Journal of Machine Learning Research 25 (2024) 1-32 Submitted 12/23; Revised 11/24; Published 11/24

Optimizing Noise for f-Differential Privacy via Anti-Concentration and Stochastic Dominance

Jordan Awan JAWAN@PURDUE.EDU
Department of Statistics
Purdue University
West Lafayette, IN 47907, USA

Aishwarya Ramasethu AISHWARYA.RAMASETHU@GMAIL.COM
Department of Statistics
Purdue University
West Lafayette, IN 47907, USA

Editor: Pierre Alquier

Abstract

In this paper, we establish anti-concentration inequalities for additive noise mechanisms which achieve f-differential privacy (f-DP), a notion of privacy phrased in terms of a tradeoff function f which limits the ability of an adversary to determine which individuals were in the database. We show that canonical noise distributions (CNDs), proposed by Awan and Vadhan (2023), match the anti-concentration bounds at half-integer values, indicating that their tail behavior is near-optimal. We also show that all CNDs are sub-exponential, regardless of the f-DP guarantee. In the case of log-concave CNDs, we show that they are the stochastically smallest noise compared to any other noise distributions with the same strong privacy guarantee. In terms of integer-valued noise, we propose a new notion of discrete CND and prove that a discrete CND always exists, can be constructed by rounding a continuous CND, and that the discrete CND is unique when designed for a statistic with sensitivity 1. We further show that the discrete CND at sensitivity 1 is stochastically smallest compared to other integer-valued noises. Our theoretical results shed light on the different types of privacy guarantees possible in the f-DP framework and can be incorporated in more complex mechanisms to optimize performance.

Keywords: canonical noise distribution, total variation, discrete noise, log-concave distribution, sub-exponential distribution

1. Introduction

Differential privacy (DP), introduced by Dwork et al. (2006b), is the state-of-the-art framework for formal privacy protection. DP methods require the introduction of noise into data analyses, which obscures the contribution of any particular individual. Since its inception, DP has grown in popularity and is now employed by leading tech giants like Google (Erlingsson et al., 2014) and Apple (Tang et al., 2017), as well as by the US Census Bureau (Abowd, 2018).

While there are several variants of DP, they all quantify the privacy risk in terms of a similarity measure between the distributions of outputs, when the mechanism is applied to two databases that differ by just one individual's data. The differences between these variants are primarily in the specific similarity measure. Recently, Dong et al. (2022) proposed f-DP, which is rooted in hypothesis testing. The f parameter in f-DP is a function which offers a more expressive quantification of the privacy risk compared to other notions of DP. The f-DP framework has the benefit of lossless conversions to (ϵ, δ)-DP, as well as divergence-based notions of DP (Bun and Steinke, 2016; Mironov, 2017).

The most basic technique for achieving DP is independent noise addition to a statistic of interest. Popular examples of such noise distributions are Laplace and Gaussian. In this paper, we investigate the limitations on what types of noise can be used to satisfy f-DP, and optimize the noise distributions in various settings.

Our Contributions: To understand the limits on the magnitude of noise required to satisfy f-DP, we develop an anti-concentration inequality, which gives an upper limit on the concentration of the noise distribution about its center. This upper bound is determined by the total variation distance between the noise distribution and a shifted version of itself, which can be bounded in terms of the f-DP guarantee.

We show that canonical noise distributions (CNDs), proposed by Awan and Vadhan (2023) match our anti-concentration bound, which implies that their tail behavior is near-optimal. We also show that all CNDs are sub-exponential, indicating that it is never necessary to use heavier-tailed distributions to achieve f-DP. In the special case of log-concave CNDs, a subclass previously investigated by Awan and Dong (2022), we show that the log-concave CND is stochastically smaller than any other additive noise with the same privacy guarantee.

Lastly, we propose a new concept of discrete CND, which generalizes CNDs to integer-valued noises. Unlike continuous CNDs, which are solely defined in terms of a tradeoff function f, a discrete CND is also defined for a specific sensitivity. We show that a discrete CND always exists for any f and any sensitivity, and that one can be constructed by rounding a continuous CND to integer values. In the case that the sensitivity is 1, we prove that there is a unique discrete CND for each tradeoff function and that this noise mechanism is stochastically smaller than any other integer-centered discrete noise satisfying f-DP.

Related Work: Various ways to optimize noise for differential privacy have been previously explored. Ghosh et al. (2012) show that for a wide variety of loss functions, the geometric mechanism (discrete Laplace) minimizes the expected Bayesian loss under ϵ-DP, regardless of what prior is used. Gupte and Sundararajan (2010) prove a similar result, showing that the geometric mechanism optimizes the minimax loss under ϵ-DP. Qin et al. (2022b) consider necessary conditions for general discrete distributions to satisfy either $(\epsilon, 0)$-DP or (ϵ, δ)-DP. In the extended version of their paper, they optimize the Wasserstein distance between the mechanism's input and output distributions, and show that a discrete staircase distribution is optimal (Qin et al., 2022a).

Geng and Viswanath (2015b) focus on the minimization of loss functions in $(\epsilon, 0)$-DP, with particular emphasis on the mean absolute error and mean squared error losses. They established an optimal noise-adding mechanism which they call the staircase mechanism. Geng and Viswanath (2015a) consider optimizing the same loss functions in (ϵ, δ)-DP, and found that in the high privacy regime, the discrete Laplace and uniform noise were nearly optimal for the mean absolute error and mean squared error losses. Soria-Comas and Domingo-Ferrer (2013) construct an optimal data-independent noise-adding mechanism under ϵ-DP, where the optimality is measured by moving the most probability mass towards zero. Their solution results in a staircase distribution, similar to Geng and Viswanath (2015b). In the setting of ϵ-local-DP, a stronger form of differential privacy than the one considered in this paper, Kairouz et al. (2016) showed that for any sub-linear utility function, there exists a staircase mechanism which maximizes this utility.

The literature above is focused on finding the optimal noise under either $(\epsilon, 0)$-DP or (ϵ, δ)-DP. A broader privacy framework is f-Differential Privacy (Dong et al., 2022). Awan and Vadhan (2023) was the first work to build a noise-adding mechanism for arbitrary f-DP, which they called a canonical noise distribution (CND). CNDs were designed to tightly match the f-DP privacy bound and were shown to lead to the most powerful hypothesis tests on binary data (Awan and Vadhan, 2023). They have also seen applications in other DP testing problems (Awan and Wang, 2022; Kazan et al., 2023). Awan and Dong (2022) extended the results on CNDs by establishing the existence and construction of both log-concave CNDs and multivariate CNDs.

Organization: In Section 2, we review the basics of f-differential privacy and define the notion of optimality considered in this paper. In Section 3, we develop our general anti-concentration inequality for additive noise. In Section 4, we consider the case of continuous noise. We review background on CNDs in Section 4.1, show that CNDs are near-optimal in Section 4.2, establish that CNDs are sub-exponential in Section 4.3, and derive optimality results for log-concave CNDs in Section 4.4. In Section 5, we introduce the concept of discrete CNDs. In Section 5.1, we establish existence and construction of discrete CNDs, as well as the uniqueness of the discrete CND at sensitivity 1. In Section 5.2, we prove the optimality of the discrete CND at sensitivity 1. We conclude with discussion in Section 6. Technical details and proofs are postponed to the appendix. Code to reproduce the numerical results of the paper is available at `https://github.com/JordanAwan/OptimizingNoiseForFDP`.

2. Differential Privacy Background

Differential privacy is a framework that ensures that an adversary cannot accurately determine whether an individual's data is present in a database, based on the output of a privacy mechanism. A privacy mechanism M is a randomized algorithm that takes as input a database D and outputs a random variable $M(D)$. We write $M : \mathscr{D} \to \mathscr{Y}$ to indicate that it takes in a database $D \in \mathscr{D}$ and the random variable $M(D)$ takes values in $\mathscr{Y}$, where $\mathscr{D}$ is the space of possible databases. A mechanism M satisfies differential privacy if for two databases D and D', differing in one entry, the distributions of $M(D)$ and $M(D')$ are similar. The concept "differing in one entry" is formalized in terms of an *adjacency metric* $d(D, D') \leq 1$, which has the property that for all $k \in \mathbb{Z}^{\geq 0}$ if $d(D, D') \leq k$ then there exists a sequence $D = D_0, D_1, D_2, \ldots, D_k = D'$ such that $d(D_i, D_{i+1}) \leq 1$ for $i = 0, 1, \ldots, k-1$ (Cho and Awan, 2024). While the similarity of $M(D)$ and $M(D')$ was previously formalized in terms of divergences on probability measures, Dong et al. (2022) established that it is most natural to use the language of hypothesis testing, using the concept of a *tradeoff function*.

For two distributions P and Q, the *tradeoff function* between P and Q is $T(P, Q) : [0, 1] \to [0, 1]$, where $T(P, Q)(\alpha) = \inf\{1 - \mathbb{E}_Q(\phi) \mid \mathbb{E}_P(\phi) \leq 1 - \alpha\}$, and where the infimum is over all measurable tests ϕ. The tradeoff function returns the smallest type II error for testing $H_0 = P$ versus $H_1 = Q$ at specificity $\geq \alpha$ (i.e., $1 - \alpha$ is an upper bound on the type I error), and captures the difficulty of distinguishing between P and Q. The tradeoff function is closely related to the receiver operator characteristic curve (ROC curve), which evaluates sensitivity/power (one minus type II error) as a function of type I error.[1]

We say that a function $f : [0, 1] \to [0, 1]$ is a tradeoff function if there exist distributions P and Q such that $f(\alpha) = T(P, Q)(\alpha)$ for all $\alpha \in [0, 1]$. A function $f : [0, 1] \to [0, 1]$ is a tradeoff function if and only if f is convex, continuous, non-decreasing, and $f(x) \leq x$ for all $x \in [0, 1]$ (Dong

1. In Dong et al. (2022), the tradeoff function was originally defined as the smallest type II error as a function of type I error. Our choice to flip the tradeoff function along the x-axis follows that of Awan and Dong (2022) and is for mathematical convenience. We note that there is a typo in Awan and Dong (2022) where they mistakenly have the inequality in the wrong direction in the definition of $T(P, Q)$.

et al., 2022, Proposition 2.2). We say that a tradeoff function f is *nontrivial* if $f(\alpha) < \alpha$ for some $\alpha \in (0,1)$.

Definition 1 (f-DP: Dong et al., 2022) *Let f be a tradeoff function. A mechanism M satisfies f-DP if $T(M(D), M(D')) \geq f$, for all $D, D' \in \mathscr{D}$ such that $d(D, D') \leq 1$.*

If M satisfies f-DP, where $f = T(P, Q)$, then intuitively determining whether the database is D or D' is at least as hard as distinguishing between P and Q. Thus, if P and Q are difficult to distinguish, then the individual's private data is unlikely to be compromised. Since the differential privacy guarantee is symmetric (i.e., both $T(M(D), M(D'))$ and $T(M(D'), M(D))$ must be greater than f), it suffices to restrict the f-DP guarantee to *symmetric* tradeoff functions, meaning that if $f = T(P, Q)$, then $f = T(Q, P)$. Due to this, we limit the scope of the paper to symmetric tradeoff functions.

While f-DP consists of a wide variety of privacy guarantees, depending on the tradeoff function chosen, two important sub-families are worth pointing out. The common (ϵ, δ)-DP version of differential privacy is a special case of f-DP: let $\epsilon \in [0, \infty)$ and $\delta \in [0, 1]$, then M satisfies (ϵ, δ)-DP if and only if it satisfies $f_{\epsilon,\delta}$-DP, where

$$f_{\epsilon,\delta}(\alpha) = \max\{0, 1 - \delta - \exp(\epsilon) + \exp(\epsilon)\alpha, \exp(-\epsilon)(\alpha - \delta)\}.$$

When $\delta = 0$, we call $(\epsilon, 0)$-DP *pure DP*, which was the original definition of differential privacy. When $\epsilon = 0$, then $f_{0,\delta}$-DP can be thought of as δ-Total Variation DP, as it is equivalent to bounding the total variation distance between $M(D)$ and $M(D')$ by δ; recall that for two distributions P and Q, the *total variation distance* is $\mathrm{TV}(P, Q) = \sup_A |P(A) - Q(A)|$.

Another sub-family of f-DP is Gaussian-DP (GDP). We say that M satisfies μ-GDP if it satisfies G_μ-DP, where $G_\mu = T(N(0,1), N(\mu, 1))$. Besides offering the intuitive appeal of being defined in terms of shifted Gaussians, GDP also has a number of useful technical properties: the family G_μ-DP is closed under group privacy and composition, and these operations commute under G_μ-DP. Furthermore, Dong et al. (2022) established a "central limit theorem for composition," which shows that under mild conditions the composition of many DP mechanisms asymptotically satisfies μ-GDP for some μ.

For a symmetric tradeoff function f, denote by c_f the unique value such that $f(1 - c_f) = c_f$, which can be visualized as the intersection of f with the line $1 - \alpha$. The value c_f is a central concept to many of the results in this paper; we highlight a few basic properties here.

Lemma 2 *Let f be a symmetric tradeoff function. Then,*

1. *$c_f \in [0, 1/2]$ and if f is nontrivial then $c_f \in [0, 1/2)$,*
2. *if $f = T(P, Q)$, then $\mathrm{TV}(P, Q) = 1 - 2c_f$,*
3. *$f_{0,(1-2c_f)} \leq f \leq f_{\epsilon_f,0}$, where $\epsilon_f = \log\left(\frac{1-c_f}{c_f}\right)$.*

Property 2 of Lemma 2 makes an explicit connection between the tradeoff function and the total variation distance, which will be heavily used throughout the paper. By property 3 of Lemma 2, for a given symmetric tradeoff function f, we can construct nontrivial upper and lower bounds for f using only the value c_f. In particular, it shows that for any f, there is a pure-DP guarantee that implies f-DP, namely $(\epsilon_f, 0)$-DP where $\epsilon_f = \log((1 - c_f)/c_f)$.

