Log-Concave and Multivariate Canonical Noise Distributions for Differential Privacy

Jordan A. Awan
Department of Statistics
Purdue University
jawan@purdue.edu

Jinshuo Dong
Department of Computer Science
Northwestern University and IDEAL*
jinshuo@northwestern.edu

Abstract

A canonical noise distribution (CND) is an additive mechanism designed to satisfy f-differential privacy (f-DP), without any wasted privacy budget. f-DP is a hypothesis testing-based formulation of privacy phrased in terms of *tradeoff functions*, which captures the difficulty of a hypothesis test. In this paper, we consider the existence and construction of both log-concave CNDs and multivariate CNDs. Log-concave distributions are important to ensure that higher outputs of the mechanism correspond to higher input values, whereas multivariate noise distributions are important to ensure that a joint release of multiple outputs has a tight privacy characterization. We show that the existence and construction of CNDs for both types of problems is related to whether the tradeoff function can be decomposed by functional composition (related to group privacy) or mechanism composition. In particular, we show that pure ϵ-DP cannot be decomposed in either way and that there is neither a log-concave CND nor any multivariate CND for ϵ-DP. On the other hand, we show that Gaussian-DP, $(0, \delta)$-DP, and Laplace-DP each have both log-concave and multivariate CNDs.

1 Introduction

Differential privacy (DP), proposed by Dwork et al. [2006], is the state-of-the-art framework in formal privacy protection and is being implemented by tech companies, government agencies, and academic institutions. Over time, the DP community has developed many new DP mechanisms as well as new frameworks. Recently, f-DP (Dong et al. [2022] was proposed as a generalization of DP, allowing for tight calculations of group privacy, composition, subsampling, and post-processing. It was shown in Dong et al. [2022] that f-DP is provably the tightest version of DP that respects the post-processing property of DP. In particular, f-DP can be losslessly converted to Rényi-DP (or any f-divergence version of DP) as well as (ϵ, δ)-DP, but not vice-versa [Dong et al., 2022]. Furthermore, f-DP is equivalent (can be losslessly converted back and forth) to the privacy profile [Balle et al., 2018, 2020] and the privacy loss random variables [Sommer et al., 2019, Zhu et al., 2022].

f-DP is defined in terms of a *tradeoff function* or *receiver operator curve (ROC)*, which encapsulates the difficulty of conducting a hypothesis test between two distributions. If $f = T(P, Q)$ is the tradeoff function for testing between the distributions P and Q, then if a mechanism M satisfies f-DP, this means that given the output of M when run on one of two adjacent databases, it is at least as hard to determine which database was used, as it is to test between P and Q.

While f-DP has the many desirable theoretical properties listed above in its favor, there are limited techniques for working with f-DP, and few constructive mechanisms for an arbitrary f-DP guarantee. A notable exception is a *canonical noise distribution* (CND) from the recent paper Awan and Vadhan

*The Institute for Data, Econometrics, Algorithms, and Learning

Preprint. Under review.

[2021], which builds a one-dimensional additive noise mechanism designed to exactly satisfy f-DP, with no wasted privacy budget. Along with the intuitive idea that a CND is optimal in that it optimizes the privacy loss budget, Awan and Vadhan [2021] showed that CNDs are crucial to the construction of optimal DP hypothesis tests and free DP p-values. However, the CND construction given in Awan and Vadhan [2021] does not result in a smooth distribution, and in particular is not *log-concave*. Log-concavity is a desirable property because it implies that the distribution has a monotone likelihood ratio; this means that higher observed values are always more likely to have come from a higher input value than a lower one. Log-concavity thus makes the DP output much more interpretable, easily analyzed, and also has makes the calculation of the privacy cost simpler [Dong et al., 2021]. Furthermore, the results of Awan and Vadhan [2021] are limited to 1-dimensional distributions.

In this paper, we develop new properties of CNDs and f-DP, motivated by the following two questions,

1. Can we construct log-concave CNDs?
2. Can we construct multivariate CNDs?

Our Contributions The existence of both log-concave 1-dimensional CNDs and multivariate CNDs are intricately linked with properties related to group privacy and mechanism composition. Two highly desirable properties of a tradeoff function are *infinite divisibility* and *infinite decomposability*, meaning that the tradeoff function can be exactly achieved by n-fold group privacy or n-fold mechanism composition, respectively. We prove that a tradeoff function has a log-concave CND if and only if the tradeoff function is infinitely divisible, and give a construction for the *unique* log-concave CND in this case. We also show that if a tradeoff function is either infinitely divisible or decomposable, then we can construct a multivariate CND.

Along with the positive results listed above, we also include impossibility results. In particular, $(\epsilon, 0)$-DP is neither divisible nor decomposable, and in fact has neither a log-concave CND nor any multivariate CND. In contrast to $(\epsilon, 0)$-DP, two families that satisfy both infinite divisibility and infinite decomposability are μ-GDP and $(0, \delta)$-DP. While $(0, \delta)$-DP has limited applicability due to its weak protection for events with small probability, μ-GDP and related DP definitions (such as zero concentrated DP) have been gaining popularity. The results of this paper provide a new perspective supporting the adoption of GDP as the default privacy measure instead of $(\epsilon, 0)$-DP.

Organization In Section 2, we review concepts in f-DP and canonical noise distributions. In Section 3, we study 1-dimensional CNDs. In Section 3.1, we prove that the Tulap distribution is the *unique* CND for $(\epsilon, 0)$-DP. In Section 3.2, we propose the concept of infinite divisibility and prove that a tradeoff function has a log-concave CND if and only if it is infinitely divisible; we also give a construction to produce the log-concave CND from a family of infinitely divisible tradeoff functions. We prove that piece-wise linear tradeoff functions are generally not infinitely divisible in Section 3.3, and in particular $(\epsilon, 0)$-DP and several related tradeoff functions do not have log-concave CNDs. In Section 4, we propose a multivariate extension of CND. We give two general constructions of multivariate CNDs in Section 4.1 depending on whether a tradeoff function is decomposable or infinitely divisible. We give several examples of multivariate CNDs in Sections 4.2-4.5 for Gaussian DP, $(0, \delta)$-DP, (ϵ, δ)-DP, and Laplace-DP. In Section 4.6, we show that there is no multivariate CND for $(\epsilon, 0)$-DP, which implies that $(\epsilon, 0)$-DP is not decomposable. We conclude with discussion in Section 5. Proofs and technical details are found in the Appendix.

Related Work While there are many complex DP mechanisms, many use the fundamental building block of additive mechanisms (e.g., functional mechanism [Zhang et al., 2012], objective perturbation [Chaudhuri et al., 2011, Kifer et al., 2012], stochastic gradient descent [Abadi et al., 2016], and the sparse vector technique [Dwork et al., 2009, Zhu and Wang, 2020], to name a few). There have been many different additive mechanisms proposed in the literature, for different privacy purposes. We highlight the works that show some optimality property for the proposed noise distributions. This work is most directly building off of Awan and Vadhan [2021], who proposed the concept of canonical noise distributions as a method of quantifying what it means to fully use the privacy budget. There are also other works, which derive optimal mechanisms with respect to other metrics. Ghosh et al. [2012] showed that a discrete Laplace distribution is the universal utility maximizer for a general class of utility functions in pure-DP. Geng and Viswanath [2015b] proposed the staircase mechanism which they showed optimizes the ℓ_1 or ℓ_2 error for pure-DP. For (ϵ, δ)-DP, Geng and Viswanath [2015a] showed that either the staircase or a uniform distribution can achieve the optimal rate in terms of ℓ_1 and ℓ_2 error. Steinke and Ullman [2016] showed that the ℓ_∞-mechanisms is rate optimal when measuring utility in terms of ℓ_∞ error. Awan and Slavković [2020] derive optimal mechanisms

among the class of K-Norm Mechanisms, proposed by Hardt and Talwar [2010], in terms of various scale-independent measures, for a fixed statistic and sample size.

2 Differential privacy basics

Differential privacy ensures that given the output of a private mechanism, it is difficult for an adversary to determine whether an individual is present in the database or not. To satisfy DP, a privacy expert employs a *mechanism* M, which is a set of probability distributions M_D on a common space $\mathscr{Y}$, indexed by possible databases $D \in \mathscr{D}$. Let $d(D, D')$ be an integer-valued metric on the space of databases $\mathscr{D}$, which represents the number of entries that D and D' differ in. We call D and D' *adjacent* if $d(D, D') \leq 1$. While there are now many variants of DP, they all center around the idea that given a randomized algorithm M, for any two adjacent databases D, D', the distributions of $M(D)$ and $M(D')$ should be "similar." While many DP variants measure similarity in terms of divergences, f-DP formalizes similarity in terms of hypothesis tests. Intuitively, for two adjacent databases D and D', a mechanism M satisfies f-DP if given the output of M, it is difficult to determine whether the original database was D or D'. This is formalized in terms *tradeoff functions*.

