, 1]$, by the uniqueness of Neyman-Pearson optimizer, $\varphi_{\mathcal{G}}^* = \varphi^*$ at all α, and therefore $F = \frac{dQ}{dP}$ itself is $\mathcal{G}$ measurable[69]. So, $G = \mathbb{E}_P[F|\mathcal{G}] = F, P$ almost surely. $\qquad\square$

5.1 A GAUSSIAN SHIFT EXAMPLE ON A SUB-σ ALGEBRA

Some consequences in practice. Now, we describe the theory developed above with a Gaussian shift example of how privacy (trade-off function) degrades upon coarsening the σ algebra[70].

Consider a Gaussian shift family $\{P_\theta(\cdot) = P(\cdot - \theta) : \theta \in \mathbb{R}\}$ with $P = P_0 = N(0, 1)$ and for $\mu > 0$ let $Q = P_\mu = N(\mu, 1)$ on $(\Omega, \mathcal{F}) = (\mathbb{R}, \mathcal{B}_\mathbb{R})$. Now, coarse to the σ-algebra to define $\mathcal{G}$ as[71]

$$\mathcal{G} = \sigma\big([n, n+1) : n \in \mathbb{Z}\big) = \{\bigcup_{n \in A} [n, n+1) : A \subset \mathbb{Z}\}, \tag{110}$$

So this σ-algebra is naturally isomorphic to the σ-algebra $2^\mathbb{Z}$ (power set) under the isomorphism $A \to \bigcup_{n \in A}[n, n+1)$, and the test functions φ for discriminating between P and Q on $(\Omega, \mathcal{G})$ only allowed to depend on the bin index $N = \lfloor X \rfloor \in \mathbb{Z}$ of the observed value[72].

Proposition 2. *Let* $P_0 = N(0, 1), P_\mu = N(\mu, 1)$ *for* $\mu > 0$ *on* $(\Omega, \mathcal{F}) \equiv (\mathbb{R}, \mathcal{B}_\mathbb{R})$ *with a sub-σ algebra* $\mathcal{G} := \sigma([n, n+1) : n \in \mathbb{Z})$. *Then,* $T(P_0, P_\mu)(\alpha) = \Phi(\Phi^{-1}(1 - \alpha) - \mu) \forall \alpha \in [0, 1]$, *and*

$$T_{\mathcal{G}}(P_0, P_\mu)(\alpha_k) = \Phi(\Phi^{-1}(1 - \alpha_k) - \mu) \text{ for all } \alpha_k = 1 - \Phi(k) \text{ for all } k \in \mathbb{Z}, \text{ and} \tag{111}$$

$$T_{\mathcal{G}}(P_0, P_\mu)(\alpha) = T_{\mathcal{G}}(P_0, P_\mu)(\alpha_{k+1}) - \lambda^* q_k, \text{ for } \alpha \in [\alpha_{k+1}, \alpha_k] \text{ for all } k \in \mathbb{Z} \tag{112}$$

$$and \ \lambda^* = \frac{\alpha - \alpha_{k+1}}{\alpha_k - \alpha_{k+1}}, \text{ with } q_k = \Phi(k + 1 - \mu) - \Phi(k - \mu) \tag{113}$$

Proof. The proof follows from 22. First, observe that the laws $\mathbb{P}^{\mathcal{G}}$ and $\mathbb{Q}^{\mathcal{G}}$ on $(\Omega, \mathcal{G}) \equiv (\mathbb{Z}, 2^\mathbb{Z})$ are

$$p_n := P(N = n) = \Phi(n+1) - \Phi(n), \qquad q_n := Q(N = n) = \Phi(n+1 - \mu) - \Phi(n - \mu), \tag{114}$$

where Φ is the standard normal CDF. For $\mu > 0$ the likelihood ratios F_n are increasing in $n \in \mathbb{Z}$[73].