Similar to other notions of differential privacy, f-DP has a number of important properties:

Group Privacy: The guarantee of f-DP is nominally for *groups of size 1*, since the adjacent databases differ in one entry. However, if M satisfies f-DP, then it also satisfies $f^{\circ k}$-DP for groups of size k (where $d(D, D') \leq k$), where $f^{\circ k} := f \circ \cdots \circ f$ and f appears k times. Note that while this guarantee is tight in general, for a specific f-DP mechanism there may exist a stronger guarantee for groups.

Composition: Another property of f-DP is that it can quantify the cumulative privacy cost of multiple privacy mechanisms. Specifically, if $M_1, \ldots, M_k$ each satisfy f_1-DP,$\ldots$, f_k-DP, respectively, and $(M_j(D))_j$ are mutually independent given D, then the joint release of $(M_1(D), \ldots, M_k(D))$ satisfies $f_1 \otimes \cdots \otimes f_k$-DP, where "$\otimes$" is the *tensor product* of tradeoff functions: if $f = T(P_1, Q_1)$ and $g = T(P_2, Q_2)$, then $f \otimes g = T(P_1 \times P_2, Q_1 \times Q_2)$. See Dong et al. (2022) for a more complex sequential composition result, where each $M_i(D)$ may depend on the results of $M_1(D), \ldots, M_{i-1}(D)$.

Invariance to Postprocessing: If $M : \mathscr{D} \to \mathscr{Y}$ satisfies f-DP and $T : \mathscr{Y} \to \mathscr{Z}$ is a (possibly randomized) function, then $T \circ M : \mathscr{D} \to Z$ also satisfies f-DP. This property is important to ensure that the privacy guarantee cannot be "undone" by a data-independent attack.

Additive Noise Mechanisms: The class of privacy mechanisms that we study in this paper are *additive noise mechanisms*, which add noise scaled to the *sensitivity* of a statistic. A statistic $S : \mathscr{D} \to \mathbb{R}$ has sensitivity $\Delta > 0$ if $|S(D) - S(D')| \leq \Delta$ for all $d(D, D') \leq 1$. An additive f-DP mechanism at sensitivity Δ is a random variable $N \sim F$ which satisfies $T(S(D)+N, S(D')+N) \geq f$ for all $d(D, D') \leq 1$ and all statistics S with sensitivity Δ. More succinctly, this is equivalent to $T(N, N+m) \geq f$ for all $m \in [-\Delta, \Delta] \cap \mathrm{Range}(S)$. If N is a continuous noise, then we can rescale both S and N by Δ.

Two of the most common additive noise mechanisms are the Laplace and Gaussian mechanisms: adding $\mathrm{Laplace}(0, \Delta/\epsilon)$ to a real-valued statistic S with sensitivity Δ satisfies $(\epsilon, 0)$-DP, and adding $N(0, \Delta^2/\mu^2)$ to the same statistic satisfies μ-GDP.

While elementary, additive noise mechanisms are widely used in DP, either by themselves or as part of more complex mechanisms. For example, using Hamming distance as the adjacency metric, a count statistic has sensitivity 1, the sample mean of n data points lying in $[a, b]$ has sensitivity $(b - a)/n$, and the sample variance of n data points lying in $[a, b]$ has sensitivity $(b - a)^2/n$ (Du et al., 2020); all of these statistics can be privatized to achieve f-DP by using an additive noise mechanism scaled to the sensitivity.

Stochastic Dominance and Optimal Noise Mechanisms: Let X and Y be two random variables. Stochastic dominance of random variables is a partial order, which asserts that one variable takes larger values compared to another. The following are equivalent definitions of (first order) stochastic dominance (Levy, 1992): 1) X stochastically dominates Y, 2) $F_X(t) \leq F_Y(t)$ for all $t \in \mathbb{R}$, 3) for every non-decreasing function ϕ, $\mathbb{E}\phi(X) \geq \mathbb{E}\phi(Y)$, 4) there exists a non-negative random variable W (not necessarily independent of Y) such that $X \overset{d}{=} Y + W$.

If N and N' are two additive noise mechanisms, we say that N is *stochastically smaller* than N' if $|N'|$ stochastically dominates $|N|$: $P(|N| \leq t) \geq P(|N'| \leq t)$ for all $t \in \mathbb{R}^{\geq 0}$. If we have that $P(|N| \leq t) \geq P(|N'| \leq t)$ for all N' in some class of noise distributions $\mathscr{A}$, we say that N is the *stochastically smallest* noise distribution in $\mathscr{A}$ or simply that N is *optimal* in $\mathscr{A}$.

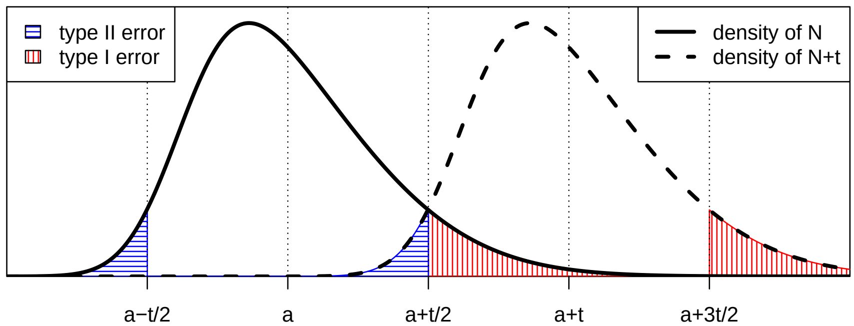

Figure 1: Illustration of Lemma 3, and the connection to hypothesis testing. When testing $H_0 : N$ versus $H_1 : N+t$, and using the rejection region $(a+t/2, \infty)$, the type I error probability is illustrated by the vertically shaded red region and the type II error probability is given by the horizontally shaded blue region.

3. A General Anti-Concentration Inequality

Concentration inequalities such as Markov, Chebyshev, and Bernstein, give upper bounds on the amount of mass in the tails of a distribution. In this section, we establish *anti-concentration inequalities* for additive f-DP mechanisms, which instead give upper bounds on how much mass is near the center of the distribution. Our inequalities establish these bounds in terms of total variation distance and f-DP tradeoff functions. The most basic version of our anti-concentration inequality is in Lemma 3, which bounds the amount of mass in the center of the distribution in terms of the total variation distance between the random variable and a shifted version of itself.

Lemma 3 *[Anti-Concentration Inequality] Let N be a real-valued random variable. Then for all $t \in \mathbb{R}^{>0}$, $\sup_{a\in\mathbb{R}} P(-t/2 < N - a \leq t/2) \leq \mathrm{TV}(N, N+t)$.*

Proof [Proof Sketch] We consider a hypothesis test of $H_0 : N$ versus $H_1 : N+t$, using the rejection region $(a+t/2, \infty)$. The type I and type II errors of this test correspond to the tail probabilities $P(N - a > t/2)$ and $P(N - a \leq t/2)$; see Figure 1. Combining this with the inequality (type I) + (type II) $\leq 1 - \mathrm{TV}(N, N+t)$ gives the result. ■

Remark 4 *Since total variation is a symmetric and translation-invariant relation, it follows that* $\mathrm{TV}(N, N+t) = \mathrm{TV}(N+t, N) = \mathrm{TV}(N, N-t)$*, which are alternative expressions for the right side of Lemma 3.*

If a random variable N is an additive noise mechanism at sensitivity 1 for f-DP, then $T(N, N+1) \geq f$. The following theorem establishes that in this case, we can express the upper bound $\mathrm{TV}(N, N+t)$ from Lemma 3 in terms of evaluations of f.

Theorem 5 *[Anti-Concentration for Additive Noise] Let f be a symmetric nontrivial tradeoff function and let N be a random variable such that $T(N, N+1) \geq f$. Then for $t \in \mathbb{Z}^{>0}$,*

$$\sup_{a\in\mathbb{R}} P(-t/2 < N - a \leq t/2) \leq \begin{cases} 1 - 2f^{\circ k}(c_f) & \text{if } t = 2k+1, \\ 1 - 2f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases}$$

where $k \in \mathbb{Z}^{\geq 0}$ and the notation $f^{\circ k}$ represents $f \circ \cdots \circ f$, where f appears k times.

Proof [Proof Sketch] By Lemma 31, we have that $T(N, N+t) \geq f^{\circ t}$. Then we apply the formula $\mathrm{TV}(N, N+t) \leq 1 - 2c_{f^{\circ t}}$ to the result of Lemma 3. Finally, Lemma 28 gives the more explicit formula for $c_{f^{\circ t}}$, which depends on the parity of t. The dependence on the parity comes from the expression $c_{f^{\circ t}} = F_f(-t/2)$, where F_f is the CND for f from Proposition 9 below, which has a different recursive expression depending on whether $t/2$ is an integer or half-integer. Note that Lemmas 28 and 31 appear in the appendix. ∎

Theorem 5 establishes that at all half-integer values, an additive noise distribution can only be so concentrated around zero. This result provides a minimum on the magnitude of noise that must be added to achieve f-DP. Lemma 3 and Theorem 5 are similar to anti-concentration inequalities of Lévy's concentration function (Krishnapur, 2016), which are of the form $\sup_{a\in\mathbb{R}} P(|X - a| \leq t) \leq g(t)$ for some function of t. Note that when N is continuous, the results of Lemma 3 and Theorem 5 simplify to this form.

Remark 6 *In Lemma 3 and Theorem 5, we could have written the probabilities as $P(-t/2 \leq N - a < t/2)$. In fact, a more general version of these anti-concentration inequalities is as follows: for any $c \in [0, 1]$ and any $a \in \mathbb{R}$,*

$$cP(N - a = -t/2) + P(-t/2 < N - a < t/2) + (1-c)P(N - a = t/2) \leq \mathrm{TV}(N, N+1),$$

where the "c" corresponds to the randomized test $\phi(x) = I(x > a + t/2) + cI(x = a + t/2)$ of $H_0 : N$ versus $H_1 : N + t$.

In the following sections, we will see that different canonical noise distributions match the bounds of Lemma 3 and Theorem 5, indicating that they are near-optimal to achieve f-DP.

4. Continuous Canonical Noise Distributions

While f-DP was introduced by Dong et al. (2022), the first method of constructing a privacy mechanism for an arbitrary f-DP guarantee was proposed by Awan and Vadhan (2023). Awan and Vadhan (2023) proposed the concept of a *canonical noise distribution* (CND), which captured the idea that the noise distribution satisfies f-DP and the privacy guarantee is tight. They also gave a construction for a CND for an arbitrary tradeoff function, and showed that CNDs are connected to certain optimal hypothesis tests. In this section, we review background on CNDs, show that they match the inequality of Theorem 5, prove that CNDs are sub-exponential, and establish optimality properties of log-concave CNDs.

4.1 Background on Canonical Noise Distributions

In this section, we review some of the key results of Awan and Vadhan (2023). A canonical noise distribution is an additive mechanism, which tightly matches the f-DP bound and has other desirable properties such as symmetry and a monotone likelihood ratio for testing $T(N, N+1)$. Awan and Vadhan (2023) proposed the following definition to capture these desirable properties:

Definition 7 (Canonical Noise Distribution: Awan and Vadhan (2023)) *Let f be a symmetric tradeoff function, and let N be a continuous random variable with cumulative distribution function (cdf) F. Then, F is a* canonical noise distribution *(CND) for f if*

1. *For any $m \in [0,1]$, $T(N, N+m) \geq f$,*
2. *$f(\alpha) = T(N, N+1)(\alpha)$ for all $\alpha \in (0,1)$,*
3. *$T(N, N+1)(\alpha) = F(F^{-1}(\alpha) - 1)$ for all $\alpha \in (0,1)$,*
4. *$F(x) = 1 - F(-x)$ for all $x \in \mathbb{R}$; that is, N is symmetric about zero.*

In Definition 7, the four properties can be interpreted as follows: 1) adding ΔN to a statistic of sensitivity Δ satisfies f-DP, 2) if $S(D)$ and $S(D')$ differ by exactly the sensitivity Δ, then $T(S(D) + \Delta N, S(D') + \Delta N) = f$ showing that the privacy guarantee is tight, 3) for $S(D)$ and $S(D')$ that differ by exactly the sensitivity Δ, an optimal rejection region is of the form $[x, \infty)$, which implies a monotone likelihood ratio property, and 4) since the privacy guarantee f-DP is symmetric, it is sensible to restrict attention to symmetric distributions.

The following recurrence is an important technical property of CNDs:

Lemma 8 (Awan and Vadhan (2023)) *Let f be a symmetric nontrivial tradeoff function and let F be a CND for f. Then $F(x) = 1 - f(1 - F(x-1))$ when $F(x-1) > 0$ and $F(x) = f(F(x+1))$ when $F(x+1) < 1$.*

Lemma 8 implies that if a CND is specified on an interval of length 1, then it is fully determined by the recurrence. Awan and Vadhan (2023) showed that the above recurrence relation can be used to construct a CND for any nontrivial symmetric tradeoff function, by starting with the linear function $c_f(1/2 - x) + (1 - c_f)(x + 1/2)$ on the interval $[-1/2, 1/2]$. This linear function is chosen to ensure that the resulting distribution is symmetric about zero and has a continuous cdf.

Proposition 9 (CND Construction: Awan and Vadhan (2023)) *Let f be a symmetric nontrivial tradeoff function. We define $F_f : \mathbb{R} \to \mathbb{R}$ as*

$$F_f(x) = \begin{cases} f(F_f(x+1)) & x < -1/2, \\ c_f(1/2 - x) + (1 - c_f)(x + 1/2) & -1/2 \leq x \leq 1/2, \\ 1 - f(1 - F_f(x-1)) & x > 1/2. \end{cases}$$

Then $N \sim F_f$ is a canonical noise distribution for f.

Proposition 9 demonstrates that CNDs exist for any symmetric nontrivial tradeoff function, and gives an explicit construction. Furthermore, Awan and Vadhan (2023) give an algorithm to sample from the constructed CND, enabling it to be employed in practice, which we have implemented in R code for this paper.