For two distributions P and Q, the *tradeoff function* (or ROC) between P and Q is $T(P, Q) : [0, 1] \to [0, 1]$, where $T(P, Q)(\alpha) = \inf\{1 - \mathbb{E}_Q \phi \mid \mathbb{E}_P(\phi) \geq 1 - \alpha\}$, where the infinimum is over all measurable tests ϕ. The tradeoff function returns the optimal type II error for testing $H_0 = P$ versus $H_1 = Q$ at specificity (one minus type I error) α, and captures the difficulty of distinguishing between P and Q. [2] 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 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 2.1 (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 $D, D' \in \mathscr{D}$ which satisfy $d(D, D') \leq 1$.

Intuitively, a mechanism satisfies f-DP, where $f = T(P, Q)$, if testing $H_0 : M(D)$ versus $H_1 : M(D')$ is at least as hard as testing $H_0 : P$ versus $H_1 : Q$. Without loss of generality we can assume that f is *symmetric*, meaning that if $f = T(P, Q)$, then $f = T(Q, P)$. This is due to the fact that adjacency of databases is a symmetric relation [Dong et al., 2022, Proposition 2.4]. So, we limit the focus of this paper on symmetric tradeoff functions.

A key property of differential privacy is that it also implies privacy guarantees for groups. Dong et al. [2022, Theorem 2.14] showed that if a mechanism is f-DP, then it satisfies $f^{\circ k}$-DP, when the adjacency measure is changed to allow for a difference in k entries (where $f^{\circ k}$ means the functional composition of f with itself, k times). We call this *group privacy*, which is a central topic in differential privacy. Note that the bound $f^{\circ k}$ is not necessarily the tightest privacy guarantee for a particular mechanism.

Mechanism Composition quantifies the cumulative privacy cost of the output of k mechanisms. To express the tradeoff function resulting from composition, Dong et al. [2022] proposed the *tensor product* of tradeoff functions: if $f = T(P, Q)$ and $g = (P', Q')$, then $f \otimes g := T(P \times P', Q \times Q')$, which they show is well defined, commutative, and associative. They prove that if we have k mechanisms $M_1, \ldots, M_k$, which each satisfy f_1-DP, f_2-DP,$\ldots$, f_k-DP respectively, then the composition $(M_1, \ldots, M_k)$ satisfies $f_1 \otimes \cdots \otimes f_k$-DP (see Dong et al. [2022, Theorem 3.2] for a more precise statement).

The traditional framework of (ϵ, δ)-DP is a subclass of f-DP: Let $\epsilon \geq 0$ and $\delta \in [0, 1]$. A mechanism satisfies (ϵ, δ)-DP if it satisfies $f_{\epsilon,\delta}$-DP, where $f_{\epsilon,\delta}(\alpha) = \max\{0, 1 - \delta - e^{\epsilon} + e^{\epsilon}\alpha, \exp(-\epsilon)(\alpha - \delta)\}$. An important special case is $(\epsilon, 0)$-DP, which was the original definition of DP.

Another popular subclass is Gaussian-DP (GDP): For $\mu \geq 0$, a mechanism satisfies μ-GDP if it satisfies G_μ-DP, where $G_\mu = T(N(0, 1), N(\mu, 1))$. Gaussian-DP was proposed in Dong et al. [2022] and has several desirable properties, such as being closed under group privacy and closed under composition. Dong et al. [2022] also established a central limit theorem for tradeoff functions as the number of compositions approaches infinity, showing that under general assumptions the tradeoff function of the composed mechanisms approaches G_μ for some μ.

[2] In Dong et al. [2022], the tradeoff function was originally defined as a function of type I error. Our choice to flip the tradeoff function along the x-axis is for mathematical convenience. The ROC function is usually defined as the power (one minus type II error) as a function of type I error.

2.1 Canonical noise distributions

To satisfy DP, additive mechanisms must introduce noise proportional to the *sensitivity* of the statistic of interest. Let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. A statistic $S : \mathscr{D} \to \mathbb{R}^d$ has $\|\cdot\|$-*sensitivity* $\Delta > 0$ if $\|S(D) - S(D')\| \leq \Delta$ for all $d(D, D') \leq 1$. When $d = 1$, we use $|\cdot|$ as the default norm. Any additive mechanism, which releases $S(D) + \Delta N$, satisfies f-DP if $T(N, N+v) \geq f$ for all $\|v\| \leq 1$. The concept *canonical noise distribution* (CND) was proposed by Awan and Vadhan [2021] to capture when an additive mechanism satisfies f-DP, and "fully uses the privacy budget."

Definition 2.2 (Canonical noise distribution: Awan and Vadhan [2021])**.** Let f be a symmetric tradeoff function. A continuous random variable N with cumulative distribution function (cdf) 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 2.2, property 1 ensures that the additive mechanism using a CND satisfies f-DP, property 2 ensures that the privacy guarantee is tight, property 3 gives a closed form for the tradeoff function in terms of the CND's cdf, which is equivalent to enforcing a monotone likelihood ratio property, and property 4 imposes symmetry which is mostly for convenience.

An important property of CNDs is that they satisfy the following recurrence relation:

Lemma 2.3 (Awan and Vadhan [2021])**.** *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 $F(x + 1) < 1$.*

In Awan and Vadhan [2021], they showed that the above recurrence relation can be used to construct a CND for any nontrivial symmetric tradeoff function.

Proposition 2.4 (CND construction: Awan and Vadhan [2021])**.** *Let f be a symmetric nontrivial tradeoff function, and let $c \in [0, 1]$ be the solution to $f(1 - c) = c$. We define $F_f : \mathbb{R} \to \mathbb{R}$ as*

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

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

While Proposition 2.4 gives a general construction of a CND for an arbitrary f, the resulting distribution is generally not smooth or log-concave. Awan and Vadhan [2021] showed that in the case of G_μ, this construction does not recover the Gaussian distribution, which is the log-concave CND.

3 One-dimensional CNDs

In this section, we expand on the results of Awan and Vadhan [2021], by producing new results for one-dimensional CNDs. In Section 3.1, we show that the Tulap distribution is the *unique* CND for $(\epsilon, 0)$-DP. In Section 3.2, we propose the concept of an *infinitely divisible tradeoff function* and show that a tradeoff function has a log-concave CND if and only if it is infinitely divisible. We also give a construction to produce the unique log-concave CND for an infinitely divisible family of tradeoff functions. In Section 3.3, we determine when a piece-wise linear tradeoff function is divisible, and show that $f_{\epsilon,0}$ and related tradeoff functions are not infinitely divisible, and hence do not have log-concave CNDs.

3.1 CNDs for $(\epsilon, 0)$-DP

In Awan and Vadhan [2021], it was shown that in general, the CND is not unique, but it was not clear whether there existed alternative CNDs for $f_{\epsilon,0}$ or $f_{\epsilon,\delta}$. We begin this section by showing that

the Tulap distribution, which was shown to be a CND for $f_{\epsilon,\delta}$ by Awan and Vadhan [2021] is in fact the *unique* CND for $f_{\epsilon,0}$. The Tulap distribution was proposed by Awan and Slavković [2018] for the purpose of designing uniformly most powerful hypothesis tests for Bernoulli data. In the case of $(\epsilon, 0)$-DP, the Tulap distribution coincides with one of the staircase mechanisms [Geng and Viswanath, 2015b]. It is also closely related to the discrete Laplace distribution (also known as the geometric mechanism), which is optimal for a wide range of utility functions in Ghosh et al. [2012].

Proposition 3.1. *Let $\epsilon > 0$. The distribution $\mathrm{Tulap}(0, \exp(-\epsilon), 0)$ is the unique CND for $f_{\epsilon,0}$.*

Proof Sketch. By Lemma 2.3, the only choice in a CND is on $[-1/2, 1/2]$. If the density is non-constant on $[-1/2, 1/2]$, we show that the likelihood ratio is not bounded by e^{ϵ}, violating ϵ-DP. □

Proposition 3.1 is a surprising result in that one may expect a more natural CND than the Tulap distribution, which has a discontinuous density. However, we now know that there are no other CNDs for $(\epsilon, 0)$-DP. In particular, there is no log-concave CND, which is the topic of the next subsection.