$$F_n := \frac{q_n}{p_n} = \frac{\int_n^{n+1} \exp(-\frac{(x-\mu)^2}{2}) dx}{\int_n^{n+1} \exp(-\frac{(x)^2}{2}) dx} = \frac{\int_n^{n+1} F_\mu(x) dP_0(x)) dx}{\int_n^{n+1} dP_0(x)} = \mathbb{E}_P[F_\mu|\mathcal{G}](x) \text{ for } x \in [n, n+1), \tag{115}$$

where we denote $F_\mu(x) = \frac{dP_\mu}{dP_0}(x) = \frac{\exp(-\frac{(x-\mu)^2}{2})}{\exp(-\frac{x^2}{2})}$ as the Radon-Nikodym derivative, and a $\mathcal{G}$ measurability requirement of $G = \mathbb{E}_P[F_\mu|\mathcal{G}]$ reveals that $\mathbb{E}_P[F_\mu|\mathcal{G}]$ is constant on each of those partitions $E_n = [n, n+1)$ generating the σ-algebra $\mathcal{G}$ Durrett (2019). Moreover, from the integrability requirement $\mathbb{E}_P[F_\mu \varphi] = \mathbb{E}_P[G\varphi]$ for bounded $\mathcal{G}$ measurable functions φ that constant value is the weighted average as computed above (take $\varphi \equiv 1$). Now, to show that F_n is increasing in n, observe that for any $\mu > 0$, $\frac{dP_\mu}{dP_0}(x) = \exp\left(\mu x - \frac{\mu^2}{2}\right)$ is monotone in x whenever $\mu > 0$[74]. As a consequence, we have $F_\mu(n) \leq F_n \leq F_\mu(n+1) \leq F_{n+1} \leq F_\mu(n+2)$ for every $n \in \mathbb{Z}$.

---

[69]By the restriction of $\varphi_{\mathcal{G}}^* = \varphi^*$ at all $\alpha \in [0, 1]$, we need $\{F > \tau\}_{\tau \geq 0}$ and $\{F = \tau\}_{\tau \geq 0}$ to be $\mathcal{G}$ measurable, and therefore F itself is G measurable, since it is by definition non-negative.

[70]The results extend verbatim with appropriate substitution for any shift family $\{P_\theta(\cdot) = P(\cdot - \theta) : \theta \in\}$ where P_0 is a symmetric-log concave probability measure on $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ Dong et al. (2022).

[71]σ-algebra generated by a countable partition $\{E_n\}_{n \in \mathbb{Z}}$ of Ω can be written as $\{\{\bigcup_{A \subset \mathbb{Z}} E_n : A \subset \mathbb{Z}\}\}$.

[72]The intuition is that the practitioner does not observe the actual output X, but only see the processed output $\lfloor X \rfloor$ taking integer values, and is supposed to base his discriminating procedure based on this information.

[73]A similar computation would show that F_n would be decreasing in n when the shift $\mu < 0$.

[74]As shown in Dong et al. (2022), this condition of monotonicity (increasing) of $\frac{dP_\mu}{dP_0}(x)$ in x for all (but fixed) $\mu > 0$ is equivalent to the condition that P_0 is log concave. This is the reason this Gaussian result extends verbatim to symmetric-log concave distribution P_0 with the shift family $\{P_\theta(\cdot) = P(\cdot - \theta) : \theta \in \mathbb{R}\}$.

So the Neyman-Pearson optimizer on $(\Omega, \mathcal{G}) \equiv (\mathbb{Z}, 2^{\mathbb{Z}})$ is a threshold in N[75] (along with some randomization as required). For the non-randomized part, consider an threshold $k \in \mathbb{Z}$,

$$\text{type I error} = \alpha_k := P(N \geq k) = \sum_{n \geq k} p_n = 1 - \Phi(k), \text{ and} \tag{116}$$

$$\text{type II error} = T_{\mathcal{G}}(P, Q)(\alpha_k) := Q(N < k) = \sum_{n < k} q_n = \Phi(k - \mu). \tag{117}$$

Thus, the trade-off function $T_{\mathcal{G}}(P^{\mathcal{G}}, Q^{\mathcal{G}})$ at the points with type I errors $\{\alpha_k\}_{k \in \mathbb{Z}} \subset [0, 1]$[76], have the respective minimum value of type II errors as $\{\Phi(k - \mu))\}_{k \in \mathbb{Z}}$ or equivalently as the subset of points $\{\alpha_k, \Phi(\Phi^{-1}(1 - \alpha_k) - \mu)\}_{k \in \mathbb{Z}} \subset \{\alpha, T_{\mathcal{G}}(P, Q)(\alpha)\}_{\alpha \in [0,1]} \subset [0, 1]^2$. Now, to have the entire $T_{\mathcal{G}}(P, Q)$, one linearly interpolates the points above on the curve. More precisely, for any $\alpha \in [\alpha_{k+1}, \alpha_k]$, define $\varphi_{\mathcal{G}}^*(N) = \mathbf{1}(N > k) + \lambda^* \mathbf{1}(N = k)$, where $\lambda^* \in [0, 1]$ is given by