4.2 Near-Optimality of CNDs

While Awan and Vadhan (2023) demonstrated that CNDs are essential to constructing optimal hypothesis testing procedures, they did not give any direct results to show that CNDs add a minimal amount of noise. In this section, we show the concentration of a CND about zero is equal to the upper bound given in Theorem 5 at all $t \in \mathbb{Z}^{\geq 0}$. This implies that at half-integer values, the CND is more concentrated than any alternative additive noise satisfying f-DP. We also show that

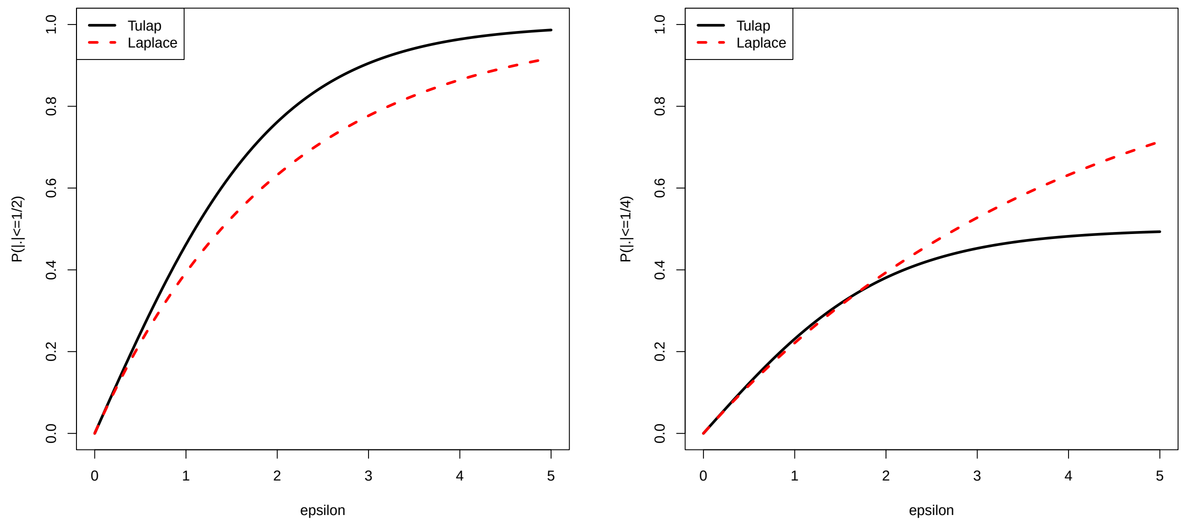

Figure 2: Comparison of central probabilities for Tulap and Laplace distributions, as ϵ varies. Left: $P(|\cdot| \leq 1/2)$, Right: $P(|\cdot| \leq 1/4)$. See Example 1 for details.

the mass of any other f-DP noise distribution is no more than a radius of 1/2 closer to its center than a CND.

Lemma 10 gives an explicit formula for the concentration of a CND, which we see matches the bound in Theorem 5.

Lemma 10 *Let f be a symmetric, nontrivial tradeoff function, and let $N \sim F_N$ be a CND for f. Then for all $t \in \mathbb{Z}^{\geq 0}$,*

$$P(|N| \leq t/2) = \begin{cases} 1 - 2f^{\circ k}(c_f) & \text{if } t = 2k+1, \\ 1 - 2f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases}$$

where $k \in \mathbb{Z}^{\geq 0}$

Combining Lemma 10 with Theorem 5, we get the following corollary, indicating that at half-integer values, CNDs are more concentrated than any alternative f-DP noise mechanism.

Corollary 11 *Let f be a symmetric, nontrivial tradeoff function. Let $N \sim F$ be a CND for f, and let N' be a random variable such that $T(N', N'+1) \geq f$. Then for every $t \in \mathbb{Z}^{\geq 0}$,*

$$P(|N| \leq t/2) \geq \sup_{a \in \mathbb{R}} P(-t/2 < N' - a \leq t/2).$$

It is important to acknowledge the limitations of Corollary 11, as the inequality need not hold at values which are not half-integers. See Example 1 below:

Example 1 *Recall from Awan and Dong (2022) that* $\mathrm{Tulap}(0, \exp(-\epsilon), 0)$ *is the unique CND for $(\epsilon, 0)$-DP, which can be constructed using Proposition 9 (Awan and Vadhan, 2023) (the name is an abbreviation of truncated-uniform-Laplace, which references an alternative construction of the* *distribution). It is also common knowledge that* $\mathrm{Laplace}(0, 1/\epsilon)$ *is another additive noise that can be used to satisfy* $(\epsilon, 0)$*-DP. Let* $N \sim \mathrm{Tulap}(0, \exp(-\epsilon), 0)$ *and let* $L \sim \mathrm{Laplace}(0, 1/\epsilon)$*. We see in the left plot of Figure 2 that* $P(|N| \leq 1/2) \geq P(|L| \leq 1/2)$ *for all values of* ϵ*. On the other hand, the right plot of Figure 2 shows that while* $P(|N| \leq 1/4) \geq P(|L| \leq 1/4)$ *for small values of* ϵ*, when* $\epsilon > 2$*,* $P(|N| \leq 1/4) \leq P(|L| \leq 1/4)$*. This indicates that the result of Corollary 11 cannot be extended for non-integer values of* t*.*

A consequence of Corollary 11 is that CNDs are nearly optimal compared to another f-DP noise distribution N', in the sense that at half-integer values, the cdf of $|N|$ dominates the cdf of $|N' - a|$. Corollary 12 provides another perspective, demonstrating that $|N' - a|$ stochastically dominates $|N| - 1/2$. This result can also be interpreted as saying that the mass of a competing f-DP noise distribution is no more than a radius of $1/2$ closer to its center than a CND.

Corollary 12 *Let* f *be a symmetric, nontrivial tradeoff function. Let* $N \sim F$ *be a CND for* f*, and let* N' *be a random variable such that* $T(N', N' + 1) \geq f$*. Then for every* $t \in \mathbb{R}^{\geq 0}$*,*

$$P(|N| \leq t + 1/2) \geq \sup_{a \in \mathbb{R}} P(-t < N' - a \leq t).$$

If N' *is a continuous random variable, this simplifies to* $P(|N| \leq t+1/2) \geq \sup_{a \in \mathbb{R}} P(|N' - a| \leq t)$*.*

Remark 13 *While Lemma 10 establishes that the CND for a tradeoff function* f *matches the anti-concentration bounds of Theorem 5, these are not the only distributions that can do so. For example, we see in Theorem 5 that the bounds only depend on the values* $f^{\circ k}(c_f)$ *and* $f^{\circ k}(1/2)$*, which often do not fully determine the tradeoff function* f*. Thus, if* $g \geq f$ *satisfies* $c_g = c_f$*,* $g^{\circ k}(c_f) = f^{\circ k}(c_f)$ *and* $g^{\circ k}(1/2) = f^{\circ k}(1/2)$ *for all* $k \geq 0$*, then a CND for* g *satisfies* f*-DP and also matches the anti-concentration bounds of Theorem 5.*

4.3 CNDs are Sub-Exponential

In addition to comparing CNDs with other distributions, we can also directly analyze the tail behavior of CNDs. In the following result, we establish that all CNDs are sub-exponential, implying that their tail behavior has at slowest an exponential decay. As demonstrated in Example 2 this result indicates that even if the target f-DP guarantee is designed in terms of a heavier tailed distribution (such as Cauchy), a CND only requires sub-exponential tails.

Tail behavior of a noise mechanism is important to understand the dispersion, the probability of obtaining extreme values, and whether moments exist. With heavy-tailed distributions, such as the Cauchy distribution, it is possible that the moments do not exist. On the other hand, for sub-exponential random variables, all moments are guaranteed to be finite.

A random variable X with mean μ is (σ^2, α)*-sub-exponential* if $\mathbb{E} \exp(\lambda(X - \mu)) \leq \exp(\lambda^2 \sigma^2/2)$ for all $|\lambda| < 1/\alpha$. The bound on the moment generating function, implies an exponential decaying bound on the tails of the distribution by the Chernoff bound. In the following theorem, $\lfloor t \rfloor := \max\{n \in \mathbb{Z} | n \leq t\}$ is the floor function.

Theorem 14 *Let* f *be a symmetric nontrivial tradeoff function and let* $N \sim F$ *be a CND for* f*. Set* $\epsilon_f = \log((1 - c_f)/c_f)$*. Then the following hold:*

1. $P(|N| > t) \leq \exp(-\epsilon_f \lfloor t \rfloor) \leq \exp(-\epsilon_f (t - 1))$ *for* $t \geq 0$*,*
2. $\mathbb{E}|N|^n \leq \epsilon_f^{-n} \exp(\epsilon_f) n!$ *for* $n \in \mathbb{Z}^{>0}$*,*

3. N is $(4\exp(\epsilon_f)/\epsilon_f^2, 2/\epsilon_f)$-sub-exponential.

Proof [Proof Sketch] Using part 3 of Lemma 2, we upper bound f with the piece-wise linear tradeoff function $f_{\epsilon_f,0}$. Using this along with the recurrence in Lemma 8, we obtain the inequality in part 1. Part 2 is obtained by integrating the result of part 1, and part 3 is obtained from part 2 by using the Bernstein condition for sub-exponential random variables. ■

An interesting consequence of Theorem 14 is that there is no privacy-relevant reason to add noise with tails heavier than exponential. We explore this in terms of "Cauchy-DP" in the following example.

Example 2 (Cauchy-DP) *For $m \geq 0$, let $C_m = T(\text{Cauchy}(0,1), \text{Cauchy}(m,1))$. Intuition suggests that C_m-DP is a stronger notion of DP than $(\epsilon, 0)$-DP, since Cauchy noise has heavier tails than Laplace or Tulap noise. However in this example, we establish that C_m-DP is equivalent to $(\epsilon, 0)$-DP, up to a change of variables.*

Reimherr and Awan (2019, Example 6) previously proved that adding Cauchy noise to a statistic of sensitivity m satisfies $(\epsilon_L(m), 0)$-DP, where

$$\epsilon_L(m) = \log\left(\frac{4 + (m + \sqrt{m^2+4})^2}{4 + (m - \sqrt{m^2+4})^2}\right)$$

(while their result is stated for multivariate t-distributions with degrees of freedom $\nu > 1$, the argument also works for the Cauchy distribution, which is a univariate t-distribution with 1 degree of freedom). This implies that $C_m \geq f_{\epsilon_L(m),0}$. This means that "Cauchy-DP" is a stronger notion of privacy than $(\epsilon, 0)$-DP, which agrees with our intuition.

On the other hand, using part 3 of Lemma 2, we have that $C_m \leq f_{\epsilon_U(m),0}$, where $\epsilon_U(m) = \log\left((1-c_m)/c_m\right)$. We know that $1 - 2c_m = \text{TV}(\text{Cauchy}(0,1), \text{Cauchy}(m,1)) = (2/\pi)\arctan\left(\frac{m}{2}\right)$ (Nielsen and Okamura, 2022), where c_m is shorthand for c_{C_m}. It follows that

$$\epsilon_U(m) = \log\left(\frac{3\pi - \arctan(m/2)}{2\arctan(m/2) - \pi}\right).$$

In Figure 3, we plot C_1 as well as $f_{\epsilon_U,0}$ and $f_{\epsilon_L,0}$. Theorem 14 tells us that any CND for C_m will have sub-exponential tails, indicating that a CND can obtain the same privacy guarantee as the Cauchy distribution, but with less noise.

In the right plot of Figure 3 we compare $P(|N_i| \leq t)$ where $N_2 \sim \text{Cauchy}(0,1)$ and N_1 is drawn from the CND for C_1 constructed in Proposition 9. Empirically, for $t \geq 1/2$, $P(|N_1| \geq t) \geq P(|N_2| \geq t)$ and the difference is at most $\approx .00425$ for $t \leq 1/2$. Thus, the CND would be preferred over the Cauchy distribution in most applications. To illustrate the difference in the tails, we simulated 100 random variables from both distributions. The maximum absolute value of the CND variables was 4.28 whereas the maximum absolute value of the Cauchy random variables was 118.43. We see that the subexponential tails of the CND significantly reduce the chances of observing extreme events.

4.4 Optimality of Log-Concave CNDs

Among the additive noise distributions, log-concave distributions form an important subclass. Log-concave CNDs are also continuous CNDs and because of this they inherit all of the properties of Sections 4.2 and 4.3. In this section, we give a new construction for the (unique) log-concave CND,

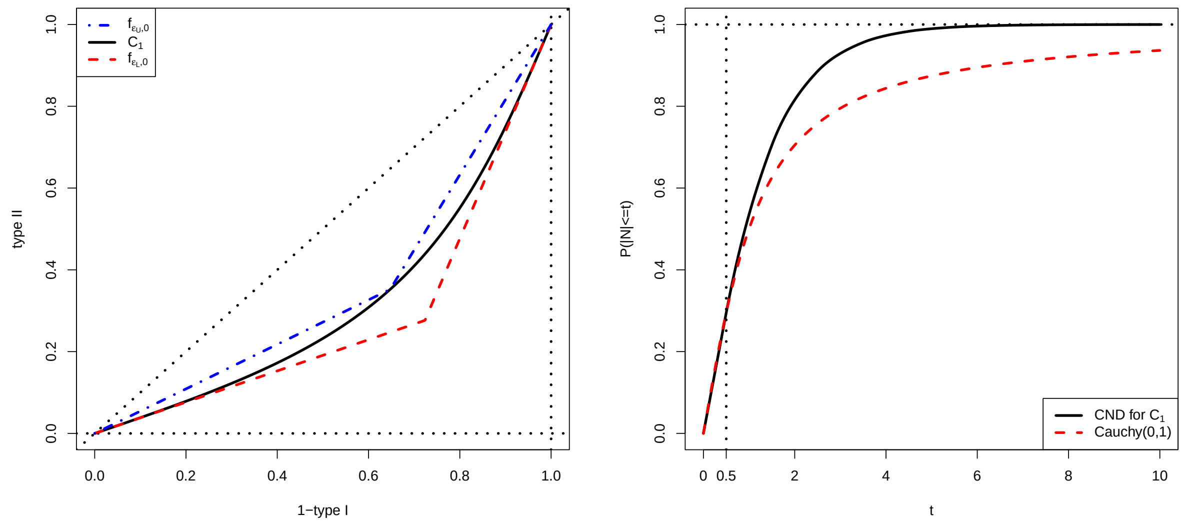

Figure 3: Left: Tradeoff function $C_1 = T(\text{Cauchy}(0,1), \text{Cauchy}(1,1))$ as well as $f_{\epsilon_U,0}$ and $f_{\epsilon_L,0}$, from Example 2. Right: Comparison of $P(|N| \leq t)$ where $N \sim \text{Cauchy}(0,1)$ or N is a CND for C_1. Vertical line is at $t = 1/2$ and horizontal line is at 1.

when it exists, and we also show that log-concave CNDs are stochastically smallest among additive noises with the same privacy guarantee.