3.2 Infinite divisibility and log-concavity

It has been shown in Dong et al. [2022] and Dong et al. [2021] that tradeoff functions built from location family log-concave distributions have very nice properties for f-DP. Log-concave distributions are have a monotone likelihood ratio property which gives a simple closed form expression for the tradeoff function in terms of the cdf of the log-concave distribution. It is easily observed that a tradeoff function with a log-concave CND satsifies a property that we call *infinite divisibility*. We prove that in fact a tradeoff function has a log-concave CND if and only if it is infinitely divisble. Our proof also results in a construction to produce the unique log-concave CND.

A continuous random variable X is *log-concave* if its density can be written as $g_X(x) \propto \exp(C(x))$, where C is a concave function. We call a (symmetric) tradeoff function f *log-concave* if there exists a log-concave CND N for f. Recall that if $N \sim F$ is a CND for f, then $f(\alpha) = F(F^{-1}(\alpha) - 1)$. If N is also log-concave, then $f_t(\alpha) := F(F^{-1}(\alpha) - t)$ is a tradeoff function for every $t \in [0, \infty)$, and the family $\{f_t \mid t \in [0, \infty)\}$ is a monoid satisfying the assumptions of Definition 3.2.

Definition 3.2. A tradeoff function f is *infinitely divisible* if there exists a monoid, under the operation of functional composition, $\{f_t \in \mathscr{F} \mid t \geq 0\}$ containing f such that

1. $f_t \circ f_s = f_{t+s}$ for all $s, t \geq 0$,
2. f_s is nontrivial for all $s > 0$, and
3. $f_s \to f_0 = \mathrm{Id}$ as $s \downarrow 0$.

The discussion above established that log-concave CNDs are infinitely divisible. The key result of this section is that a tradeoff function is log-concave if and only if it is infinitely divisible. We saw that it is easy to construct the infinitely divisible family given a log-concave CND. Surprisingly, we give a construction to derive the log-concave CND from the infinitely divisible family as well. This result shows an intimate relationship between properties of a tradeoff function and the possible CNDs for that tradeoff function. We will see in Section 4.1 that the property of infinite divisibility shows up again in the construction of multivariate CNDs.

Theorem 3.3. *A nontrivial tradeoff function $f \in \mathscr{F}$ is log-concave if and only if it is infinitely divisible. In particular,*

1. *If f is log-concave with log-concave CND $N \sim F$, then $\{f_t \mid t \geq 0\}$ defined by $f_t = F(F^{-1}(\alpha) - t)$ satisfies the assumptions of Definition 3.2.*
2. *Let f be infinitely divisible, with monoid $\{f_t \in \mathscr{F} \mid t \geq 0\}$, as defined in Definition 3.2, such that $f = f_1$. Let F_s be any CND for f_s (such as constructed in Proposition 2.4). Then the following limit exists $F^*(t) := \lim_{s \to 0} F_s(\frac{1}{s}t)$ and $N \sim F^*$ is the unique log-concave CND for f. Furthermore, $F^*(st)$ is the unique log-concave CND for f_s, for all $s > 0$.*

Proof Sketch. It is easy to verify property 1. For property 2, we consider a subsequence $s_n = 1/n!$ and observe that $F_{1/n!}(n!t)$ is a CND for f at every n, but that as n increases, the number of points

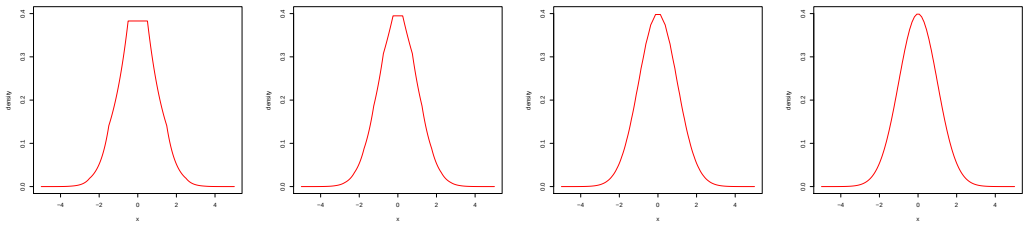

Figure 1: An illustration of Theorem 3.3 when applied to $G_1 = T(N(0,1), N(1,1))$. From left to right, we have the density corresponding to $F_{G_{2^{-n}}}(2^n t)$ for $n = 0, 1, 2, 3$.

at which the CND is uniquely determined also increases, by Lemma 2.3. In the limit, this sequence converges to a unique cdf, which we show has the properties of a log-concave CND. □

Example 3.4. We will illustrate the limit of Theorem 3.3 on G_1. Let $F_{G_{2^{-n}}}$ be the constructed cdf from Proposition 2.4 for $n = 0, 1, 2, 3$. The density functions corresponding to $F_{G_{2^{-n}}}(2^n t)$ are plotted in Figure 1. We see that as n increases, the pdfs approach that of a standard normal, which we know is the log-concave CND for $f = G_1$.

When the construction of Theorem 3.3 is applied to $f_{\epsilon,0}$, the cdf F^* converges to a Laplace cdf. This seems to reflect the fact that under the limit of group privacy, $(\epsilon, 0)$-DP converges to Laplace-DP Dong et al. [2022, Proposition 2.15].

Finally, we illustrate why properties 2 and 3 of Definition 3.2 are necessary for Theorem 3.3

Example 3.5 (Non examples for Theorem 3.3)**.** First consider why it is necessary to have $f_s \to \mathrm{Id}$. Set $f_s(\alpha) = I(\alpha = 1)$ for all $s > 0$. Note that $f_s \circ f_t = f_{s+t}$, but that the construction of Theorem 3.3 results in a point mass at zero, which is not a CND as it is not continuous.

Next, suppose that all of the tradeoff functions are trivial, then $f_s(\alpha) = \alpha$ for all $s > 0$, and $f_s \circ f_t = f_{s+t}$. However, there are no CNDs in this case.

3.3 Piece-wise linear tradeoff functions are generally not infinitely divisible

We showed in Theorem 3.3 that if a tradeoff function is infinitely divisible, then we can construct a log-concave CND. However, it is not always obvious whether a tradeoff function is infinitely divisible or not. We show that in the case of piece-wise linear tradeoff functions, we can upper bound the number of possible divisions in terms of the number of break points. In particular, the piece-wise linear tradeoff functions considered in this section are not infinitely divisible.

We can characterize the piece-wise linear convex functions in terms of the 2nd derivative behavior: A convex function is piece-wise linear if and only if its 2nd derivative is defined everywhere except for finitely many points, and is zero whenever it is defined.

Part 1 of Proposition 3.6 shows that a piece-wise linear tradeoff function f, which satisfies $f(x) = 0$ implies $x = 0$, can be sub-divided only a finite number of times. A consequence of this is that $f_{\epsilon,0}$ and several related tradeoff functions are not infinitely divisible and hence do not have log-concave CNDs. In fact, not only is $f_{\epsilon,0}$ not infinitely divisible, but there is in fact *no division* $f_{\epsilon,0} = f \circ g$ into symmetric tradeoff functions, except where either f or g is the identity!

Proposition 3.6.

1. *Let f be a nontrivial piece-wise linear tradeoff function with $k \geq 1$ breakpoints and such that $f(x) = 0$ implies that $x = 0$. Then there is no tradeoff function g such that $g^{\circ(k+1)} = f$.*

2. *Let $\epsilon > 0$. There does not exist nontrivial symmetric tradeoff functions f_1 and f_2 such that $f_{\epsilon,0} = f_1 \circ f_2$.*

3. *Let f be the tradeoff function obtained by an arbitrary sequence of mechanism compositions, functional compositions, or subsampling (without replacement) of $f_{\epsilon,0}$ (could be different ϵ values for each). Then f is not infinitely divisible and so does not have a log-concave CND.*

Proof Sketch. We show in Lemma A.12 that divisions of a piece-wise linear tradeoff function are themselves piece-wise linear, and that the functional composition of piece-wise linear tradeoff functions increases the number of breakpoints. This then limits the number of divisions a piece-wise linear tradeoff function can have in terms of the number of its breakpoints. □

Example 3.7 ($f_{0,\delta}$ is log-concave)**.** What if $f(x) = 0$ does not imply that $x = 0$? The tradeoff functions $f_{0,\delta}$ fit within this setting, and the results of Proposition 3.6 do *not* apply here. In fact, $f_{0,\delta}$ is infinitely divisible with log-concave CND $U(-1/(2\delta), 1/(2\delta))$. That is $f_{0,\delta} = T(U, U + \delta)$ where $U \sim U(-1/2, 1/2)$. While $f_{\epsilon,\delta}$ for $\delta > 0$ also does not satisfy the assumption that $f_{\epsilon,\delta}(x)$ implies $x = 0$, it is not clear at this time whether $f_{\epsilon,\delta}$ is log-concave or not.