$$\lambda^* = \frac{\alpha - \alpha_{k+1}}{\alpha_k - \alpha_{k+1}}, \text{ and } T_{\mathcal{G}}(P, Q)(\alpha) = T_{\mathcal{G}}(P, Q)(\alpha_{k+1}) - \lambda^* q_k. \tag{118}$$

$\square$

This finishes the description of the computation of the trade-off function $T_{\mathcal{G}}(P, Q)$ for a Gaussian shift family, which is clearly equal to $T(P, Q)$ at the points of type I errors $\{\alpha_k\}_{k \in \mathbb{Z}}$ without requiring randomization, and linear interpolated otherwise. Therefore, it always lies above the original Gaussian trade-off curve $T(P, Q)$. Moreover, the $\mu < 0$, case be dealt in exactly similar way.
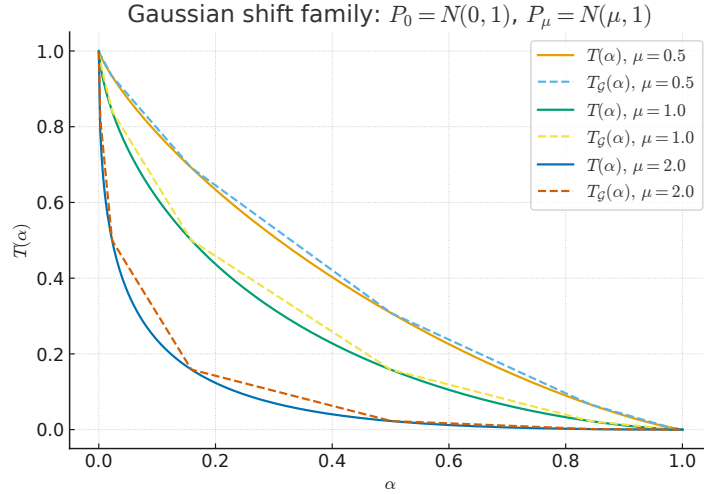

Figure 6: Trade-off function $T(\alpha)$ and its coarsened version $T_{\mathcal{G}}(\alpha)$ for the Gaussian shift family $P_0 = N(0, 1)$, $P_\mu = N(\mu, 1)$.

6 CONCLUSION AND FUTURE WORK

In this paper, we introduce the hypothesis testing-based framework of infinitely divisible privacy 6, encompassing both the Gaussian differential privacy of Dong et al. (2022) and the newly proposed Poisson differential privacy 9, which are not just natural baselines for count statistics but also appear under repeated compositions of (nearly perfect) differentially private outputs. Along the way, we resolve the $s^2 = 2k$ conjecture of Dong et al. (2022) as a consequence of contiguity 3.

In practice, the number of private operations made on a dataset is uncertain. We show that under randomization of the number of operations, the limiting trade-off can even *escape* the universal class

---

[75]To follow the theorem 22, one could equivalently threshold at the value based on F_n as well.

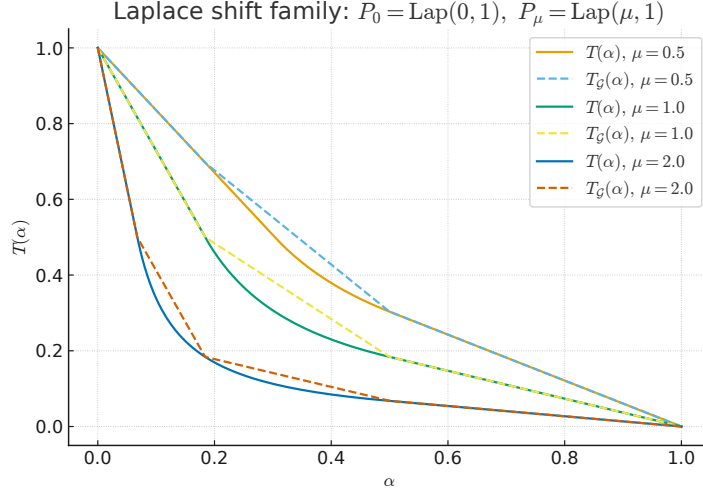[76]These are the values of type I errors which do not require randomization in Neyman-Pearson lemma.