A continuous random variable is log-concave if its density can be expressed in the form $g(x) \propto \exp(C(x))$, where C is a concave function. Log-concave distributions have many nice properties. In particular, a random variable N with cdf F satisfies $T(N, N+t)(\alpha) = F(F^{-1}(\alpha)-t)$ for every $t > 0$ if and only if it is log-concave (Dong et al., 2022). This is equivalent to stating that log-concave distributions have the monotone likelihood ratio property when testing between any two shifted versions of N.

Awan and Vadhan (2023) established conditions for the existence and construction of log-concave CNDs in terms of whether a tradeoff function is *infinitely divisible*. A collection of tradeoff functions $\{f_t | t \geq 0\}$ is infinitely divisible if it has the following properties: 1) $f_s \circ f_t = f_{s+t}$ for all $s, t \geq 0$, 2) f_s is nontrivial for all $s > 0$, and 3) $f_s \to f_0 = \text{Id}$ as $s \downarrow 0$, where $\text{Id}(\alpha) = \alpha$. Infinite divisibility formalizes whether the tradeoff functions in the collection can be achieved by group privacy. Awan and Dong (2022) showed that a symmetric nontrivial tradeoff function has a log-concave CND if and only if it belongs to an infinitely divisible collection of tradeoff functions. Furthermore, Awan and Dong (2022) gave a construction for the unique log-concave CND of an infinitely divisible collection of tradeoff functions, which was expressed as a limit of CNDs constructed via Proposition 9. In the following lemma, we give a more explicit construction of the same log-concave distribution. Interestingly, the log-concave CND only depends on a single value from each of the f_t's: $f_t(1/2) = c_{f_{2t}}$. Lemma 15 also offers another way of understanding why the log-concave CND for f_1 is unique.

Lemma 15 *Let $\{f_t | t \geq 0\}$ be an infinitely divisible collection of tradeoff functions, and let F be the log-concave CND for f_1. Then, for $t \in \mathbb{R}^{>0}$, $F(-t) = f_t(1/2) = c_{f_{2t}}$.*

Most importantly, Lemma 15 can be used to construct the log-concave CND F from the family of tradeoff functions, but it can also be used to compute the value c_{f_t}, when f_t is log-concave:

Example 3 (Laplace c_f) *Let $f_t(\alpha) = T(\text{Laplace}(0,1), \text{Laplace}(t,1))$, which is infinitely divisible, and recall that* Laplace(0, 1) *is the log-concave CND for f_1. Then, we can calculate $c_{f_t} = F(-t/2) = (1/2)\exp(-t/2)$, which is simply an evaluation of the Laplace cdf.*

In Theorem 16, we establish that the log-concave CND N is stochastically smallest, compared to any other noise distribution which satisfies $T(N', N'+t) \geq f_t$ for all $t \geq 0$.

Theorem 16 *Let $\{f_t | t \geq 0\}$ be an infinitely divisible collection of tradeoff functions, and let F be the log-concave CND for f_1. Let $N \sim F$ and let N' be any random variable such that $T(N', N'+t) \geq f_t$ for all $t \geq 0$. Then $|N' - a|$ stochastically dominates $|N|$ for all $a \in \mathbb{R}$. It follows that*

- *$P(|N| \leq t) \geq P(|N' - a| \leq t)$ for all $a \in \mathbb{R}$ and all $t \in \mathbb{R}^{\geq 0}$,*
- *For any non-decreasing function ϕ, $\mathbb{E}_N \phi(|N|) \leq \mathbb{E}_{N'} \phi(|N' - a|)$ for all $a \in \mathbb{R}$,*
- *There exists a non-negative, random variable W such that $|N'| \overset{d}{=} |N| + W$. If N' is symmetric about zero, then $N' \overset{d}{=} N + \text{sign}(N) \cdot W$.*

While the condition $T(N', N'+t) \geq f_t$ for all $t \geq 0$ is not required to achieve f-DP (it is only required that $T(N', N'+t) \geq f$ for $t \in [-1, 1]$), this stronger condition is related to *individual privacy accounting* (Rogers et al., 2016; Feldman and Zrnic, 2021; Koskela et al., 2023), which tracks each individual's privacy budget along several sequential compositions, such as in DP stochastic gradient descent. By using infinitely divisible tradeoff functions along with the condition $T(N', N'+t) \geq f_t$ for all $t \geq 0$, we can use f_t as a lower bound for privacy accounting within the $\{f_t | t \geq 0\}$ family. Theorem 16 shows that among these types of noise distributions, the log-concave CND is optimal.

Stochastic dominance is a very strong ordering of random variables. While most prior works focus on a specific objective criterion to optimize such as mean absolute error and mean squared error (Geng and Viswanath, 2015a), optimal hypothesis testing (Awan and Slavković, 2018), and Wasserstein distance (Qin et al., 2022b)), stochastic dominance implies that the mechanism optimizes all symmetric objectives centered at the non-private value, which are non-decreasing away from the center.

Example 4 *Examples of log-concave distributions include Gaussian, Laplace, Logistic, Uniform, and Beta, as well as truncated versions of these distributions. Of these, the Gaussian distribution is the log-concave CND for GDP, Laplace is the log-concave CND for Laplace-DP (Awan and Dong, 2022), and uniform is the log-concave CND for $(0, \delta)$-DP (Awan and Dong, 2022). Dong (2020) showed that the logistic tradeoff function is a lower bound on the privacy guarantee for exponential mechanisms (McSherry and Talwar, 2007).*

5. Discrete Canonical Noise Distributions

In Section 4, the canonical noise distributions we considered were all continuous, as were those considered in the previous literature (Awan and Vadhan, 2023; Awan and Dong, 2022). However, there are also important use-cases where integer-valued noise is preferable. For example, the US Census Bureau used the discrete Gaussian mechanism (Canonne et al., 2020) to privatize the 2020

US Census products. Other papers have advocated for discrete noise such as the Skellam distribution (Agarwal et al., 2021), binomial distribution (Dwork et al., 2006a; Agarwal et al., 2018), and discrete Laplace/geometric mechanism (Ghosh et al., 2012). Even when one may theoretically prefer a continuous noise distribution, there may still be a benefit of using a discrete noise distribution to ensure that an implementation on finite computers has guaranteed privacy properties (e.g., floating point calculations are vulnerable to privacy attacks; see Mironov, 2012).

In this section, we propose a definition for a "discrete CND," which captures the idea that 1) it is an integer-valued distribution, 2) for integer-valued statistics with finite sensitivity, it can be used to achieve f-DP, 3) the privacy guarantee is "tight," and 4) it satisfies a monotone-likelihood ratio property analogous to property 3 in Definition 7.

Definition 17 *Let $\Delta \in \mathbb{Z}^{>0}$, let f be a symmetric tradeoff function, and let N be a random variable with cdf F. Then F is a* discrete CND *for f at sensitivity Δ if*

1. *$T(N, N+t) \geq f$, for all $t \in \{-\Delta, \ldots, \Delta\}$,*
2. *$f(F(t+\Delta)) = F(t)$ for all $t \in \mathbb{Z}$, such that $F(t+\Delta) < 1$,*
3. *N takes values in $\mathbb{Z}$, and is symmetric about zero.*

Property 1 of Definition 17 implies that for an integer-valued statistic $S(D)$, with sensitivity Δ, $T(S(D)+N, S(D')+N) \geq f$ for any adjacent databases D, and D'. Property 2 implies a monotone likelihood ratio property for N, and that the tradeoff function for the discrete CND is "tight" for f. Essentially, using the rejection set (t, ∞) where $t \in \mathbb{Z}$, the type I error for $T(N-\Delta, N)$ is $1 - F(t+\Delta)$ and type II error is $F(t)$.

Note that in property 2) of Definition 17, the tradeoff function $T(N, N+\Delta)$ only matches f for rejection regions of the form (t, ∞), where $t \in \mathbb{Z}$. In general, this means that the tradeoff function $T(N, N+\Delta)$ agrees with f at the "sharp points," but may be greater than f at other values of α. See Figure 4 for an illustrative example.

Remark 18 *One key difference between Definition 17 for discrete CNDs and Definition 7 for continuous CNDs is the dependence on the sensitivity. For continuous CNDs, the sensitivity was not included in the definition, as the noise can simply be scaled up or down according to the sensitivity. However, in the discrete case, scaling affects the support of the distribution. For example, if N takes values on the integers and we try to scale it by 2 to adjust for statistics of sensitivity 2, then $2N$ takes values on the even numbers. In this case, when shifting by a smaller amount, we fail to guarantee f-DP: $T(2N, 2N+1) = 0$. Instead, the discrete CND must be designed for a specific sensitivity, as indicated in Definition 17.*

5.1 Existence, Construction, and Uniqueness of Discrete CNDs

In Proposition 19, we demonstrate that a discrete CND exists for any nontrivial symmetric tradeoff function and any sensitivity, and we can construct a discrete CND by rounding a re-scaled continuous CND. In the case of sensitivity 1, we show in Proposition 20 that the discrete CND is *unique*.

In the following proposition, the round function is formally defined as $\mathrm{round}(t) = \lfloor t + 1/2 \rfloor$.

Proposition 19 (Existence and Construction via Rounding) *Let f be a symmetric nontrivial tradeoff function, let N_c be a (continuous) CND for f with cdf F, and let $\Delta \in \mathbb{Z}^{>0}$. Then $N = \mathrm{round}(\Delta N_c)$, which has cdf $F_N(t) = F_{N_c}(\lfloor t + 1/2 \rfloor / \Delta)$, is a discrete CND for f at sensitivity Δ.*

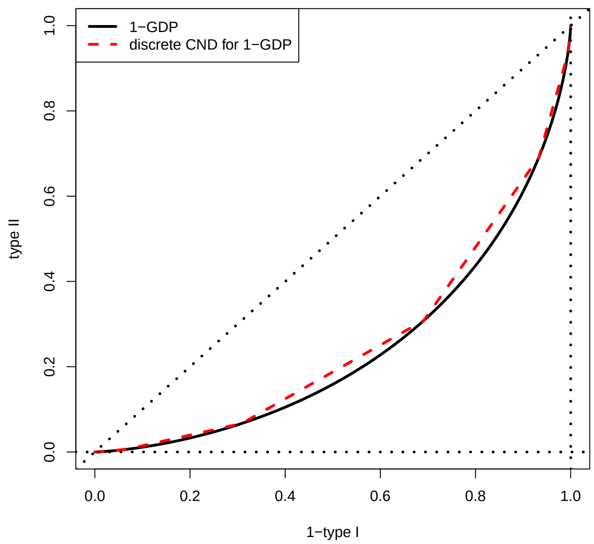

Figure 4: Tradeoff function of 1-GDP as well as the tradeoff function $T(N, N+1)$, where N is the unique discrete CND for 1-GDP at sensitivity 1.

In particular, combining Proposition 19 with Proposition 9, we have an explicit construction for a discrete CND, and by using Proposition F.6 from Awan and Vadhan (2023) (reprinted as Lemma 41 in the appendix), we also have a sampling algorithm for this discrete CND.

Example 5 (Staircase Distribution) *In the case of (ϵ, δ)-DP, the discrete CND constructed by rounding the constructed CND of Proposition 9 has a staircase shape, which has been identified in other DP literature as a near-optimal distribution (Geng and Viswanath, 2015b; Qin et al., 2022a). The pmf of this distribution is illustrated in Figure 5.*

Example 6 (Discrete CNDs Not by Rounding) *While Proposition 19 tells us that we can construct a discrete CND by rounding a continuous CND, there are also discrete CNDs that do not arise in this manner. Recall that for $(\epsilon, 0)$-DP, the unique (continuous) CND is the Tulap distribution (Awan and Dong, 2022). It follows that for each ϵ and Δ, there is a unique discrete CND for $(\epsilon, 0)$-DP generated by rounding the continuous CND. However, in the case of $\epsilon = 1$ and $\Delta = 2$, we demonstrate that there is an infinite family of discrete CNDs:*

In the case of $\Delta = 2$, constraints 2 and 3 of Definition 17 leave only one degree of freedom: the choice of $F(0)$; this is because symmetry enforces that $F(-1) = 1 - F(0)$ and the recurrence of 2 fixes all other values of F. One can then verify that for

$$0.5938455 \approx \frac{2\exp(\epsilon)}{3\exp(\epsilon)+1} \leq F(0) \leq \frac{\exp(\epsilon)+1}{\exp(\epsilon)+3} \approx 0.6502446,$$

one obtains a discrete CND for $(1, 0)$-DP. Since there is more than one discrete CND for $(1, 0)$-DP at $\Delta = 2$, and there is only one continuous CND for $(1, 0)$-DP, it follows that not all discrete CNDs are a rounding of continuous CNDs.

Interestingly, for the special case of $\Delta = 1$, Proposition 20 below establishes that there is a *unique* discrete CND for a nontrivial symmetric tradeoff function.