4 Multivariate CNDs

In this section, we generalize the definition of CND to dimensions greater than one. While in the univariate case, *sensitivity* is measured using the absolute distance between two statistic values, in $\mathbb{R}^d$, there are many choices of norms which can be used to measure the sensitivity [Awan and Slavković, 2020]. So, we will specify the sensitivity norm when talking about a multivariate CND. In Definition 4.1 we define a multivariate CND to be a natural generalization of properties 1-4 of Definition 2.2.

Definition 4.1. Let f be a symmetric tradeoff function, and let $\|\cdot\|$ be a norm on $\mathbb{R}^d$. A continuous random vector N with density g is a *canonical noise distribution* (CND) for f, with respect to $\|\cdot\|$, if

1. For all $v \in \mathbb{R}^d$ such that $\|v\| \leq 1$ we have that $T(N, N + v) \geq f$,
2. there exists $\|v^*\| \leq 1$ such that $T(N, N + v^*)(\alpha) = f(\alpha)$ for all $\alpha \in (0, 1)$,
3. for all v^* which satisfy property 2, and all $w \in \mathbb{R}^d$, we have that the likelihood ratio $g(w + tv^* - v^*)/g(w + tv^*)$ is a non-decreasing function of $t \in \mathbb{R}$,
4. N is symmetric about zero: $g(x) = g(-x)$ for all $x \in \mathbb{R}^d$.

When restricted to $d = 1$, Definition 4.1 recovers Definition 2.2. This is clear for properties 1, 2, and 4. Property 3 of Definition 2.2 can be interpreted as requiring that an optimal rejection set for $T(N, N + 1)$ is of the $[x, \infty)$ for some x. By the Neyman Pearson Lemma, we know that this holds if and only if the likelihood ratio $F'(x - 1)/F'(x)$ is non-decreasing in x. We see that when $d = 1$, property 3 of Definition 4.1 is equivalent to property 3 of Definition 2.2. We can interpret Property 3 of Definition 4.1 as enforcing a monotone likelihood ratio in directions parallel to v^*.

4.1 Constructions of multivariate CNDs

Composition gives a simple method to construct a multivariate CND whenever a tradeoff function can be decomposed into the composition of k tradeoff functions:

Proposition 4.2. *Suppose that $f = f_1 \otimes f_2 \otimes \cdots \otimes f_k$ all be nontrivial and symmetric tradeoff functions, and let $F_1, F_2, \ldots, F_k$ be CNDs for $f_1, \ldots, f_k$ respectively. Let $N = (N_1, \ldots, N_k)$ be the random vector where $N_i \sim F_i$ are independent. Then N is a CND for f with respect to $\|\cdot\|_\infty$.*

Interestingly, when a tradeoff function is infinitely divisible and hence has a log-concave CND by Theorem A.8, we can create a multivariate CND with respect to $\|\cdot\|_1$-sensitivity.

Theorem 4.3. *Let f be a nontrivial and symmetric log-concave tradeoff function with log-concave CND F. Let $N = (N_1, \ldots, N_k)$ be the random vector where $N_i \sim F$ are independent. Then N is a (log-concave) CND for f with respect to $\|\cdot\|_1$.*

Proof Sketch. Since the noise added is i.i.d., we can rephrase the tradeoff function as the tensor product of the individual tradeoff functions. We apply Theorem A.2 which lower bounds the tensor product of tradeoff functions with the functional composition. □

Theorem 4.3 was inspired by the i.i.d. Laplace mechanism. In Section 4.5, we show that the i.i.d. Laplace mechanism is a special case of Theorem 4.3 and gives a multivariate CND for Laplace-DP.

Note that Theorem 4.3 results in a log-concave multivariate CND, and if each of $N_1, \ldots, N_k$ are log-concave in Proposition 4.2, then that constructed multivariate CND is log-concave as well. Dong

et al. [2021] showed that log-concave distributions have many nice properties in multivariate settings as well. We leave it to future work to investigate when multivariate log-concave CNDs exist.

4.2 Multivariate CND for GDP

Recall that if $N \sim N(0, I)$ is a d-dimensional Gaussian random vector, and $v \in \mathbb{R}^d$ is any vector, then $T(N, N+v) = T(N(0,1), N(\|v\|_2, 1))$ [Dong et al., 2022, Proposition D.1(5)]. This previous result implies that $N(0, I)$ was a multivariate CND for GDP under $\|\cdot\|_2$-sensitivity. In fact, we show in Proposition 4.4 that for GDP, any multivariate Gaussian is a CND with respect to any norm.

Proposition 4.4. *Let Σ be a $d \times d$ positive definite matrix. Let $v^* \in \operatorname{argmax}_{\|u\| \leq 1} \|\Sigma^{-1/2} u\|_2$. Then $N(0, \Sigma)$ is a d-dimensional CND for $\|\Sigma^{-1/2} v^*\|_2$-GDP with respect to the norm $\|\cdot\|$.*

Remark 4.5. While a multivariate Gaussian is always a multivariate CND for GDP, there is still possibly room for improvement. For Definition 4.1, we only need a single vector to satisfy property 2. However, we could potentially ask that the bound is achieved at all u such that $\|u\| = 1$. Note that if $\|\cdot\|$ is an elliptical norm, then we do get this stronger property for the multivariate Gaussian, when we choose Σ to align with the sensitivity norm.

4.3 Multivariate CND for $(0, \delta)$-DP

First let's review a few facts about $(0, \delta)$-DP, also known as $f_{0,\delta}$-DP. First, note that $U(\frac{-1}{2\delta}, \frac{1}{2\delta})$ is a (log-concave) CND for $f_{0,\delta}$. So, we can write $f_{0,\delta} = T(U, U+\delta)$ where $U \sim U(-1/2, 1/2)$. Because of this, we have that $f_{0,\delta}$ is infinitely divisible, and $f_{0,\delta_1} \circ f_{0,\delta_2} = f_{0,\min\{\delta_1+\delta_2,1\}}$. Furthermore, $f_{0,\delta_1} \otimes f_{0,\delta_2} = f_{0,1-(1-\delta_1)(1-\delta_2)}$, as observed in Dong et al. [2022]. This means that $f_{0,\delta}$ is also infinitely decomposable, a property that we had only seen for GDP before. This decomposability implies, by Proposition 4.2 that we can build a multivariate CND for $f_{0,\delta}$ under $\|\cdot\|_\infty$-sensitivity. In fact, this construction is a multivariate CND for any sensitivity norm.

Proposition 4.6. *Let $0 < \delta \leq 1$, $d \geq 1$, and $\|\cdot\|$ be a norm on $\mathbb{R}^d$. Call $v^* \in \underset{\|v\| \leq 1}{\arg\min} \prod_{i=1}^d (1 - \delta|v_i|)$ and $A = \prod_{i=1}^d (1 - \delta|v_i^*|)$. Then $U(\frac{-1}{2\delta}, \frac{1}{2\delta})^n$ is a CND for $f_{0,1-A}$ under $\|\cdot\|$-sensitivity. In the special case of $\|\cdot\| = \|\cdot\|_\infty$, this simplifies to $A = (1-\delta)^d$.*

4.4 Multivariate CND for $f_{\epsilon,\delta}$ when $\delta > 0$

Let $\epsilon > 0$ and $\delta \in (0, 1]$. Recall that $f_{\epsilon,\delta} = f_{\epsilon,0} \otimes f_{0,\delta}$ [Dong et al., 2022]. Since $f_{0,\delta}$ is infinitely decomposable, we can write $f_{\epsilon,\delta} = f_{\epsilon,0} \otimes f_{0,\delta_1} \otimes \cdots \otimes f_{0,\delta_k}$ where $\delta = \prod_{i=1}^k (1 - \delta_i)$. By Proposition 4.2 we construct a multivariate CND for $f_{\epsilon,\delta}$ with respect to $\|\cdot\|_\infty$-sensitivity by using $\mathrm{Tulap}(0, \exp(-\epsilon), 0)$ in one coordinate, and the uniform distributions $U(\frac{-1}{2\delta_i}, \frac{1}{2\delta_i})$ in the other k coordinates.

4.5 Two multivariate CNDs for Laplace-DP

Many mechanisms designed to satisfy $(\epsilon, 0)$-DP actually satisfy the stronger privacy guarantee of Laplace-DP. In particular, variations on the Laplace mechanism are very common additive mechanisms used to achieve $(\epsilon, 0)$-DP. In this section, we show that two multivariate versions of the Laplace mechanism, the ℓ_1 and ℓ_∞ mechanisms, are multivariate CNDs for Laplace-DP.