Figure 7: Trade-off function $T(\alpha)$ and its coarsened version $T_{\mathcal{G}}(\alpha)$ for the Laplace shift family $P_0 = \mathrm{Lap}(0,1)$, $P_\mu = \mathrm{Lap}(\mu, 1)$.

of infinitely divisible trade-off functions 64 We illustrate this with a locally asymptotically mixed normal (LAMN) example (66), following the spirit of Le Cam (1986).

Then, we construct a mechanism in Theorem 13 that *optimally* attains (asymmetric) Poisson differential privacy for real-valued statistics, including count statistics, thus going beyond standard additive-noise mechanisms. Afterwards, we demonstrate that Poisson baselines naturally appear for graph-based datasets, where released statistics are inherently discretized 20.

Finally, inspired by a Bayesian perspective on DP (Strack & Yang, 2024), we quantify exactly how trade-off functions degrade under *information coarsening* and derive a coarsened Neyman-Pearson lemma 22 with equality characterizations via conditional likelihood ratios 22. We demonstrate this by explicitly working out the Gaussian case 2 that extends to symmetric-log concave family.

Several interesting directions follow, and many have been detailed in the remarks within the text. First, our algorithmic result suggests the construction of private mechanisms that are optimal with respect to a given natural, infinitely divisible trade-off curve. Second, it would be extremely interesting to illustrate the tightness of the proposed Poisson differential privacy framework on graph-based problems of modern interest Karwa & Slavković (2016). Third, it would be interesting to analyze how privacy degrades when one is allowed to stop adaptively. Finally, it would be fascinating to find an equivalence between the f-differential privacy framework of Dong et al. (2022) based on comparison of trade-off functions and the (Bayesian) framework of privacy-preserving signals Strack & Yang (2024), since coarsening the underlying σ algebra clarifies how information restrictions lead to an increase in indistinguishability.

REFERENCES

Martín Abadi, Andy Chu, Ian Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 23rd ACM Conference on Computer and Communications Security (CCS)*, pp. 308–318. ACM, 2016. doi: 10.1145/2976749.2978318.

John M. Abowd. The u.s. census bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD)*, pp. 2867–2867. ACM, 2018. doi: 10.1145/3219819.3226070.

Naman Agarwal, Peter Kairouz, and Ziyu Liu. The skellam mechanism for differentially private federated learning. In *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)*, pp. 5052–5064, 2021.

Apple Inc. Learning with privacy at scale, 2017. URL `https://machinelearning.apple.com/research/learning-with-privacy-at-scale`. Accessed 2025-10-24.

Jordan Awan and Salil Vadhan. Canonical noise distributions and private hypothesis tests. *The Annals of Statistics*, 51(2):547–572, 2023. doi: 10.1214/23-AOS2259.

Jinho Baik, Gérard Ben Arous, and Sandrine Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005. doi: 10.1214/009117905000000233.

Arup Bose, Anirban DasGupta, and Herman Rubin. A contemporary review and bibliography of infinitely divisible distributions and processes. *Sankhyā: The Indian Journal of Statistics, Series A*, 64(3):763–819, 2002. ISSN 0972-7671.

Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of cryptography conference*, pp. 635–658. Springer, 2016.

Clément L. Canonne, Gautam Kamath, and Thomas Steinke. The discrete gaussian for differential privacy. In *Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, pp. 2576–2587, 2020.

Erhan Çınlar. *Probability and Stochastics*, volume 261 of *Graduate Texts in Mathematics*. Springer, New York, 2011. ISBN 978-0-387-87858-4. doi: 10.1007/978-0-387-87859-1.

Rachel Cummings and collaborators. Advancing differential privacy: Where we are now and future directions for real-world deployment. *Harvard Data Science Review*, 2024. doi: 10.1162/99608f92.d3197524. URL `https://doi.org/10.1162/99608f92.d3197524`.

Víctor H. de la Peña and Evarist Giné. *Decoupling: From Dependence to Independence*. Probability and Its Applications. Springer, New York, 1999. ISBN 0-387-98616-2. doi: 10.1007/978-1-4612-0537-1. URL `https://doi.org/10.1007/978-1-4612-0537-1`.