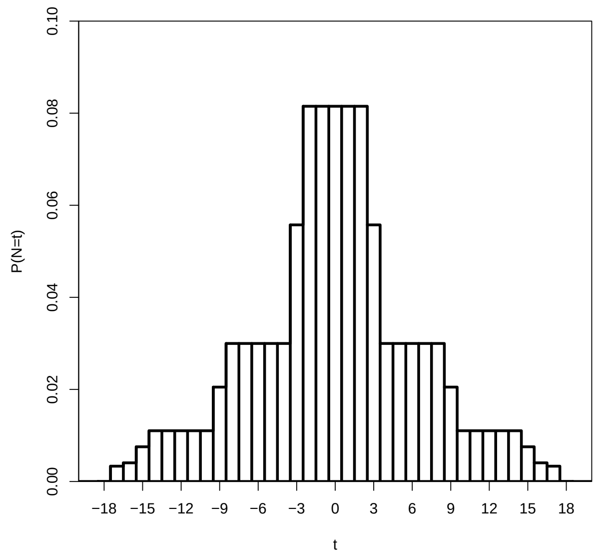

Figure 5: Discrete staircase distribution, which is a discrete CND for $(1, .05)$-DP at $\Delta = 6$.

Proposition 20 (Uniqueness) *If $\Delta = 1$, then there is a unique discrete CND for a symmetric nontrivial tradeoff function f. Furthermore, it can be realized by rounding any (continuous) CND for f, and has pmf $P(N = x) = f^{\circ|x|}(1 - c_f) - f^{\circ|x|}(c_f)$ for $x \in \mathbb{Z}$.*

This result follows from the observation that once $F(0)$ is chosen, the recurrence of property 2 of Definition 17 fully determines the cdf, and the symmetry via property 3 of Definition 17 removes the freedom in the choice of $F(0)$. Note further that combining Propositions 19 and 20 tells us that rounding any CND (including the one constructed in Proposition 9) will result in the unique discrete CND at sensitivity 1.

Example 7 (Discrete Laplace) *The unique discrete CND for $f_{\epsilon,0}$, at sensitivity 1 is the discrete Laplace / geometric mechanism, with pmf*

$$P(N = x) = \frac{\exp(\epsilon) - 1}{\exp(\epsilon) + 1} \exp(-\epsilon|x|), \quad x \in \mathbb{Z}.$$

This distribution is well-known in the differential privacy literature, and has been identified as the optimal noise-adding mechanism in pure differential privacy (Ghosh et al., 2012) for count statistics (which have sensitivity 1).

Example 8 (Rounded Gaussian Versus Discrete Gaussian) *Proposition 19 tells us that we can round a Gaussian random variable to produce a discrete CND for Gaussian-DP (since Gaussian noise is a CND). Furthermore, by Proposition 20, the rounded Gaussian is the unique discrete CND at sensitivity 1. The rounded Gaussian has been referred to as the discrete normal distribution, and has been independently studied in the statistics literature (Roy, 2003); it has pmf $P(N = x) = \Phi((x + 1/2)/\sigma) - \Phi((x - 1/2)/\sigma)$. On the other hand, the discrete Gaussian distribution (Canonne et al., 2020) (also sometimes called the discrete normal distribution (Kemp, 1997)) has gained*

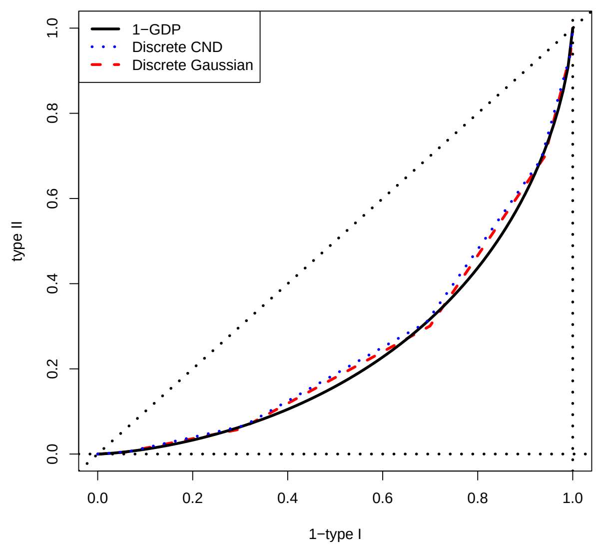

Figure 6: Tradeoff function for $T(N, N+1)$, where N is a discrete Gaussian distribution with $\mu = 0$ and $\sigma = 1$, against the tradeoff function G_1, as well as the discrete CND for G_1 at sensitivity 1. Note that $T(N, N+1)$ is below G_1 at some values, violating the GDP guarantee.

popularity as a privacy mechanism, and has been employed in the 2020 Decennial Census. The discrete Gaussian with parameter $\mu = 0$ and scale parameter σ has pmf

$$P(N = x) = \frac{\exp(-x^2/(2\sigma^2))}{\vartheta_3(0, \exp(-1/(2\sigma^2)))},$$

where $\vartheta_3(0, q) = \sum_{k=-\infty}^{\infty} q^{k^2}$ is a Jacobi theta function (Szabłowski, 2001).

In this example, we demonstrate that while the discrete Gaussian distribution with parameter $\sigma^2 = 2/\rho$ satisfies ρ-zCDP for integer-valued statistics of sensitivity 1 (the same zCDP guarantee as the continuous Gaussian mechanism with the same scale parameter, Canonne et al., 2020), the discrete Gaussian with $\sigma = 1/\mu$ does not satisfy μ-GDP. For example, with $\mu = \sigma = 1$, and setting $f_{DG} = T(N, N+1)$, where N is a discrete Gaussian with $\sigma = 1$, we have by the monotone likelihood ratio property of the discrete Gaussian that

$$c_{f_{DG}} = (1/2)(1 - P(N = 0)) = (1/2)(1 - [\vartheta_3(0, \exp(-1/2(\sigma^2)))]^{-1}) \approx 0.301,$$

which is smaller than $c_{G_1} = \Phi(-1/2) \approx 0.309$, which implies that $f_{DG} \not\geq G_1$; see Figure 6.

Remark 21 *While a tradeoff function has a unique discrete CND for sensitivity 1, there may be multiple tradeoff functions with the same discrete CND for sensitivity 1. In particular, we see from Proposition 19 that the discrete CND only depends on $\{f^{\circ n}(1 - c_f)\}_{n=0}^{\infty}$ or equivalently $\{f^{\circ n}(c_f)\}_{n=-1}^{\infty}$. So, if two tradeoff functions agree on these points, then they will have the same discrete CND for sensitivity 1.*

5.2 Optimality of Discrete CND at Sensitivity 1

In this section, we establish a special case of Theorem 5 for integer-valued noise, and show that the unique discrete CND at $\Delta = 1$ matches the bound. Using this result, we establish that the discrete CND at $\Delta = 1$ is stochastically smaller than any other integer-centered discrete noise.

While Theorem 5 offers an anti-concentration inequality for any additive noise, the inequality is not enforced at all values of the support. For integer-valued noises for statistics with sensitivity 1, Theorem 5 can give an inequality at every integer.

Corollary 22 *Let N be an integer-valued random variable such that $T(N, N+1) \geq f$, where f is a nontrivial symmetric tradeoff function. Then for all $t \in \mathbb{Z}^{\geq 0}$,*

$$\sup_{a \in \mathbb{Z}} P(|N - a| \leq t) \leq 1 - 2f^{\circ t}(c_f).$$

Similar to Lemma 10, Lemma 23 shows that the unique discrete CND at sensitivity 1 matches the bound of Corollary 22.

Lemma 23 *Let N be the unique discrete CND for a symmetric nontrivial tradeoff function f at sensitivity $\Delta = 1$. Then for $t \in \mathbb{Z}^{\geq 0}$, $P(|N| \leq t) = 1 - 2f^{\circ t}(c_f)$.*

It follows from Lemma 23 that the unique discrete CND is stochastically smaller than any alternative DP noise, a similar result to Theorem 16.

Theorem 24 *[Stochastic Optimality of Discrete CND for Sensitivity 1] Let N be the unique discrete CND for a symmetric nontrivial tradeoff function f at sensitivity $\Delta = 1$, and let N' be any integer-valued random variable such that $T(N', N'+1) \geq f$. Then, $|N' - a|$ stochastically dominates $|N|$ for all $a \in \mathbb{Z}$. This implies that,*

- *$P(|N| \leq t) \geq P(|N' - a| \leq t)$ for all $a \in \mathbb{Z}$ and all $t \in \mathbb{Z}^{\geq 0}$,*
- *For any non-decreasing function ϕ, $\mathbb{E}_N \phi(|N|) \leq \mathbb{E}_{N'} \phi(|N' - a|)$ for all $a \in \mathbb{Z}$,*
- *There exists a non-negative, integer-valued, random variable W such that $|N'| \overset{d}{=} |N| + W$. If N' is symmetric about zero, then $N' \overset{d}{=} N + \mathrm{sign}(N) \cdot W$.*

Theorem 24 is similar to the results of Ghosh et al. (2012). The key differences are 1) our results are for general f-DP instead of only $(\epsilon, 0)$-DP, and 2) our results are phrased in terms of stochastic dominance rather than universal utility maximizers. In the case of $(\epsilon, 0)$-DP both our Theorem 24 and the results of Ghosh et al. (2012) suggest that the discrete Laplace mechanism, discussed in Example 7 is the optimal discrete noise for count statistics.

Example 9 *Let S be an integer-valued statistic with sensitivity 1. Suppose that we are considering two privacy mechanisms $M_1(D) = S(D) + N$ and $M_2(D) = S(D) + N'$, where N is the discrete CND for f at sensitivity 1 and N' is another noise distribution satisfying $T(N', N'+1) \geq f$.*

- *Let ψ be a symmetric loss function, which is increasing away from zero (e.g., $\psi(x) = |x|$ or $\psi(x) = x^2$). Then by Theorem 24, $\mathbb{E}\psi(M_1(D) - S(D)) \leq \mathbb{E}\psi(M_2(D) - S(D))$. This is a similar result to those in Ghosh et al. (2012) and Geng and Viswanath (2015b), indicating that the discrete CND at sensitivity 1 is a universal utility optimizer.*
- *If $\mathbb{E}(N') = 0$, then $\mathrm{Var}(N) \leq \mathrm{Var}(N')$, by taking $\phi(x) = x^2$ in Theorem 24.*

Example 10 (Rounded Gaussian Versus Discrete Gaussian, Continued) *In Example 8, we saw that the discrete Gaussian with $\sigma = 1/\mu$ does not satisfy μ-GDP for sensitivity 1 statistics. Even if there were a different scale parameter $\sigma(\mu)$, for which the discrete Gaussian does satisfy μ-GDP, Theorem 24 tells us that the rounded Gaussian—the discrete CND at sensitivity 1 for μ-GDP—will be stochastically smaller than the discrete Gaussian.*

Example 11 (Non-Integer-Valued Center) *If the noise N' in Theorem 24 does not have an integer-valued mean, then the results of Theorem 24 do not imply that* $\mathrm{Var}(N) \leq \mathrm{Var}(N')$. *For example, recall that $U(-1,1)$ is a CND for $f_{0,1/2}$ (Awan and Dong, 2022). Then $N = \mathrm{round}(U)$ is the discrete CND for $f_{0,\delta}$ at sensitivity 1 and $N' = \lfloor U \rfloor$ is another integer-valued noise distribution which satisfies $T(N', N'+1) \geq f$ by postprocessing. We can calculate that $P(N=-1) = P(N=1) = 1/4$, $P(N=0) = 1/2$ and $P(N'=0) = P(N'=-1) = 1/2$. From this, we see that* $\mathrm{Var}(N) = 1/2 > 1/4 = \mathrm{Var}(N')$.

This counter-example demonstrates that the result of Theorem 24 cannot be extended to permit non-integer values of a, and that the variance is not necessarily minimized by the discrete CND at sensitivity 1 if a competing discrete noise has a non-integer mean.

Example 12 (Non-Optimality for Continuous CNDs) *While Theorem 24 may seem intuitive, it is worth pointing out that the analogous result does not hold for continuous CNDs. Recall that* $\mathrm{Tulap}(0, e^{-\epsilon}, 0)$ *is the unique CND for $f_{\epsilon,0}$ (Awan and Dong, 2022). Recall further that $L \sim$* $\mathrm{Laplace}(1/\epsilon)$ *is an additive mechanism that can be used to satisfy $f_{\epsilon,0}$-DP. However, there is no stochastic dominance relationship between Tulap and Laplace. For example, at $\epsilon = 5$, the variance of the Tulap is ≈ 0.097, whereas the variance of the Laplace is ≈ 0.08. On the other hand, $P(|N| \leq 1/2)$ is equal to $(\exp(\epsilon)-1)/(\exp(\epsilon)+1) \approx .987$ for Tulap and $1 - \exp(-\epsilon/2) \approx .918$ for Laplace. Furthermore, Geng and Viswanath (2015b) derived the minimum variance and minimum mean absolute error additive (continuous) mechanisms for $(\epsilon, 0)$-DP, which are not canonical noise distributions. So, while a (continuous) CND optimizes the privacy budget, and matches the anti-concentration inequality (Lemma 10) at half-integer values, it does not necessarily optimize other objectives.*

6. Discussion

Additive noise is a fundamental technique to achieve differential privacy, often either being employed by itself, or as part of a more complex privacy mechanism. Due to this, it is essential to understand the limits of what types of noise can be used to satisfy DP, and to optimize the noise distribution.

In this paper, we have explored the constraints on noise distributions for differential privacy, establishing upper bounds on the amount of mass that can be concentrated near the center of the distribution. We showed that canonical noise distributions (CNDs) match these bounds, which leads to their near-optimality. We also showed that log-concave CNDs are optimal compared to other noise distributions with the same privacy property. To address integer-valued statistics, we proposed a definition of a discrete CND which extended the original canonical noise distribution definition. We showed that discrete CNDs always exist and give an explicit construction. In the case that they are being added to a statistic of sensitivity 1, we showed that the discrete CND is unique and the smallest of any other integer-centered discrete noise.

In addition to our theoretical contributions, the R code for this paper includes a general method to sample from the CND constructed in Proposition 9 as well as functions to evaluate its cdf and quantile functions. These methods can enable researchers to implement and study CNDs in future works.