The Laplace distribution, denoted $\mathrm{Laplace}(m, s)$ is a distribution on $\mathbb{R}$ with density $\frac{1}{2s} \exp(\frac{-1}{s}|x - m|)$. We say a mechanism satisfies ϵ-Laplace-DP if it satisfies L_ϵ-DP, where $L_\epsilon := T(N, N+\epsilon)$ and $N \sim \mathrm{Laplace}(0, 1)$. It is easily seen that $\mathrm{Laplace}(0, 1/\epsilon)$ is a log-concave CND for ϵ-Laplace-DP.

i.i.d. Laplace Mechanism The i.i.d. Laplace mechanism is defined as follows: Let $\epsilon > 0$ be given. If $T : \mathscr{X} \to \mathbb{R}^k$ has $\|\cdot\|_1$-sensitivity of Δ, then the i.i.d. Laplace mechanism releases $T(X) + \Delta N$, where $N = (N_1, \ldots, N_k)$ is the random vector with i.i.d. entries $N_i \sim \mathrm{Laplace}(0, 1/\epsilon)$. It is well known that the i.i.d. Laplace mechanism satisfies $f_{\epsilon,0}$-DP [Dwork et al., 2014, Theorem 3.6]. Since N_1 is a log-concave CND for L_ϵ, Theorem 4.3 shows that N is a CND for L_ϵ, with respect to $\|\cdot\|_1$-sensitivity. As $L_\epsilon \geq f_{\epsilon,0}$ and $L_\epsilon(\alpha) > f_{\epsilon,0}(\alpha)$ for some values of α, we can more precisely capture the privacy cost of the i.i.d. Laplace mechanism using tradeoff functions rather than ϵ-DP.

ℓ_∞-Mechanism The ℓ_∞-mechanism, proposed in Steinke and Ullman [2016] is a special case of the K-norm mechanisms [Hardt and Talwar, 2010], with density proportional to $\exp(-\epsilon\|x\|_\infty)$. Steinke and Ullman [2016] showed that the ℓ_∞ mechanism can improve the sample complexity of answering multiple queries, when accuracy is measured by ℓ_∞-norm. Awan and Slavković [2020] showed that the ℓ_∞ mechanism is near optimal in certain applications of private linear and logistic regression. It is well known that when using ℓ_∞-sensitivity, the ℓ_∞-mechanism satisfies ϵ-DP. In this section, we show that the ℓ_∞-mech is a CND for L_ϵ, with respect to ℓ_∞-sensitivity.

Proposition 4.7. *Let $\epsilon > 0$, and $d \geq 1$. Let X be a d-dimensional random vector with density $g(x) = \frac{\exp(-\epsilon\|x\|_\infty)}{d!(2/\epsilon)^d}$. Then X is a CND for the tradeoff function L_ϵ with respect to $\|\cdot\|_\infty$.*

Proof Sketch. First we show that with the shift of $v^* = (1, 1, \ldots, 1)$, the privacy loss random variable coincides with that of L_ϵ. Then, we show that $v^* = (1, 1, \ldots, 1)^\top$ is the worst case of any shift v to minimize the tradeoff functions. To deal with the case that some of the entries of v are zero, we establish a convergence theorem for tradeoff functions in Theorem A.8 of the Appendix. □

4.6 No multivariate CND for $f_{\epsilon,0}$

By the construction of Proposition 2.4, we know that a one-dimensional CND exists for any nontrivial tradeoff function. It turns out that the same cannot be said for the multivariate setting. In Theorem 4.8, we show that there is *no multivariate CND for $f_{\epsilon,0}$ with respect to any norm*. In fact, we prove the stronger result that it is not even possible to satisfy properties 1 and 2 of Definition 4.1

Theorem 4.8. *Let $d \geq 2$ and let $\|\cdot\|$ be any norm on $\mathbb{R}^d$. Then for any $\epsilon > 0$, there is no random vector satisfying properties 1 and 2 of Definition 4.1 for $f_{\epsilon,0}$ with respect to the norm $\|\cdot\|$. In particular, there is no multivariate CND for $f_{\epsilon,0}$.*

Proof Sketch. Suppose to the contrary, then $(\epsilon, 0)$-DP imposes strict bounds on the likelihood ratio of the distribution. These bounds allow us to find an arbitrarily long sequence of points, sufficiently far apart, where the density is bounded below. This ultimately shows that the density is not integrable. □

Combining Theorem 4.8 with Proposition 4.2, we infer in Corollary 4.9 that $f_{\epsilon,0}$ cannot be written as the tensor product of any two nontrivial tradeoff functions. This means that if we want to design two independent mechanisms such that the joint release exactly satisfies $(\epsilon, 0)$-DP, then one of the mechanisms must be perfectly private.

Corollary 4.9. *Let $\epsilon > 0$ be given. There does not exist nontrivial symmetric tradeoff functions f_1 and f_2 such that $f_{\epsilon,0} = f_1 \otimes f_2$.*

Remark 4.10. Theorem 4.8 along with Theorem 4.3 gives an alternative argument that $f_{\epsilon,0}$ is not log-concave/infinitely decomposable.

5 Discussion

Motivated by the goals of constructing log-concave CNDs and multivariate CNDs, we found some fundamental connections between these constructions and the operations of mechanism composition and functional composition of the tradeoff functions. Surprisingly, the constructions for both log-concave and multivariate CNDs relied on whether a tradeoff function could be decomposed either according to functional composition, or according to mechanism composition. An interesting result of our work was that for $(\epsilon, 0)$-DP there is a unique 1-dimensional CND and no multidimensional CNDs, which implies that $f_{\epsilon,0}$ can neither be decomposed according to functional composition or mechanism composition. This highlights the limitations of pure-DP as a privacy definition. On the other hand, Gaussian-DP, Laplace-DP, and $(0, \delta)$-DP were seen to have much better properties.

While the framework of GDP and related notions (e.g., zero-concentrated DP) have many desirable properties, including those developed in this paper, there are still many reasons why one may be interested in other DP frameworks. In some applications, having a stronger notion of DP is needed to protect events with small probability, such as pure-DP or Laplace-DP; in this case, our work shows that Laplace-DP is a much better behaved notion of privacy than pure-DP. One may also propose other alternative DP definitions based on a family of tradeoff functions, and our research gives some

fundamental insights on what properties that family must have in order for log-concave or multivariate CNDs to be constructed.

We showed that a multivariate extension of CND can capture the same properties as in the 1-dimensional case. Awan and Vadhan [2021] showed that in one dimension, CNDs can be used to obtain DP hypothesis tests with optimal properties. An open question is whether our definition of a multivariate CND has any connections to optimal hypothesis testing.

Most of the constructions of multivariate CNDs presented in this paper are product distributions. Even the multivariate CNDs for GDP are a linear transformation of i.i.d. random variables. The ℓ_∞-mechanism is the exception, providing a truly nontrivial CND for Laplace-DP. It is worth exploring whether there are general techniques to produce nontrivial multivariate CNDs like the ℓ_∞-mechansism, as well as exploring the merits of such CNDs.

While many of the multivariate CNDs constructed for tradeoff functions, only held for specific sensitivity norms, a more general question would be on the existence and construction of multivariate CNDs for an arbitrary tradeoff/norm pair.

Acknowledgments

This work was supported in part by NSF SES 2150615, awarded to Purdue University.

References

Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 308–318, 2016.

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 Aleksandra Slavković. Structure and sensitivity in differential privacy: Comparing k-norm mechanisms. *Journal of the American Statistical Association*, pages 1–20, 2020.

Jordan Awan and Salil Vadhan. Canonical noise distributions and private hypothesis tests. *arXiv preprint arXiv:2108.04303*, 2021.

Jordan Awan, Ana Kenney, Matthew Reimherr, and Aleksandra Slavković. Benefits and pitfalls of the exponential mechanism with applications to Hilbert spaces and functional pca. In *International Conference on Machine Learning*, pages 374–384. PMLR, 2019.

Borja Balle, Gilles Barthe, and Marco Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in Neural Information Processing Systems*, 31, 2018.

Borja Balle, Gilles Barthe, and Marco Gaboardi. Privacy profiles and amplification by subsampling. *Journal of Privacy and Confidentiality*, 10(1), 2020.

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

Jinshuo Dong, Weijie Su, and Linjun Zhang. A central limit theorem for differentially private query answering. *Advances in Neural Information Processing Systems*, 34, 2021.

Jinshuo Dong, Aaron Roth, and Weijie Su. Gaussian differential privacy. *Journal of the Royal Statistical Society: Series B*, 08 2022. doi: 10.1111/rssb.12455. arXiv:1905.02383.

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

Cynthia Dwork, Moni Naor, Omer Reingold, Guy N Rothblum, and Salil Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pages 381–390, 2009.

Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

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.

Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 705–714, 2010.

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.

David M Sommer, Sebastian Meiser, and Esfandiar Mohammadi. Privacy loss classes: The central limit theorem in differential privacy. *Proceedings on privacy enhancing technologies*, 2019(2): 245–269, 2019.

Thomas Steinke and Jonathan Ullman. Between pure and approximate differential privacy. *Journal of Privacy and Confidentiality*, 7(2):3–22, 2016.

Jun Zhang, Zhenjie Zhang, Xiaokui Xiao, Yin Yang, and Marianne Winslett. Functional mechanism: Regression analysis under differential privacy. *Proceedings of the VLDB Endowment*, 5(11), 2012.

Yuqing Zhu and Yu-Xiang Wang. Improving sparse vector technique with renyi differential privacy. *Advances in Neural Information Processing Systems*, 33:20249–20258, 2020.

Yuqing Zhu, Jinshuo Dong, and Yu-Xiang Wang. Optimal accounting of differential privacy via characteristic function. In *International Conference on Artificial Intelligence and Statistics*, pages 4782–4817. PMLR, 2022.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] In Section 5 we discuss directions for future work.
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] See Section A.1 in the Appendix
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [Yes]
   (b) Did you include complete proofs of all theoretical results? [Yes] Included in the Appendix
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [N/A] No experiments were run, as the results are justified by theory.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]

(c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...

   (a) If your work uses existing assets, did you cite the creators? [N/A]

   (b) Did you mention the license of the assets? [N/A]

   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A]

   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]

   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A]

5. If you used crowdsourcing or conducted research with human subjects...

   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]

   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]

   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]

A Appendix

A.1 Broader impacts

Privacy is an important societal problem, and there is a natural tradeoff between the privacy afforded to the individuals of the dataset, and the utility of the published result. One may be concerned that differential privacy techniques reduce the utility of the results too much, in exchange for the privacy protection. In our work, by providing a better understanding of differential privacy, and by developing new mechanisms to achieve differential privacy, we make it possible to achieve higher utility at the same privacy cost; or alternatively, we can maintain the same utility while giving a stronger privacy protection. In our view, optimizing the privacy-utility tradeoff is universally beneficial to society, and we do not foresee any negative societal impacts of this work.

A.2 Relations between functional composition and tensor product

Both the functional composition and the tensor product of tradeoff functions are essential concepts in our constructions of CNDs. In the remainder of this section, we establish some new relations between the two operations, which provide an interesting insight into the connection between group privacy and composition. First, we recall a lemma from Dong et al. [2022]:

Lemma A.1 (Lemma A.5: Dong et al. [2022])**.** *Suppose that $T(P, Q) \geq f$ and $T(Q, R) \geq g$. Then $T(P, R) \geq g \circ f$.*

Lemma A.2. *Let f and g be any two symmetric tradeoff functions. Then $f \otimes g \geq f \circ g$.*

Proof. First note that if either f or g is equal to Id, then the result is trivial. Now, suppose that both f and g are nontrivial, and let $N_1 \sim F$ and $N_2 \sim G$ be independent, where F is a CND for f and G is a CND for g.

By definition of the tensor product of tradeoff functions [Dong et al., 2022, Definition 3.1], we have that

$$T\left(\begin{pmatrix} 0 + N_1 \\ 0 + N_2 \end{pmatrix}, \begin{pmatrix} 1 + N_1 \\ 1 + N_2 \end{pmatrix}\right) = f \otimes g, \tag{1}$$

since $T(0 + N_1, 1 + N_1) = f$ and $T(0 + N_2, 1 + N_2) = g$, by definition of CND.

It is also true that

$$T\left(\begin{pmatrix}0+N_1\\0+N_2\end{pmatrix},\begin{pmatrix}1+N_1\\0+N_2\end{pmatrix}\right)=f,$$
$$T\left(\begin{pmatrix}1+N_1\\0+N_2\end{pmatrix},\begin{pmatrix}1+N_1\\1+N_2\end{pmatrix}\right)=g.$$

Then by Lemma A.1,

$$T\left(\begin{pmatrix}0+N_1\\0+N_2\end{pmatrix},\begin{pmatrix}1+N_1\\1+N_2\end{pmatrix}\right)\geq g\circ f=f\circ g, \tag{2}$$

where the last equality follows since f and g are symmetric, using [Dong et al., 2022, Lemma A.4]. Comparing Equations (1) and (2), we have that $f\otimes g\geq f\circ g$. □

Remark A.3. As a special case of Lemma A.2, we have that $f\otimes f\geq f\circ g$, which has an interesting interpretation: Suppose two situations: 1) your data is present once in two databases, and an f-DP mechanism is applied to each database once. This gives $f\otimes f$-DP cumulative privacy cost to you. 2) your data is present twice in one database, and an f-DP mechanism is applied once to the database. Since your data is present twice, by group privacy the incurred privacy cost to you is $f\circ f$-DP. Lemma A.2 says that you would prefer to be in the two separate databases. The intuition behind this can be understood as follows: in the second scenario, the privacy expert could choose to split the database into two: each one containing a copy of your data, and apply an f-DP mechanism to both. The nominal privacy cost of this would be f-DP (considering groups of size 1), as changing one entry affects only one of the two calculations. However, for groups of size two, the privacy cost is $f\circ f$-DP. This shows that all of the mechanisms in scenario 1 could also be applied to scenario 2, but in general there are mechanisms in scenario two that are not possible in scenario 1 (since in scenario 1, the databases cannot be merged).

Before we move on, we give a Lemma, extending Awan and Vadhan [2021, Lemma E.8] to arbitrary k. Lemma A.5 shows that given a CND N for f, we can easily produce a CND for $f^{\circ k}$ by rescaling N by $\frac{1}{k}$. To establish Lemma A.5, we need another technical lemma, which appeared within the proof Awan and Vadhan [2021, Lemma E.12]. We say that a cdf F is *invertible* at t if $F^{-1}(F(t))=t$.

Lemma A.4 (Awan and Vadhan [2021])**.** *Let f be a nontrivial symmetric tradeoff function, and let F be a CND for f. Call $M:=\inf\{t\mid 0<F(t)\}$. Then if $M<-1/2$ and $\alpha>1-f(1)$, then F is invertible at $F^{-1}(\alpha)-1$.*

Proof. Let $\alpha>1-f(1)$, or equivalently $1-\alpha<f(1)$. Note that F is invertible at $F^{-1}(\alpha)$, and by Awan and Vadhan [2021, Lemma E.3] F is also invertible at $F^{-1}(\alpha)-1$ unless $F^{-1}(\alpha)-1<M$. So, we need to show that $F^{-1}(\alpha)\geq M+1$:

$$\begin{aligned}
M+1&=\inf\{t+1\mid 0<F(t)\} &(3)\\
&=\inf\{t\mid 0<F(t-1)\} &(4)\\
&=\inf\{t\mid 1>1-F(t-1)\} &(5)\\
&=\inf\{t\mid 1-f(1)<1-f(1-F(t-1))\ \&\ 0<F(t-1)\} &(6)\\
&=\inf\{t\mid 1-f(1)<F(t)\ \&\ 0<F(t-1)\}, &(7)
\end{aligned}$$

where (6) uses the fact that $1-f$ is strictly decreasing at 1; (7) uses the fact that $0<F(t-1)$ to apply the recursion of Lemma 2.3. Now, suppose that $F(t-1)=0$: then $t-1\leq M$ and because $M<-1/2$, $F(t)<1$. So, $0=F(t-1)$ implies that $0=f(F(t))$. But this in turn implies that $F(t)\leq 1-f(1)$. We see that $1-f(1)<F(t)$ implies that $0<F(t-1)$. So,

$$\begin{aligned}
M+1&=\inf\{t\mid 1-f(1)<F(t)\} &(8)\\
&\leq\inf\{t\mid \alpha\leq F(t)\} &(9)\\
&=F^{-1}(\alpha), &(10)
\end{aligned}$$

where (9) uses the fact that $\alpha>1-f(1)$. We see that $F^{-1}(\alpha)-1\geq M$ and conclude that F is invertible at $F^{-1}(\alpha)-1$. □

Lemma A.5. *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 $k \in \mathbb{N}^+$.*

Proof. For any k, denote $F_k(x) = F(kx)$ and $F_k^{-1}(x) = \frac{1}{k}F^{-1}(x)$, where F_k^{-1} is the quantile function of F_k. Symmetry and continuity of F_k are clear.