Victor H. de la Peña, Tze Leung Lai, and Qi-Man Shao. *Self-Normalized Processes: Limit Theory and Statistical Applications*. Probability and Its Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, 1 edition, 2009. ISBN 978-3-540-85635-1. doi: 10.1007/978-3-540-85636-8.

Jinshuo Dong, Aaron Roth, and Weijie J. Su. Gaussian differential privacy. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 84(1):3–37, 2022. doi: 10.1111/rssb.12401.

David L. Donoho and Richard C. Liu. Geometrizing rates of convergence, iii. *Annals of Statistics*, 19(2):668–701, 1991. doi: 10.1214/aos/1176348115.

Joseph L. Doob. *Stochastic Processes*. Wiley Publications in Statistics. John Wiley & Sons, New York, 1953. ISBN 978-0471218135.

John C. Duchi and Feng Ruan. The right complexity measure in locally private estimation: It is not the fisher information. *Annals of Statistics*, 52(1):1–51, 2024. doi: 10.1214/22-AOS2227.

Rick Durrett. *Probability: Theory and Examples*, volume 49 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 5 edition, 2019. ISBN 978-1-108-47368-2. doi: 10.1017/9781108591034.

Cynthia Dwork and Aaron Roth. *The Algorithmic Foundations of Differential Privacy*, volume 9 of *Foundations and Trends in Theoretical Computer Science*. Now Publishers Inc., Hanover, MA, USA, 2014. ISBN 978-1-60198-818-8. doi: 10.1561/0400000042. URL https://doi.org/10.1561/0400000042.

Cynthia Dwork and Guy N. Rothblum. Concentrated differential privacy, 2016. Preprint.

Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Advances in Cryptology – EUROCRYPT 2006*, volume 4004 of *Lecture Notes in Computer Science*, pp. 486–503. Springer, 2006a. doi: 10.1007/11761679_29.

Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography (TCC 2006)*, volume 3876 of *Lecture Notes in Computer Science*, pp. 265–284. Springer, 2006b. doi: 10.1007/11681878_14.

Gerald B. Folland. *Real Analysis: Modern Techniques and Their Applications*. John Wiley & Sons, New York, 2 edition, 1999. ISBN 9780471317166.

Qiyang Han, Tiefeng Jiang, and Yandi Shen. Contiguity under high-dimensional gaussianity with applications to covariance testing. *Annals of Applied Probability*, 33(6A):4272–4321, 2023. doi: 10.1214/22-AAP1917.

Arnold Janssen, Hartmut Milbrodt, and Helmut Strasser. *Infinitely Divisible Statistical Experiments*, volume 27 of *Lecture Notes in Statistics*. Springer, New York, NY, 1985. ISBN 978-0-387-96055-5. doi: 10.1007/978-1-4615-7261-9. Reprinted as eBook, Springer, 2012.

P. Jeganathan. On the asymptotic theory of estimation when the limit distribution of the log-likelihood ratios is mixed normal. *Sankhyā: The Indian Journal of Statistics, Series A*, 44(2):173–198, 1982.

Iain M. Johnstone and Alexei Onatski. Testing in high-dimensional spiked models. *The Annals of Statistics*, 48(3):1231–1254, 2020. doi: 10.1214/18-AOS1697.

Peter Kairouz, Sewoong Oh, and Pramod Viswanath. The composition theorem for differential privacy. In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, volume 37 of *Proceedings of Machine Learning Research*, pp. 1376–1385. PMLR, 2015.

Olav Kallenberg. *Foundations of Modern Probability*, volume 99 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 3 edition, feb 2021. ISBN 978-3-030-61870-4. doi: 10.1007/978-3-030-61871-1. URL https://link.springer.com/book/10.1007/978-3-030-61871-1.

Noureddine El Karoui. Tracy–widom limit for the largest eigenvalue of a large class of complex sample covariance matrices. *Ann. Probab.*, 35(2):663–714, 2007. doi: 10.1214/009117906000000917.

Vishesh Karwa and Aleksandra B. Slavković. Inference using noisy degrees: Differentially private β-model and synthetic graphs. *Annals of Statistics*, 44(1):87–112, 2016. doi: 10.1214/15-AOS1358.