Limitations and Future Work: For continuous additive noise, the upper bounds on anti-concentration hold only at half-integer values. It would be worth exploring whether other bounds can be derived at other values. Alternatively, it would be interesting to investigate specific objective criteria, such as mean squared error, or mean absolute error, which have not been studied in the f-DP framework.

In terms of discrete noise, our definition of a discrete CND imposed that the distribution be symmetric about zero. The optimality result of Theorem 24 also only compared against other noise distributions centered at an integer value. Future research could weaken the discrete CND definition to not require symmetry about zero. While introducing a potential bias, this would allow for a wider range of noise mechanisms to be considered.

We also point out that the optimality of the discrete CND, Theorem 24, only applies in the sensitivity 1 case. Near-optimality results similar to Corollaries 11 and 12 can likely be derived for the discrete CND when sensitivity is greater than 1.

While this paper focused on additive noise mechanisms, our results can be used to optimize the performance of more complex mechanisms which use noise addition as an intermediate step (e.g., stochastic gradient descent (Abadi et al., 2016), objective perturbation (Chaudhuri et al., 2011; Kifer et al., 2012)). Future researchers may investigate these privacy mechanisms to determine whether optimizing the additive noise distributions results in improved performance of the complete privacy mechanism.

This paper focused on the central DP model, where a mechanism is applied to the whole dataset D which is held by a single curator. As Dong et al. (2022) discussed, there is a natural version of f-local DP, where a mechanism is applied to a single entry at a time; this offers a stronger privacy protection because the mechanism can be applied to each individual's data before sending the results to the curator. With this formulation, all of our results still apply in the local DP setting. We leave it to future researchers to investigate any particular nuances that arise in the local DP setting, such as the design of iterative mechanisms.

Multivariate Extension: The results of this paper are limited to univariate noise distributions. Future researchers may develop a multivariate extension to Lemma 3, and explore whether there are optimality properties of the multivariate CNDs proposed by Awan and Dong (2022).

While our techniques can be applied to multivariate distributions by restricting to 1-dimensional projections, as demonstrated below, this has limited utility. Let $N \in \mathbb{R}^d$ be a continuous f-DP additive noise distribution with respect to the norm $\|\cdot\|$; that is, $T(N, N+v) \geq f$ for all $\|v\| \leq 1$ (Awan and Dong, 2022). Then,

$$\begin{aligned}
P(\|N\| \leq t/2) &\leq \inf_{\|v\|=1} P(|\mathrm{Proj}_v N| \leq t/2) && (1)\\
&\leq \inf_{\|v\|=1} \mathrm{TV}(\mathrm{Proj}_v N, \mathrm{Proj}_v N + t) && (2)\\
&\leq \inf_{\|v\|=1} \mathrm{TV}(N, N + tv) && (3)\\
&\leq \begin{cases} 1 - 2f^{\circ k}(c_f) & \text{if } t = 2k+1 \\ 1 - 2f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases} && (4)
\end{aligned}$$

where Proj_v is the orthogonal projection onto the subspace spanned by the vector v, (2) applies Lemma 3, (3) applies the data processing inequality for total variation, and (4) uses the same argument as in the proof of Theorem 5. While this result does give some indication of the concentration of the distribution of N, it is generally not tight as the bound does not involve the dimension. The following example illustrates this gap in the case of Gaussian noise.

Example 13 *The distribution $N(0,I)$ is a multivariate CND with respect to $\|\cdot\|_\infty$ for $G_{\sqrt{d}}$-DP (Awan and Dong, 2022). We can directly calculate $P(\|N\|_\infty \leq t/2) = [1-2\Phi(-t/2)]^d$. On the other hand, the bound* (3) *simplifies to $P(\|N\|_\infty \leq t/2) \leq 1-2\Phi(-t/2)$, and we see that the bound is tight only in the case that $d=1$.*

We believe that the reason our method does not produce a tight bound in multivariate settings is because it is based on a two-point hypothesis test, similar in spirit to the Le Cam method for minimax lower bounds. It is possible that using a packing argument, similar to the Fano or Assouad methods, could produce a tighter bound that incorporates the dimension.

Acknowledgments

We are thankful to Zhanyu Wang for feedback and discussion on an early draft of this manuscript. We are also grateful to the anonymous reviewers, whose feedback has greatly improved the presentation of this paper. This work was supported in part by the National Science Foundation, Grant No. SES-2150615 to Purdue University.

Appendix A. Proofs and Technical Details

In this section, we include the proofs and technical details for the results in the paper.

Lemma 25 *Let f be a symmetric tradeoff function. Then,*

1. *$c_f \in [0,1/2]$ and if f is nontrivial then $c_f \in [0,1/2)$,*
2. *if $f = T(P,Q)$, then $\mathrm{TV}(P,Q) = 1-2c_f$,*
3. *$f_{0,(1-2c_f)} \leq f \leq f_{\epsilon_f,0}$, where $\epsilon_f = \log\left(\frac{1-c_f}{c_f}\right)$.*

Proof

1. Since $f(0)=0$, $f(1)\geq 0$, and f is an increasing convex function, it follows that there is a unique solution $f(1-c_f)=c_f$. Since $f(\alpha)\leq\alpha$, it follows that $c_f\in[0,1/2]$. Furthermore, if $c_f = 1/2$, then by convexity, it follows that $f(\alpha)=\alpha$ for all α, implying that it is not nontrivial.

2. Recall the equivalence between total variation and $f_{0,\delta}$: if δ is the smallest value such that $f_{0,\delta}\leq f$, then $\mathrm{TV}(P,Q)=\delta$. Recall that the $f_{0,\delta}$ curves have slope 1 and have value $1-\delta$ at $\alpha=1$. Since f is symmetric, we see that the tightest $f_{0,\delta}$ curve supports f at the point $(1-c_f,c_f)$. We calculate that the line with point $(1-c_f,c_f)$ with slope 1 has value $2c_f$ at $\alpha=1$. We conclude that $\mathrm{TV}(P,Q)=1-2c_f$.

3. As discussed in the proof of part 2, $f_{0,1-2c_f}\leq f$. For $f_{\epsilon,0}$, we call $c_\epsilon := c_{f_{\epsilon,0}}$ as shorthand. We can calculate that $c_\epsilon = (1+e^\epsilon)^{-1}$, which implies that $\epsilon = \log\left(\frac{1-c_\epsilon}{c_\epsilon}\right)$. Note that $f_{\epsilon_f,0}$ is the continuous, piece-wise linear tradeoff function with break points $(0,0)$, $(1-c_f,c_f)$, and $(1,1)$. Since f is a tradeoff function, we have $f(0)=0=f_{\epsilon_f,0}(0)$, $f(1-c_f)=c_f=f_{\epsilon_f,0}$, and $f(1)\leq 1 = f_{\epsilon_f,0}$. Thus, since f is convex, it follows that $f(\alpha)\leq f_{\epsilon_f,0}(\alpha)$ for all $\alpha\in[0,1]$.

∎

Lemma 26 *Let F be a CND for a symmetric nontrivial tradeoff function f. Then $F(-1/2) = c_f$.*

Proof Let $N \sim F$.

- (Case 1) If $F(-1/2) = 0$, then by symmetry we have that $F(1/2) = 1$. Consider the rejection region $[1/2, \infty)$ to test $T(N, N+1)$, which has type I error $P(N > 1/2) = 1 - F(1/2) = 0$ and type II error $P(N + 1 < 1/2) = P(N < -1/2) = F(-1/2) = 0$. We see that the tradeoff function includes the point $(1 - 0, 0)$, implying that $T(N, N+1) = 0$. This implies that $c_f = 0 = F(-1/2)$.
- (Case 2) Suppose that $F(-1/2) > 0$, which implies that $F(1/2) < 1$. Then, by Lemma 8, we have that $F(-1/2) = f(F(1/2)) = f(1 - F(-1/2))$ by symmetry. We see that $F(-1/2) = c_f$, since this is the unique solution to $c_f = f(1 - c_f)$.

■

Lemma 27 *[Lemma A.5: Awan and Dong, 2022] Let F be a CND for a nontrivial symmetric tradeoff function f. Then $F(k\cdot)$ is a CND for $f^{\circ k}$ for any $k \in \mathbb{Z}^{>0}$.*

Lemma 28 *Let f be a nontrivial tradeoff function, and let $t \in \mathbb{Z}^{>0}$. Then*

$$c_{f^{\circ t}} = \begin{cases} f^{\circ k}(c_f) & \textit{if } t = 2k+1, \\ f^{\circ k}(1/2) & \textit{if } t = 2k, \end{cases}$$

where $k \in \mathbb{Z}^{\geq 0}$ and the notation $f^{\circ k}$ represents $f \circ \cdots \circ f$, where f appears k times.

Proof Let F_f be the constructed CND for f, by Proposition 9. Then by Lemma 27, we know that $F_f(t\cdot)$ is a CND for $f^{\circ t}$. By Lemma 26, we have that $F_f(-t/2) = c_{f^{\circ t}}$. Using the recursive definition in Proposition 9, we can also write

$$\begin{aligned} c_{f^{\circ t}} &= F_f(-t/2) \\ &= \begin{cases} F_f(-k - 1/2) & \text{if } t = 2k+1, \\ F_f(-k) & \text{if } t = 2k, \end{cases} \\ &= \begin{cases} f^{\circ k}(F_f(-1/2)) & \text{if } t = 2k+1, \\ f^{\circ k}(F_f(0)) & \text{if } t = 2k, \end{cases} \\ &= \begin{cases} f^{\circ k}(c_f) & \text{if } t = 2k+1, \\ f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases} \end{aligned}$$

where we substituted $F_f(-1/2) = c_f$ and $F_f(0) = 1/2$ by Proposition 9. ■

Lemma 29 *[Anti-Concentration Inequality] Let N be a real-valued random variable. Then for all $t \in \mathbb{R}^{>0}$, $\sup_{a \in \mathbb{R}} P(-t/2 < N - a \leq t/2) \leq \mathrm{TV}(N, N+t)$.*

Proof Let $a \in \mathbb{R}$ be given. Let $N' \overset{d}{=} N - a$. For the hypothesis $H_0 : N'$ versus $H_1 : N' + t$, consider the following rejection region $(t/2, \infty)$. The type I and type II errors are $P(N' > t/2)$ and $P(N' + t \leq t/2) = P(N' \leq -t/2)$ respectively. Then

$$\begin{aligned}
P(-t/2 < N - a \leq t/2) &= P(-t/2 < N' \leq t/2) && (5)\\
&= 1 - P(N' > t/2) - P(N' \leq -t/2) && (6)\\
&= 1 - \text{type I} - \text{type II} && (7)\\
&\leq \mathrm{TV}(N', N' + t) && (8)\\
&= \mathrm{TV}(N, N + t), && (9)
\end{aligned}$$

where (8) uses the inequality: type I + type II $\geq 1 - \mathrm{TV}(N', N' + t)$, and (9) uses the fact that total variation is translation-invariant. ■

Lemma 30 (Lemma A.5: Dong et al., 2022) *Let f and g be tradeoff functions (not necessarily symmetric). If $T(P, Q) \geq f$ and $T(Q, R) \geq g$, then $T(P, R) \geq g \circ f$.*

Lemma 31 *Let N be a real-valued random variable, and call $f = T(N, N + 1)$ (not necessarily symmetric). Then for $t \in \mathbb{Z}^{>0}$, $T(N, N + t) \geq f^{\circ t}$.*

Proof We consider the sequence of tradeoff functions $T(N, N+1), T(N+1, N+2), \ldots T(N+t-1, N+t)$, which all have the tradeoff function $T(N, N + 1) \geq f$. By Dong et al. (2022, Lemma A.5), $T(N, N + t) \geq f^{\circ t}$. ■

Theorem 5 *[Anti-Concentration for Additive Noise] Let f be a symmetric nontrivial tradeoff function and let N be a random variable such that $T(N, N + 1) \geq f$. Then for $t \in \mathbb{Z}^{>0}$,*

$$\sup_{a \in \mathbb{R}} P(-t/2 < N - a \leq t/2) \leq \begin{cases} 1 - 2f^{\circ k}(c_f) & \text{if } t = 2k + 1, \\ 1 - 2f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases}$$

where $k \in \mathbb{Z}^{\geq 0}$ and the notation $f^{\circ k}$ represents $f \circ \cdots \circ f$, where f appears k times.

Proof We begin with Lemma 3:

$$\begin{aligned}
\sup_{a \in \mathbb{R}} P(-t/2 < N - a \leq t/2) &\leq \mathrm{TV}(N, N + t) && (10)\\
&\leq 1 - 2c_{f^{\circ t}} && (11)\\
&= \begin{cases} 1 - 2f^{\circ k}(c_f) & \text{if } t = 2k + 1, \\ 1 - 2f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases} && (12)
\end{aligned}$$

where (11) uses Lemma 31 to lower bound $T(N, N + t) \geq f^{\circ t}$ and property 2 of Lemma 2 to get $\mathrm{TV}(N, N + t) \leq 1 - 2c_{f^{\circ t}}$. Finally, (12) applies Lemma 28. ■

Lemma 32 *Let f be a symmetric, nontrivial tradeoff function, and let $N \sim F_N$ be a CND for f. Then for all $t \in \mathbb{Z}^{\geq 0}$,*

$$P(|N| \leq t/2) = \begin{cases} 1 - 2f^{\circ k}(c_f) & \textit{if } t = 2k+1, \\ 1 - 2f^{\circ k}(1/2) & \textit{if } t = 2k, \end{cases}$$

where $k \in \mathbb{Z}^{\geq 0}$

Proof We calculate

$$\begin{aligned} P(|N| \leq t/2) &= P(-t/2 \leq N \leq t/2) \\ &= 1 - 2F_N(-t/2) \\ &= 1 - 2\begin{cases} f^{\circ k}(F_N(-1/2)) & \text{if } t = 2k+1, \\ f^{\circ k}(F_N(0)) & \text{if } t = 2k, \end{cases} \\ &= \begin{cases} 1 - 2f^{\circ k}(c_f) & \text{if } t = 2k+1, \\ 1 - 2f^{\circ k}(1/2) & \text{if } t = 2k, \end{cases} \end{aligned}$$

where in the last line we used Lemma 26 to replace $F_N(-1/2) = c_f$, and symmetry to justify that $F_N(0) = 1/2$. ■

Corollary 33 *Let f be a symmetric, nontrivial tradeoff function. Let $N \sim F$ be a CND for f, and let N' be a random variable such that $T(N', N'+1) \geq f$. Then for every $t \in \mathbb{Z}^{\geq 0}$,*

$$P(|N| \leq t/2) \geq \sup_{a \in \mathbb{R}} P(-t/2 < N' - a \leq t/2).$$

Proof The result follows by combining Lemma 10 with Theorem 5. ■

Corollary 34 *Let f be a symmetric, nontrivial tradeoff function. Let $N \sim F$ be a CND for f, and let N' be a random variable such that $T(N', N'+1) \geq f$. Then for every $t \in \mathbb{R}^{\geq 0}$,*

$$P(|N| \leq t + 1/2) \geq \sup_{a \in \mathbb{R}} P(-t < N' - a \leq t).$$

If N' is a continuous random variable, this simplifies to $P(|N| \leq t + 1/2) \geq \sup_{a \in \mathbb{R}} P(|N' - a| \leq t)$.