For induction, assume that for some $k > 1$, we have that F_{k-1} is a CND for $f^{\circ(k-1)}$. In particular, we have that

$$\begin{aligned} f^{\circ(k-1)} &= F_{k-1}(F_{k-1}^{-1}(\alpha) - 1) \\ &= F[(k-1)\{[1/(k-1)]F^{-1}(\alpha) - 1\}] \\ &= F(F^{-1}(\alpha) - (k-1)). \end{aligned}$$

Let $M := \inf\{t \mid 0 < F_{k-1}(t)\}$. By symmetry of F_{k-1}, we know that $M \leq 0$. If $M \geq -1/2$, then we have that $f^{\circ k}(\alpha) \leq f^{\circ(k-1)}(\alpha) = F_{k-1}(F_{k-1}^{-1}(\alpha) - 1) = 0$ for all $\alpha \in (0,1)$; we also have $F_k(F_k^{-1}(\alpha) - 1) = F(F^{-1}(\alpha) - k) = 0 = f^{\circ k}(\alpha)$. Furthermore, $T(F_k(\cdot), F_k(\cdot - 1)) = T(F(\cdot), F(\cdot - k)) = 0 = g$, since $F(\cdot)$ and $F(\cdot - k)$ have disjoint support. Finally, note that $T(F_k(\cdot), F_k(\cdot - m)) \geq 0 = T(F_k(\cdot), F_k(\cdot - 1))$, since 0 is a trivial lower bound for any tradeoff function. We conclude that when $M \geq -1/2$, F_k is a CND for $f^{\circ k}$.

Now suppose that $M < -1/2$ and let $\alpha \in (0,1)$. If $\alpha \leq 1 - f^{\circ(k-1)}(1)$, then $f^{\circ k}(\alpha) = f(f^{\circ(k-1)}(\alpha)) = f(0) = 0$ and $F(F^{-1}(1-\alpha) - k) \leq F(F^{-1}(1-\alpha) - (k-1)) = f^{\circ(k-1)}(\alpha) = 0 = f^{\circ k}(\alpha)$ because F is increasing. We see that $f^{\circ k} = F_k(F_k^{-1}(\alpha) - 1)$ in this case. Now assume that $\alpha > 1 - f(1)$. Then by Lemma A.4, we have that F_{k-1} is invertible at $F_{k-1}^{-1}(\alpha) - 1$. Then,

$$\begin{aligned} f^{\circ k} &= f \circ F_{k-1}(F_{k-1}^{-1}(\alpha) - 1) && (11) \\ &= F(F^{-1}[F_{k-1}(F_{k-1}^{-1}(\alpha) - 1)] - 1) && (12) \\ &= F((k-1)F_{k-1}^{-1}[F_{k-1}(F_{k-1}^{-1}(\alpha) - 1)] - 1) && (13) \\ &= F((k-1)(F_{k-1}^{-1}(\alpha) - 1) - 1) && (14) \\ &= F(F^{-1}(\alpha) - k) && (15) \\ &= F_k(F_k^{-1}(\alpha) - 1), && (16) \end{aligned}$$

where in (14), we used the fact that F_{k-1} is invertible at $F_{k-1}^{-1}(\alpha) - 1$.

We have shown that $f^{\circ k} = F_k(F_k^{-1}(\alpha) - 1)$. Since $F_k(F_k^{-1}(\alpha) - 1)$ represents the type II error of the (potentially suboptimal) test, which rejects when the observed random variable is above a certain threshold, we have that $T(N_k, N_k + 1) = T(N, N + k) \leq f^{\circ k}$, where $N \sim F$ and $N_k \sim F_k$. To verify properties 2 and 3 of Definition 2.2, it remains to show that $T(N_k, N_k + 1) \geq f^{\circ k}$. Note that $T(N, N + (k-1)) = f^{\circ(k-1)}$, and $T(N + (k-1), N + k) = T(N, N + 1) = f$. By Lemma A.1, we have that $T(N_k, N_k + 1) = T(N, N + k) \geq f^{\circ(k-1)} \circ f = f^{\circ k}$, which completes the argument for parts 2 and 3 of Definition 2.2.

For property 1 of Definition 2.2, let $m \in [0,1]$. As before, we use the notation $N \sim F$, $N_k \sim F_k$ and $N_{k-1} \sim F_{k-1}$. We will show that $T(N_k, N_k + m) = T(N, N + km) \geq f^{\circ k}$. If $m \leq \frac{k-1}{k}$, then $m^* = \frac{km}{k-1} \in [0,1]$. In this case,

$$\begin{aligned} T(N, N + km) &= T\left(N, N + (k-1)\frac{km}{k-1}\right) \\ &= T(N_{k-1}, N_{k-1} + m^*) \\ &\geq T(N_{k-1}, N_{k-1} + 1) \\ &= f^{\circ(k-1)} \\ &\geq f^{\circ k}, \end{aligned}$$

where we used the fact that F_{k-1} is a CND for $f^{\circ(k-1)}$ and that $f^{\circ(k-1)} \geq f^{\circ k}$. If $m \geq \frac{k-1}{k}$, then $m^* = km - (k-1) \in [0,1]$. Then

$$\begin{aligned} T(N, N+km) &= T(N, N+(k-1)+m^*) && (17)\\ &\geq T(N, N+(k-1)) \circ T(N, N+m^*) && (18)\\ &= f^{\circ(k-1)} \circ T(N, N+m^*) && (19)\\ &\geq f^{\circ(k-1)} \circ f && (20)\\ &= f^{\circ k}, && (21) \end{aligned}$$

where for (18) we use Lemma A.1 and the fact that $T(N+(k-1), N+(k-1)+m^*) = T(N, N+m^*)$, and for (20), we use the inductive hypothesis that F_{k-1} is a CND for $f^{\circ(k-1)}$. □

Example A.6 (Composition and Group Privacy do not Commute)**.** It is an interesting question whether the following property holds: $(f \otimes g)^{\circ k} = f^{\circ k} \otimes g^{\circ k}$. This is true for GDP:

$$(G_{\mu_1} \otimes G_{\mu_2})^{\circ k} = G_{k\|\binom{\mu_1}{\mu_2}\|} = G_{\|\binom{k\mu_1}{k\mu_2}\|} = G_{\mu_1}^{\circ k} \otimes G_{\mu_2}^{\circ k}.$$

However, by studying $(0, \delta)$-DP, we see that this property does not hold in general – even for log-concave tradeoff functions. We compute that

$$(f_{0,\delta_1} \otimes f_{0,\delta_2})^{\circ k} = f_{0,1-(1-\delta_1)(1-\delta_2)}^{\circ k} = f_{0,\min\{1,k(1-(1-\delta_1)(1-\delta_2))\}},$$

whereas $f_{0,\delta_1}^{\circ k} \otimes f_{0,\delta_2}^{\circ k} = f_{0,\min\{1,k\delta_1\}} \otimes f_{0,\min\{1,k\delta_2\}} = f_{0,1-(1-\min\{1,k\delta_1\})(1-\min\{1,k\delta_2\})}$. plugging in $k = 2$ and $\delta_1 = \delta_2 = .1$, we get that the first expression yields .38, whereas the second gives .36. Interestingly, it seems that accounting for group privacy *first*, before applying composition gives the tighter privacy analysis. This is confirmed by the inequality in Proposition A.7.

Proposition A.7. *Let f and g be tradeoff functions. Then $(f \otimes g)^{\circ k} \leq f^{\circ k} \otimes g^{\circ k}$.*

Proof. We know that , $f = T(X, X+1)$ and $g = T(Y, Y+1)$, where X is a CND for f and Y is a CND for g.

$$\begin{aligned} (f \otimes g)^{\circ k} &= \big(T[X, X+1] \otimes T[Y, Y+1]\big)^{\circ k}\\ &= \big(T[(X,Y), (X+1, Y+1)]\big)^{\circ k}\\ &\leq T[(X,Y), (X+k, Y+k)] \qquad \text{by Lemma A.1 and Lemma A.5}\\ &= T[X, X+k] \otimes T[Y, Y+k]\\ &= T[X/k, X/k+1] \otimes T[Y/k, Y/k+1]\\ &= f^{\circ k} \otimes g^{\circ k}, \end{aligned}$$

since X/k and Y/k are CNDs for $f^{\circ k}$ and $g^{\circ k}$ respectively, by Lemma A.5. □

A.3 A limit theorem for tradeoff functions

Below, we introduce a limit theorem for tradeoff functions, which can be used to show a mechanism satisfies f-DP in terms of certain limits.