Achim Klenke. *Probability Theory: A Comprehensive Course*. Universitext. Springer, London, 1 edition, 2008. ISBN 978-1-84800-047-6. doi: 10.1007/978-1-84800-048-3.

Lucien Le Cam. *Asymptotic Methods in Statistical Decision Theory*. Springer Series in Statistics. Springer-Verlag, New York, 1986. ISBN 978-0-387-96328-0. doi: 10.1007/978-1-4612-4946-7. MR0856411.

Lucien Le Cam and Grace Lo Yang. *Asymptotics in Statistics: Some Basic Concepts*. Springer Series in Statistics. Springer-Verlag, New York, 2 edition, 2000. ISBN 978-0-387-95036-5. doi: 10.1007/978-1-4612-1166-2.

Erich Leo Lehmann, Joseph P Romano, et al. *Testing statistical hypotheses*, volume 3. Springer, 1986.

Xiang Li, Feng Ruan, Huiyuan Wang, Qi Long, and Weijie J. Su. A statistical framework of watermarks for large language models: Pivot, detection efficiency and optimal rules. *Annals of Statistics*, 53(1):322–351, 2025. doi: 10.1214/24-AOS2468. URL https://doi.org/10.1214/24-AOS2468.

Ilya Mironov. Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pp. 263–275. IEEE, 2017.

Arvind Narayanan and Vitaly Shmatikov. Robust de-anonymization of large sparse datasets. In *2008 IEEE Symposium on Security and Privacy (SP)*, pp. 111–125. IEEE, 2008. doi: 10.1109/SP.2008.33.

Alexandru Nica and Roland Speicher. *Lectures on the Combinatorics of Free Probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006. ISBN 9780521833479. doi: 10.1017/CBO9780511735127.

Alexei Onatski, Marcelo J. Moreira, and Marc Hallin. Asymptotic power of sphericity tests for high-dimensional data. *The Annals of Statistics*, 41(3):1204–1231, 2013. doi: 10.1214/13-AOS1100.

Aaradhya Pandey, Arnab Auddy, Haolin Zou, Arian Maleki, and Sanjeev Kulkarni. Gaussian certified unlearning in high dimensions: A hypothesis testing approach. *Preprint*, 10 2025. doi: 10.48550/arXiv.2510.13094. URL https://arxiv.org/abs/2510.13094. arXiv:2510.13094 [stat.ML].

David Pollard. Some thoughts on le Cam's statistical decision theory. *arXiv preprint arXiv:1107.3811*, 2011. doi: 10.48550/arXiv.1107.3811. URL https://arxiv.org/abs/1107.3811.

Yury Polyanskiy and Yihong Wu. *Information Theory: From Coding to Learning*. Cambridge University Press, Cambridge, feb 2025. ISBN 978-1108832908. doi: 10.1017/9781108966351. URL https://doi.org/10.1017/9781108966351.

Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 3 edition, 1999. ISBN 3-540-64325-7. doi: 10.1007/978-3-662-06400-9.

R. Tyrrell Rockafellar. *Convex Analysis*. Princeton Landmarks in Mathematics and Physics. Princeton University Press, Princeton, NJ, 1970. ISBN 9780691015866.

Angelika Rohde and Lukas Steinberger. Geometrizing rates of convergence under local differential privacy constraints. *Annals of Statistics*, 48(5):2646–2670, 2020. doi: 10.1214/19-AOS1901.

David O. Siegmund. *Sequential Analysis: Tests and Confidence Intervals*. Springer Series in Statistics. Springer New York, New York, NY, 1985. ISBN 978-0-387-96134-7. doi: 10.1007/978-1-4757-1862-1.

Dmitrii S. Silvestrov. *Limit Theorems for Randomly Stopped Stochastic Processes*. Probability and Its Applications. Springer, London, 2004. ISBN 978-1-85233-777-3. doi: 10.1007/978-0-85729-390-9.

Lukas Steinberger. Efficiency in local differential privacy. *The Annals of Statistics*, 52(5):2139–2166, 2024. doi: 10.1214/24-AOS2425.

Philipp Strack and Kai Hao Yang. Privacy-preserving signals. *Econometrica*, 92(6):1907–1938, November 2024. doi: 10.3982/ECTA22017.

Erik Torgersen. *Comparison of Statistical Experiments*, volume 36 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1991. ISBN 9780521250306. doi: 10.1017/CBO9780511666353.