Proof In this proof we use the notation $\lceil \cdot \rceil : \mathbb{R} \to \mathbb{Z}$ as the ceiling function, which returns the smallest integer greater than or equal to the input. We reparametrize the problem in terms of $s = 2t$ and consider the supremum of the ratio:

$$\begin{aligned} \sup_{s \in \mathbb{R}^{\geq 0}} \left(\frac{\sup_{a \in \mathbb{R}} P(-s/2 < N' \leq s/2)}{P(|N| \leq (s+1)/2)} \right) &\leq \sup_{s \in \mathbb{R}^{\geq 0}} \left(\frac{\sup_{a \in \mathbb{R}} P(-\lceil s \rceil/2 < N' \leq \lceil s \rceil/2)}{P(|N| \leq \lceil s \rceil/2)} \right) \\ &\leq \sup_{n \in \mathbb{Z}^{\geq 0}} \left(\frac{\sup_{a \in \mathbb{R}} P(-n/2 < N' \leq n/2)}{P(|N| \leq n/2)} \right) \\ &\leq 1, \end{aligned}$$

where we used the facts that $s \le \lceil s \rceil$ and $s + 1 \ge \lceil s \rceil$, we reparametrized the supremum in terms of $n = \lceil s \rceil$, and the final inequality follows from Corollary 11. ■

Recall that a random variable X is called (σ^2, b)-sub-exponential with mean μ if $\mathbb{E}\exp(\lambda(X - \mu)) \le \exp(\sigma^2\lambda^2/2)$ for all $|\lambda| < 1/b$. Lemma 35 is a sufficient condition to establish that a random variable is sub-exponential.

Lemma 35 (Bernstein Condition) *If $\mathbb{E}|X - \mu|^k \le \frac{1}{2}\sigma^2 b^{k-2} k!$, for all $k \ge 2$, then*

1. $\mathbb{E}\exp(\lambda(X - \mu)) \le \exp\left(\frac{\lambda^2\sigma^2}{2(1-b|\lambda|)}\right)$ *for all $|\lambda| < 1/b$, and*

2. *X is $(2\sigma^2, 2b)$-sub-exponential.*

Proof Part 1 is proved in Zhang and Chen (2021, Corollary 5.2). For part 2, let $|\lambda| < 1/(2b)$. Then, $1/(1 - b|\lambda|) < 2$. With part 1, we have $\mathbb{E}\exp(\lambda(X - \mu)) < \exp\left(\sigma^2\lambda^2\right)$. ■

Theorem 14 *Let f be a symmetric nontrivial tradeoff function and let $N \sim F$ be a CND for f. Set $\epsilon_f = \log\left((1 - c_f)/c_f\right)$. Then the following hold:*

1. $P(|N| > t) \le \exp(-\epsilon_f \lfloor t \rfloor) \le \exp(-\epsilon_f(t - 1))$ *for $t \ge 0$,*

2. $\mathbb{E}|N|^n \le \epsilon_f^{-n}\exp(\epsilon_f)n!$ *for $n \in \mathbb{Z}^{>0}$,*

3. *N is $(4\exp(\epsilon_f)/\epsilon_f^2, 2/\epsilon_f)$-sub-exponential.*

Proof

1. Recall, by part 3 of Lemma 2, that $f \le f_{\epsilon_f,0}$, where $\epsilon_f = \log\left(\frac{1-c_f}{c_f}\right)$. Then, $f(1/2) \le f_{\epsilon_f,0}(1/2) = (1/2)\exp(-\epsilon_f)$, since $c_f \le 1/2$. Recall that by Lemma 8 and symmetry, all CNDs satisfy $F_N(-z) = f^{\circ z}(F(0)) = f^{\circ z}(1/2)$ for $z \in \mathbb{Z}^{>0}$. So,
$$F_N(-z) = f^{\circ z}(1/2) \le f^{\circ z}_{\epsilon_f,0}(1/2) = (1/2)(\exp(-\epsilon_f))^z = (1/2)\exp(-\epsilon_f z).$$
Now let $t \ge 0$ (not necessarily an integer). Then,
$$\begin{aligned} P(|N| > t) &\le P(|N| > \lfloor t \rfloor) && (13)\\ &= 2F_N(-\lfloor t \rfloor) && (14)\\ &\le 2(1/2)\exp(-\epsilon_f\lfloor t \rfloor) && (15)\\ &\le \exp(-\epsilon_f(t-1)), && (16)\end{aligned}$$
where (14) used symmetry of CNDs and (15) used our earlier result.

2. Next, we calculate for $n \in \mathbb{Z}^{>0}$,
$$\begin{aligned} \mathbb{E}|N|^n &= \int_0^\infty nx^{n-1}P(|N| > x)\, dx && (17)\\ &\le \int_0^\infty nx^{n-1}\exp(-\epsilon_f(x-1))\, dx && (18)\\ &= \epsilon_f^{-1}\exp(\epsilon_f)n\mathbb{E}X^{n-1}, \quad \text{where } X \sim \mathrm{Exp}(\epsilon_f) && (19)\\ &= \epsilon_f^{-1}\exp(\epsilon_f)n\cdot(n-1)!\epsilon_f^{-(n-1)} && (20)\\ &= \epsilon_f^{-n}\exp(\epsilon_f)n!, && (21)\end{aligned}$$
where (17) uses Ross (2019, Self-Test Exercise 7.20) and (18) uses the result from part 1.

3. Recall that the Bernstein condition for mean-zero sub-exponential random variables states that if $\mathbb{E}|N|^n \leq (1/2)n!\sigma^2 b^{n-2}$, then N is $(2\sigma^2, 2b)$-sub-exponential. Comparing our result with this, we identify $\sigma^2 = 2\exp(\epsilon_f)/\epsilon_f^2$ and $b = 1/\epsilon_f$ and we have that the CND N is $(4\exp(\epsilon_f)/\epsilon_f^2, 2/\epsilon_f)$-sub-exponential.

■

Lemma 36 *Let $\{f_t | t \geq 0\}$ be an infinitely divisible collection of tradeoff functions, and let F be the log-concave CND for f_1. Then, for $t \in \mathbb{R}^{>0}$, $F(-t) = f_t(1/2) = c_{f_{2t}}$.*

Proof Because F is a log-concave CND for f_1, we have that $F_t(\cdot) := F(t\cdot)$ is a CND for f_t (Awan and Dong, 2022). Since $F(0) = 1/2$, we have that $F(-t) = F_t(-1) = f_t(1/2)$. It remains to establish the connection with $c_{f_{2t}}$.

Case 1: If $f_t(1/2) = 0$, then by symmetry of f_t, we have that $f_t(1) \leq 1/2$. This is because if $f = T(P,Q) = T(Q,P)$, then the type I error α and type II error β can be interchanged: $\beta = f(1-\alpha)$ if and only if $\alpha = f(1-\beta)$. It follows that

$$f_{2t}(1-0) = f_t^{\circ 2}(1) \leq f_t(1/2) = 0.$$

Since tradeoff functions are non-negative, we have that $f_{2t}(1-0) = 0$ and conclude $c_{f_{2t}} = 0 = f_t(1/2)$.

Case 2: If $f_t(1/2) > 0$, then consider the following:

$$\begin{aligned} f_{2t}(1 - f_t(1/2)) &= f_t^{\circ 2}(1 - F_t(-1)) && (22)\\ &= f_t^{\circ 2}(F_t(1)) && (23)\\ &= F_t(-1) && (24)\\ &= F(-t) && (25)\\ &= f_t(1/2), && (26) \end{aligned}$$

where in (24), we use the fact that $F_t(1) = 1 - F_t(-1) = 1 - F(-t) = 1 - f_t(1/2) < 1$ and apply the second recursion formula from Lemma 8. ■

Theorem 16 *Let $\{f_t | t \geq 0\}$ be an infinitely divisible collection of tradeoff functions, and let F be the log-concave CND for f_1. Let $N \sim F$ and let N' be any random variable such that $T(N', N'+t) \geq f_t$ for all $t \geq 0$. Then $|N' - a|$ stochastically dominates $|N|$ for all $a \in \mathbb{R}$. It follows that*

- *$P(|N| \leq t) \geq P(|N' - a| \leq t)$ for all $a \in \mathbb{R}$ and all $t \in \mathbb{R}^{\geq 0}$,*
- *For any non-decreasing function ϕ, $\mathbb{E}_N \phi(|N|) \leq \mathbb{E}_{N'} \phi(|N' - a|)$ for all $a \in \mathbb{R}$,*
- *There exists a non-negative, random variable W such that $|N'| \overset{d}{=} |N| + W$. If N' is symmetric about zero, then $N' \overset{d}{=} N + \mathrm{sign}(N) \cdot W$.*

Proof Let $t \in \mathbb{R}^{\geq 0}$ and let $a \in \mathbb{R}$. Then,

$$\begin{aligned}
P(|N' - a| < t) &\leq P(-t < N' - a \leq t) && (27)\\
&\leq \mathrm{TV}(N', N' + 2t) && (28)\\
&\leq 1 - 2c_{f_{2t}} && (29)\\
&= 1 - 2F_N(-t) && (30)\\
&= -F(-t) + F(t) && (31)\\
&= P(|N| < t), && (32)
\end{aligned}$$

where in (28) we apply Lemma 3, in (29) we use the assumption that $T(N', N' + 2t) \geq f_{2t}$ which implies that $\mathrm{TV}(N', N' + 2t) \leq 1 - 2c_{f_{2t}}$, in (30) we use Lemma 15, and (31) and (32) use the symmetry and continuity of F_N. Finally, we have that

$$P(|N' - a| \leq t) = \lim_{s \downarrow t} P(|N' - a| < s) \leq \lim_{s \downarrow t} P(|N| < s) = P(|N| \leq t).$$

The other statements follow as standard properties of stochastic dominance. ∎

Proposition 37 (Existence and Construction via Rounding) *Let f be a symmetric nontrivial tradeoff function, let N_c be a (continuous) CND for f with cdf F, and let $\Delta \in \mathbb{Z}^{>0}$. Then $N = \mathrm{round}(\Delta N_c)$, which has cdf $F_N(t) = F_{N_c}(\lfloor t + 1/2 \rfloor / \Delta)$, is a discrete CND for f at sensitivity Δ.*

Proof Property 1 of Definition 17 follows from fact that $T(\Delta N_c, \Delta N_c + \Delta) \geq f$ and the postprocessing property of tradeoff functions. For property 2 of Definition 17, note that

$$\begin{aligned}
P(N = x) &= P(x - 1/2 \leq \Delta N_c < x + 1/2)\\
&= F_{\Delta N_c}(x + 1/2) - F_{\Delta N_c}(x - 1/2),
\end{aligned}$$

since N_c is a continuous random variable. It follows that $F_N(t) = P(N \leq t) = F_{\Delta N_c}(t + 1/2)$ for all integers t. By Lemma 8, we know that $f(F_{N_c}(t+1)) = F_{N_c}(t)$ for all $t \in \mathbb{R}$ such that $F_{N_c}(t+1) < 1$, or equivalently, $f(F_{\Delta N_c}(t + \Delta)) = F_{\Delta N_c}(t)$ for all $t \in \mathbb{R}$ such that $F_{\Delta N_c}(t + \Delta) < 1$. Property 3 follows immediately from the symmetry of N_c. ∎

Proposition 38 (Uniqueness) *If $\Delta = 1$, then there is a unique discrete CND for a symmetric nontrivial tradeoff function f. Furthermore, it can be realized by rounding any (continuous) CND for f, and has pmf $P(N = x) = f^{\circ|x|}(1 - c_f) - f^{\circ|x|}(c_f)$ for $x \in \mathbb{Z}$.*

Proof First we will show that properties 2 and 3 of Definition 17 uniquely determine the cdf, so there is at most one discrete CND. Notice that given $F(0)$, the recursion in property 2 of Definition 17 fully determines the cdf. Property 3 of Definition 17 implies that $F(-1) = P(N < 0) = P(N > 0) = 1 - F(0)$. Combining this with $f(F(0)) = F(-1)$, which is from property 2, we have that $f(F(0)) = 1 - F(0)$, which implies that $F(0) = 1 - c_f$. We see that there is at most one discrete CND at sensitivity 1. Since Proposition 19 established that the rounding of any continuous CND gives a discrete CND, it follows that this construction results in the unique discrete CND at sensitivity 1.