A. W. van der Vaart. *Asymptotic Statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. ISBN 9780521496032.

Ramon van Handel. Apc 550: Lecture notes. `https://web.math.princeton.edu/~rvan/APC550.pdf`, 2016. Course notes, accessed 2025-09-24.

Yu-Xiang Wang, Borja Balle, and Shiva Prasad Kasiviswanathan. Subsampled rényi differential privacy and analytical moments accountant. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 89 of *Proceedings of Machine Learning Research*, pp. 1226–1235. PMLR, 2019.

Larry Wasserman and Shuheng Zhou. A statistical framework for differential privacy. *Journal of the American Statistical Association*, 105(489):375–389, 2010. doi: 10.1198/jasa.2009.tm08651.

A SUPPLEMENTARY MATERIAL

We first define a Markov kernel (randomized mechanism) following Kallenberg (2021).

Definition 7 (Markov kernel). *Let $(\mathcal{X}, \mathcal{F}_\mathcal{X})$ and $(\mathcal{Y}, \mathcal{F}_\mathcal{Y})$ be measurable spaces. A* Markov kernel *(also called a* stochastic kernel *or* transition kernel*) from $(\mathcal{X}, \mathcal{F}_\mathcal{X})$ to $(\mathcal{Y}, \mathcal{F}_\mathcal{Y})$ is a map*

$$K : \mathcal{X} \times \mathcal{F}_\mathcal{Y} \to [0, 1], \text{ usually written } K(x, B) \text{ for } x \in \mathcal{X}, B \in \mathcal{F}_\mathcal{Y}, \tag{119}$$

such that for every fixed $x \in X$, the set function $B \mapsto K(x, B)$ is a probability measure on $(\mathcal{Y}, \mathcal{F}_\mathcal{Y})$. Moreover, for every fixed $B \in \mathcal{F}_\mathcal{Y}$, the function $x \mapsto K(x, B)$ is $\mathcal{F}_\mathcal{X}$-measurable.

We state the following Berry Esseen CLT, where **kl** denotes the vector $(\mathrm{kl}(f_1), \ldots, \mathrm{kl}(f_n))$ and $\boldsymbol{\kappa}_2$, $\boldsymbol{\kappa}_3$, $\overline{\boldsymbol{\kappa}}_3$ are defined similarly; in addition, $\|\cdot\|_1$ and $\|\cdot\|_2$ are the ℓ_1 and ℓ_2 norms, respectively.

Our **first** lemma summarizes some of the basic but extremely fruitful properties of the Gaussian trade-off function $f_{G,\varepsilon}(x) = T(N(0,1), N(\varepsilon,1))(x) = \Phi(\Phi^{-1}(1-x) - \varepsilon)$ for all $x \in [0,1]$, and any $\varepsilon \geq 0$. Among others, it proves that $T\left(N\left(-\frac{\mu^2}{2}, \mu^2\right), N\left(\frac{\mu^2}{2}, \mu^2\right)\right) = f_{G,|\mu|}$ for any $\mu \in \mathbb{R}$.

Lemma 23. *(Dong et al. (2022)) The Gaussian trade off functions satisfy the following properties:*

- *Monotonicity: For any pair $\varepsilon_1, \varepsilon_2 \geq 0$, $\varepsilon_1 \leq \varepsilon_2$ if and only if*

$$f_{G,\varepsilon_1}(x) \geq f_{G,\varepsilon_2}(x) \text{ for all } x \in [0,1]. \tag{120}$$

- *Closure under suprema: For any collection of $(\varepsilon_i)_{i \in I} \subset \mathbb{R}$ with index set I*

$$\inf_{i \in I} f_{G,\varepsilon_i}(x) = f_{G,\sup_{i \in I} \varepsilon_i}(x) \text{ for all } x \in [0,1] \tag{121}$$

- *Symmetry: For any $\mu_1, \nu_2 \in \mathbb{R}^p$ and $\sigma > 0$*

$$T(\nu_2 + \sigma N(0, \mathbb{I}_p), \mu_1 + \sigma N(0, \mathbb{I}_p)) = T(\mu_1 + \sigma N(0, I_p), \nu_2 + \sigma N(0, I_p)) \tag{122}$$