Let N_c be any CND for f. Note that $F_{N_c}(1/2) = 1 - c_f$ and $F_{N_c}(-1/2) = c_f$, by Lemma 26. By the recursion in Lemma 8 it follows that for an integer $x <= 0$, $F_{N_c}(x + 1/2) = f^{\circ x}(1 - c_f)$ and

$F_{N_c}(x-1/2) = f^{\circ x}(c_f)$. Using the expression from the proof of Proposition 19, we have that for an integer $x \leq 0$,

$$P(N = x-1) = F_{N_c}(x-1+1/2) - F_{N_c}(x-1-1/2) = f^{\circ x}(1-c_f) - f^{\circ x}(c_f).$$

Finally, we have that $P(N = x)$ is symmetric with the desired formula:

$$\begin{aligned}
P(N = -x) &= F_{N_c}(-x+1/2) - F_{N_c}(-x-1/2) \\
&= 1 - F_{N_c}(x-1/2) - 1 + F_{N_c}(x+1/2) \\
&= F_{N_c}(x+1/2) - F_{N_c}(x-1/2) \\
&= P(N = x),
\end{aligned}$$

where we used the fact that F_{N_c} is symmetric (i.e., $F(-x) = 1 - F(x)$). ■

Corollary 39 *Let N be an integer-valued random variable such that $T(N, N+1) \geq f$, where f is a nontrivial symmetric tradeoff function. Then for all $t \in \mathbb{Z}^{\geq 0}$,*

$$\sup_{a \in \mathbb{Z}} P(|N-a| \leq t) \leq 1 - 2f^{\circ t}(c_f).$$

Proof Since N takes integer values, we can write

$$\sup_{a \in \mathbb{Z}} P(|N-a| \leq t) = \sup_{a \in \mathbb{Z}} P\Big(-(2t+1)/2 < N - a \leq (2t+1)/2\Big) \leq 1 - 2f^{\circ t}(c_f),$$

where the inquality uses Theorem 5. ■

Lemma 40 *Let N be the unique discrete CND for a symmetric nontrivial tradeoff function f at sensitivity $\Delta = 1$. Then for $t \in \mathbb{Z}^{\geq 0}$, $P(|N| \leq t) = 1 - 2f^{\circ t}(c_f)$.*

Proof Let F_f be the CND constructed in Proposition 9, and recall that $F_N(t) = F_f(t+1/2)$ for $t \in \mathbb{Z}$, by Proposition 20. Then,

$$\begin{aligned}
P(|N| \leq t) &= P(-t \leq N \leq t) && (33) \\
&= 1 - 2F_f(-t-1/2) && (34) \\
&= 1 - 2f^{\circ t}(F(-1/2)) && (35) \\
&= 1 - 2f^{\circ t}(c_f), && (36)
\end{aligned}$$

where in (34), we use the fact that F_f is the cdf of a symmetric continuous random variable, in (35) we use the recurrence in Proposition 9, and in (36) we use the construction in Proposition 9 that $F_f(-1/2) = c_f$. ■

Theorem 24 *[Stochastic Optimality of Discrete CND for Sensitivity 1] Let N be the unique discrete CND for a symmetric nontrivial tradeoff function f at sensitivity $\Delta = 1$, and let N' be any integer-valued random variable such that $T(N', N'+1) \geq f$. Then, $|N'-a|$ stochastically dominates $|N|$ for all $a \in \mathbb{Z}$. This implies that,*

- *$P(|N| \leq t) \geq P(|N' - a| \leq t)$ for all $a \in \mathbb{Z}$ and all $t \in \mathbb{Z}^{\geq 0}$,*
- *For any non-decreasing function ϕ, $\mathbb{E}_N \phi(|N|) \leq \mathbb{E}_{N'} \phi(|N' - a|)$ for all $a \in \mathbb{Z}$,*
- *There exists a non-negative, integer-valued, random variable W such that $|N'| \stackrel{d}{=} |N| + W$. If N' is symmetric about zero, then $N' \stackrel{d}{=} N + \mathrm{sign}(N) \cdot W$.*

Proof Let $t \in \mathbb{Z}^{\geq 0}$. Then, $\sup_{a \in \mathbb{Z}} P(|N' - a| \leq t) \leq 1 - 2f^{\circ t}(c_f) = P(|N| \leq t)$, where the inequality used Corollary 22 and the equality used Lemma 23. The first two bullets follow from properties of stochastic dominance. Using $a = 0$, we get the third bullet. ∎

A.1 Derivation of the Cauchy-DP Tradeoff Function

In Example 2, we implement the CND via Proposition 9 for $C_1 = T(\mathrm{Cauchy}(0, 1), \mathrm{Cauchy}(1, 1))$. Sampling from this distribution is straightforward provided that we can evaluate C_1 and have $c_1 := c_{C_1}$, using Lemma 41 below:

Lemma 41 (Proposition F.6: Awan and Vadhan, 2023) *Let f be a symmetric nontrivial tradeoff function and let F_f be as in Proposition 9. Then the quantile function $F_f^{-1} : (0, 1) \to \mathbb{R}$ for F_f can be expressed as*

$$F_f^{-1}(u) = \begin{cases} F_f^{-1}(1 - f(1 - u)) & u < c_f, \\ \frac{u - 1/2}{1 - 2c_f} & c_f \leq u \leq 1 - c_f, \\ F_f^{-1}(f(u)) + 1 & u > 1 - c_f. \end{cases}$$

Furthermore, for any $u \in (0, 1)$, the $F_f^{-1}(u)$ takes a finite number of recursive steps to evaluate. Thus, if $U \sim U(0, 1)$ then $F_f^{-1}(U) \sim F_f$.

Recall that in Example 2, we calculated that $c_1 = 1/2 - (1/\pi) \arctan(1/2)$, so it only remains to numerically evaluate the tradeoff function C_1. By the Neyman-Pearson Lemma, we know that the family of optimal hypothesis tests reject when the likelihood ratio is above a given threshold. Given an observation $x \in \mathbb{R}$ from either $\mathrm{Cauchy}(0, 1)$ or $\mathrm{Cauchy}(1, 1)$, the likelihood ratio statistic is

$$\mathrm{LRT}(x) = \frac{\pi(1 + x^2)}{\pi(1 + (x - 1)^2)} = 1 + \frac{2x - 1}{x^2 - 2x + 2},$$

For rejection threshold $t + 1 \geq 1$, we have the corresponding rejection region $[x_1(t), x_2(t)]$, where

$$x_1(t) = \frac{t + 1 - \sqrt{-t^2 + t + 1}}{t},$$

$$x_2(t) = \frac{t + 1 + \sqrt{-t^2 + t + 1}}{t}.$$

The type I error is $1 - \alpha(t) = F_C(x_2(t)) - F_C(x_1(t))$, where F_C is the cdf of $\mathrm{Cauchy}(0, 1)$, and the type II error is $C_1(\alpha(t)) = 1 - F_C(x_2(t) - 1) + F_C(x_1(t) - 1)$. Since $\alpha(t)$ is a monotone function, given $\alpha \in [1 - c, 1]$ we can numerically solve for t such that $\alpha(t) = \alpha$ and then evaluate $C_1(\alpha(t))$.

So far, we have a method of evaluating $C_1(\alpha)$ for $\alpha \in [1 - c, 1]$. Since $\mathrm{LRT}(1 - x) = 1/\mathrm{LRT}(x)$, we have that for $\alpha \in [0, 1 - c]$, the rejection region is of the form $(-\infty, 1 - x_2(t)] \cup [1 - x_1(t), \infty)$, where t is chosen to satisfy $1 - \alpha = F_C(1 - x_2(t)) + 1 - F_C(1 - x_1(t))$. Then $C_1(\alpha) = F_C(-x_1(t)) - F_C(-x_2(t))$.

Code to evaluate this tradeoff function is provided, along with code to implement Lemma 41 above, enabling us to sample from the CND of Proposition 9.

References

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

John M Abowd. The US Census Bureau adopts differential privacy. In *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2867–2867, 2018.

Naman Agarwal, Ananda Theertha Suresh, Felix Xinnan X Yu, Sanjiv Kumar, and Brendan McMahan. cpSGD: Communication-efficient and differentially-private distributed SGD. *Advances in Neural Information Processing Systems*, 31, 2018.

Naman Agarwal, Peter Kairouz, and Ziyu Liu. The Skellam mechanism for differentially private federated learning. *Advances in Neural Information Processing Systems*, 34:5052–5064, 2021.

Jordan Awan and Jinshuo Dong. Log-concave and multivariate canonical noise distributions for differential privacy. *Advances in Neural Information Processing Systems*, 35:34229–34240, 2022.

Jordan Awan and Aleksandra Slavković. Differentially private uniformly most powerful tests for binomial data. *Advances in Neural Information Processing Systems*, 31, 2018.

Jordan Awan and Salil Vadhan. Canonical noise distributions and private hypothesis tests. *Annals of Statistics*, 51(2):547–572, 2023.

Jordan Awan and Yue Wang. Differentially private Kolmogorov-Smirnov-type tests. *arXiv preprint arXiv:2208.06236*, 2022.

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

Clément L Canonne, Gautam Kamath, and Thomas Steinke. The discrete Gaussian for differential privacy. *Advances in Neural Information Processing Systems*, 33:15676–15688, 2020.

Kamalika Chaudhuri, Claire Monteleoni, and Anand D Sarwate. Differentially private empirical risk minimization. *Journal of Machine Learning Research*, 12(3), 2011.

Young Hyun Cho and Jordan Awan. Formal privacy guarantees with invariant statistics. *arXiv preprint arXiv:2410.17468*, 2024.

Jinshuo Dong. *Gaussian Differential Privacy and Related Techniques*. PhD thesis, University of Pennsylvania, 2020.

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.

Wenxin Du, Canyon Foot, Monica Moniot, Andrew Bray, and Adam Groce. Differentially private confidence intervals. *arXiv preprint arXiv:2001.02285*, 2020.

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: 24th Annual International Conference on the Theory and Applications of Cryptographic Techniques, St. Petersburg, Russia, May 28-June 1, 2006. Proceedings 25*, pages 486–503. Springer, 2006a.

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

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *ACM SIGSAC Conference on Computer and Communications Security*, pages 1054–1067, 2014.

Vitaly Feldman and Tijana Zrnic. Individual privacy accounting via a Rényi filter. *Advances in Neural Information Processing Systems*, 34:28080–28091, 2021.

Quan Geng and Pramod Viswanath. Optimal noise adding mechanisms for approximate differential privacy. *IEEE Transactions on Information Theory*, 62(2):952–969, 2015a.

Quan Geng and Pramod Viswanath. The optimal noise-adding mechanism in differential privacy. *IEEE Transactions on Information Theory*, 62(2):925–951, 2015b.

Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. *SIAM Journal on Computing*, 41(6):1673–1693, 2012.

Mangesh Gupte and Mukund Sundararajan. Universally optimal privacy mechanisms for minimax agents. In *ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 135–146, 2010.

Peter Kairouz, Sewoong Oh, and Pramod Viswanath. Extremal mechanisms for local differential privacy. *The Journal of Machine Learning Research*, 17(1):492–542, 2016.

Zeki Kazan, Kaiyan Shi, Adam Groce, and Andrew P Bray. The test of tests: A framework for differentially private hypothesis testing. In *International Conference on Machine Learning*, pages 16131–16151. PMLR, 2023.

Adrienne W Kemp. Characterizations of a discrete normal distribution. *Journal of Statistical Planning and Inference*, 63(2):223–229, 1997.

Daniel Kifer, Adam Smith, and Abhradeep Thakurta. Private convex empirical risk minimization and high-dimensional regression. In *Conference on Learning Theory*, pages 25–1. JMLR Workshop and Conference Proceedings, 2012.

Antti Koskela, Marlon Tobaben, and Antti Honkela. Individual privacy accounting with Gaussian differential privacy. In *The Eleventh International Conference on Learning Representations*, 2023.

Manjunath Krishnapur. Anti-concentration inequalities, 2016. URL `https://math.iisc.ac.in/~manju/anti-concentration.pdf`. Lecture notes.

Haim Levy. Stochastic dominance and expected utility: Survey and analysis. *Management Science*, 38(4):555–593, 1992.

Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *IEEE Symposium on Foundations of Computer Science (FOCS'07)*, pages 94–103. IEEE, 2007.

Ilya Mironov. On significance of the least significant bits for differential privacy. In *ACM Conference on Computer and Communications Security*, pages 650–661, 2012.

Ilya Mironov. Rényi differential privacy. In *IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275. IEEE, 2017.

Frank Nielsen and Kazuki Okamura. On f-divergences between Cauchy distributions. *IEEE Transactions on Information Theory*, 2022.

Shuying Qin, Jianping He, Chongrong Fang, and James Lam. Differential private discrete noise adding mechanism: Conditions, properties and optimization. *arXiv preprint arXiv:2203.10323*, 2022a.

Shuying Qin, Jianping He, Chongrong Fang, and James Lam. Differential private discrete noise adding mechanism: Conditions and properties. In *2022 American Control Conference (ACC)*, pages 946–951. IEEE, 2022b.

Matthew Reimherr and Jordan Awan. Elliptical perturbations for differential privacy. *Advances in Neural Information Processing Systems*, 32, 2019.

Ryan M Rogers, Aaron Roth, Jonathan Ullman, and Salil Vadhan. Privacy odometers and filters: Pay-as-you-go composition. *Advances in Neural Information Processing Systems*, 29, 2016.

Sheldon Ross. *First Course in Probability, A*. Pearson Higher Ed, 10th edition, 2019.

Dilip Roy. The discrete normal distribution. *Communications in Statistics-Theory and Methods*, 32(10):1871–1883, 2003.

Jordi Soria-Comas and Josep Domingo-Ferrer. Optimal data-independent noise for differential privacy. *Information Sciences*, 250:200–214, 2013. ISSN 0020-0255.

Paweł J Szabłowski. Discrete normal distribution and its relationship with Jacobi theta functions. *Statistics & Probability Letters*, 52(3):289–299, 2001.

Jun Tang, Aleksandra Korolova, Xiaolong Bai, Xueqiang Wang, and Xiaofeng Wang. Privacy loss in Apple's implementation of differential privacy on macOS 10.12. *arXiv preprint arXiv:1709.02753*, 2017.

Huiming Zhang and Songxi Chen. Concentration inequalities for statistical inference. *Communications in Mathematical Research*, 37(1):1–85, 2021. ISSN 2707-8523.