- *Dimension freeness: For any $\mu_1, \nu_2 \in \mathbb{R}^p$ and $\Sigma \succ 0$ let $\varepsilon^2 := \langle (\mu_1 - \nu_2), \Sigma^{-1}(\mu_1 - \nu_2) \rangle$. Then*

$$T(\mu_1 + \sqrt{\Sigma} N(0, \mathbb{I}_p), \nu_2 + \sqrt{\Sigma} N(0, \mathbb{I}_p)) \equiv T(N(0,1), N(\varepsilon,1)), \tag{123}$$

Intuition and importance: The proof of this lemma is immediate from the explicit description of $f_{G,\varepsilon}(x) = \Phi(\Phi^{-1}(1-x) - \varepsilon)$ where $\Phi(\cdot)$ is the CDF of a standard Gaussian variable $\Phi(t) = \mathbb{P}[N(0,1) \leq t]$ for all $t \in \mathbb{R}$. It also requires applying the Neyman-Pearson lemma or the likelihood ratio test Polyanskiy & Wu (2025). But, the conclusions that they imply are extremely powerful.

Monotonicity: The *monotonicity* condition (120) reduces an apriori difficult functional comparison between two functions f and g at uncountably many points to a comparison of just one parameter ε.

Closure under suprema: The closure under suprema property (121) essentially says two very important things. First, it makes it easy to identify what the suprema of an apriori arbitrary collection of functions $\{f_i\}_i$ is (in fact explicitly). Second, the limiting object is a function of the same kind: it is again a Gaussian trade-off function with a different choice of parameter ε.

Symmetry: The symmetry property (122) (requires Neyman-Pearson lemma) is very interesting because the definition of a trade-off function $T(P, Q)$ 1 is asymmetric in general, between its first and second arguments. But, for a pair of shifted isotropic Gaussians, they match because of the spherical symmetry (orthogonal invariance) of the standard Gaussian density in any dimension.

Dimension freeness: The dimension freeness property (123) (requires Neyman-Pearson lemma) makes the case for Gaussian certifiability in high dimensions stronger than any other notion of certifiability. This is because, almost all the results of classical statistics that are true in low dimensions, fail to hold in high dimensions, because many of the quantities involved in controlling the errors are dimension dependent and blows up when $p \uparrow \infty$. It is often the case that finding a dimension-free quantity or even an inequality that *tensorizes* [77] help us resolve high dimensional issues.

Our **second** lemma says that the some aspects of the Gaussian trade-off functions generalize.

Lemma 24. *Dong et al. (2022)[Lemma A.2, Proposition A.3] Consider a symmetric random variable $X \overset{d}{=} -X$ with CDF F having a log-concave Lebesgue density. Consider its baseline trade-off function $f_{X,\varepsilon}(x) = T(X, \varepsilon + X)(x)$ for all $x \in [0, 1]$.*

1. *Symmetry: For any $t_1, t_2 \in \mathbb{R}$ and $\delta = |t_2 - t_1|$ the following holds for all $x \in [0, 1]$.*

$$T(t_1 + X, t_2 + X)(x) = F\left(F^{-1}(1 - x) - \delta\right) = T(t_2 + X, t_1 + X)(x). \qquad (124)$$

$$\text{So, } T(t_1 + X, t_2 + X)(x) = T(X, \delta + X)(x) = T(\delta + X, X)(x) \qquad (125)$$

2. *Monotonicity: For $\varepsilon_1, \varepsilon_2 \geq 0$. Then $\varepsilon_1 \leq \varepsilon_2$ if and only if*

$$f_{X,\varepsilon_1}(x) \geq f_{X,\varepsilon_2}(x) \text{ for all } x \in [0, 1]. \qquad (126)$$

Intuition and importance: The proof of this lemma is along the same lines of the proof given in Dong et al. (2022)[Lemma A.2, Proposition A.3]. Using log-concavity of the density of X and symmetry $X \overset{d}{=} -X$, there is an explicit description of $f_{X,\varepsilon}(x) = F(F^{-1}(1 - x) - \varepsilon)$ where $F(\cdot)$ is the CDF of the random variable X, $F(t) = \mathbb{P}[X \leq t]$ for all $t \in \mathbb{R}$, and $\varepsilon \geq 0$.

---

[77] We are referring dimension free Poincare and Logarithmic Sobolev inequalities of high-dimensional statistics that are extremely important in obtaining concentration bounds in high dimensions van Handel (2016).