Theorem A.8. *Let $P_n \overset{\text{TV}}{\to} P$ and $Q_n \overset{\text{TV}}{\to} Q$ be two sequences of distributions, which converge in total variation. Then $T(P_n, Q_n) \to T(P, Q)$ uniformly.*

Proof. By Dong et al. [2022, Lemma A.7], it suffices to prove point-wise convergence. First we will establish $T(P, Q)$ as an asymptotic lower bound on $T(P_n, Q_n)$. By Lemma A.1, we have that

$$T(P_n, Q_n) \geq T(Q, Q_n) \circ T(P, Q) \circ T(P_n, P).$$

Since $P_n \overset{\text{TV}}{\to} P$ and $Q_n \overset{\text{TV}}{\to} Q$, we have that $T(P_n, P)(\alpha) \geq [(\alpha - TV(P_n, P))]_0^1$ and $T(Q_n, Q)(\alpha) \geq [(\alpha - TV(Q_n, Q))]_0^1$, where $[x]_a^b := \max\{\min\{x, b\}, a\}$ is the clamping function. Since all tradeoff functions are increasing, the following inequality holds:

$$T(P_n, Q_n) \geq [(\alpha - \text{TV}(Q_n, Q))]_0^1 \circ T(P, Q) \circ [(\alpha - \text{TV}(P_n, P))]_0^1 \to T(P, Q),$$

and the limit holds since $\mathrm{TV}(P_n, P) \to \mathrm{Id}$, $\mathrm{TV}(Q_n, Q) \to \mathrm{Id}$, and tradeoff functions are continuous.

Next, we show that $T(P, Q)$ is an asymptotic upper bound for $T(P_n, Q_n)$. It suffices to check for $\alpha \in (0, 1)$, since tradeoff functions are continuous. Let $\alpha^* \in (0, 1)$ be given. Let ϕ be an optimal test for $T(P, Q)$ such that $\mathbb{E}_P \phi = \alpha^*$ and $\mathbb{E}_Q \phi = 1 - f(1 - \alpha^*)$. Note that if $U \sim U(0, 1)$, we can write

$$\mathbb{E}_P \phi = \mathbb{E}_{X \sim P, U} I(U \leq \phi(X)) = \mathbb{E}_U P_{X \sim P}(U \leq \phi(X)|U) = \mathbb{E}_U P(\phi^{-1}([U, 1])|U),$$

which will allow us to apply the total variation convergence. Call $\alpha_n = \mathbb{E}_{P_n} \phi$ for all n. Notice that $\alpha_n \to \alpha^*$, since

$$\begin{aligned}
|\alpha_n - \alpha^*| &= |\mathbb{E}_{P_n} \phi - \mathbb{E}_P \phi| \\
&= \Big| \mathbb{E}_U P_n(\phi^{-1}([U, 1]) \mid U) - \mathbb{E}_U P(\phi^{-1}([U, 1]) \mid U) \Big| \\
&\leq \mathbb{E}_U \Big| P_n(\phi^{-1}([U, 1]) \mid U) - P(\phi^{-1}([U, 1]) \mid U) \Big| \\
&\leq \mathbb{E}_U \mathrm{TV}(P_n, P) \\
&\to 0,
\end{aligned}$$

as $P_n \overset{\mathrm{TV}}{\to} P$. Similarly, we have that $|\mathbb{E}_{Q_n} \phi - \mathbb{E}_Q \phi| \to 0$, implying that $\mathbb{E}_{Q_n} \phi \to 1 - f(1 - \alpha^*)$. Then,

$$\begin{aligned}
T(P_n, Q_n)(\alpha_n) &\leq 1 - \mathbb{E}_{Q_n} \phi \\
&\to f(1 - \alpha^*) \\
&= T(P, Q)(\alpha^*).
\end{aligned}$$

However, we actually want to show that $T(P_n, Q_n)(\alpha^*)$ is asymptotically upper bounded by $T(P, Q)(\alpha)$. Luckily, $T(P_n, Q_n)(\alpha_n)$ and $T(P_n, Q_n)(\alpha^*)$ are close for large n, since tradeoff functions are "locally Lipschitz." We explain as follows: Since $\alpha_n \to \alpha^*$, let N be such that for all $n \geq N$, $\alpha_n \in \left(0, \alpha^* + \frac{1-\alpha^*}{2}\right)$. On the interval $(0, \alpha^* + \frac{1-\alpha^*}{2})$, we claim that $T(P_n, Q_n)$ is $\frac{2}{1-\alpha^*}$-Lipschitz. This is because the derivative (or subderivative) of a convex function is increasing, and the tangent lines of a convex function are always a lower bound. In the worst case, using the points $(\alpha^* + (1 - \alpha^*)/2, 0)$ and $(1, 1)$, the slope at $\alpha^* + (1 - \alpha^*)/2$ is at most $\frac{1-0}{1-(\alpha^*+(1-\alpha^*)/2} = \frac{2}{1-\alpha^*}$. Now that we have established that $T(P_n, Q_n)$ is $\frac{2}{1-\alpha^*}$-Lipschitz on $(0, \alpha^* - (1 - \alpha^*)/2)$, we have that for all $n \geq 0$,

$$\Big| T(P_n, Q_n)(\alpha_n) - T(P_n, Q_n)(\alpha^*) \Big| \leq \frac{2}{1 - \alpha^*} |\alpha_n - \alpha| \to 0.$$

We conclude that $T(P_n, Q_n)(\alpha^*)$ is asymptotically upper bounded by $T(P, Q)(\alpha^*)$ for all $\alpha* \in (0, 1)$. Combining the asymptotic lower and upper bounds, we have that $T(P_n, Q_n) \to T(P, Q)$. $\square$

Two immediate corollaries of the above theorem are as follows. The first, generally states that if we establish a lower bound on $T(P_n, Q_n)$ for all n, and $P_n \overset{\mathrm{TV}}{\to} P$ and $Q_n \overset{\mathrm{TV}}{\to} Q$, then the lower bound applies to $T(P, Q)$ as well. This could be generalized to a sequence of lower bounds $f_n \to f$ as well.

Corollary A.9. *Let $P_n \overset{\mathrm{TV}}{\to} P$ and $Q_n \overset{\mathrm{TV}}{\to} Q$ be two sequences of distributions such that $T(P_n, Q_n) \geq f$ for all n. Then $T(P, Q) \geq f$.*

Corollary A.10 shows that the limit of an f-DP mechanism satisfies f-DP (could also replace each f with $f_n \to f$). This is similar to the limit result of Kifer et al. [2012], but is phrased in terms of convergence in total variation rather than almost sure convergence.

Corollary A.10. *Let M_n be a sequence of mechanisms satisfying f-DP (i.e., $T(M_n(D), M_n(D')) \geq f$ for all adjacent D and D'), and suppose that $M_n(D) \overset{\mathrm{TV}}{\to} M(D)$ for all D. Then M satisfies f-DP: $T(M(D), M(D')) \geq f$.*

A.4 Proofs and technical lemmas for the main paper

For any measurable set A, let $\lambda(A)$ denote the Lebesgue measure of A.

Lemma A.11. *Let $A, B \subset [-1/2, 1/2]$ be disjoint sets with positive Lebesgue measure such that $A \cup B = [-1/2, 1/2]$. Then there exists a shift $\omega \in (-1, 1)$ such that $(B + \omega) \cap A$ has positive Lebesgue measure.*

Proof. Suppose to the contrary that for all $\omega \in (-1, 1)$, $\lambda((B + \omega) \cap A) = 0$. This implies that

$$
\begin{aligned}
0 &= \int_{-1}^{1} \lambda((B+\omega) \cap A)\, d\omega && (22)\\
&= \int_{-1}^{1} \int_{-1/2}^{1/2} I(x \in (B+\omega) \cap A)\, dx\, d\omega && (23)\\
&= \int_{-1}^{1} \int_{A} I(x \in B+\omega)\, dx\, d\omega && (24)\\
\text{(Tonelli's Theorem)} &= \int_{A} \int_{-1}^{1} I(x \in B+\omega)\, d\omega\, dx && (25)\\
&= \int_{A} \lambda(x - B)\, dx && (26)\\
&= \int_{A} \lambda(B)\, dx && (27)\\
&= \lambda(A)\lambda(B), && (28)
\end{aligned}
$$

where we used Tonelli's Theorem in (25) to change the order of integration, in (26) we used the fact $x - B \subset [-1, 1]$ since both x and B lie in $[-1/2, 1/2]$, and in (27) we used the fact that Lebesgue measure is translation invariant. We see that either $\lambda(A) = 0$ or $\lambda(B) = 0$, giving a contradiction. □

Proposition 3.1. *Let $\epsilon > 0$. The distribution $\mathrm{Tulap}(0, \exp(-\epsilon), 0)$ is the unique CND for $f_{\epsilon,0}$.*

Proof. Let g be the density of an arbitrary CND for $f_{\epsilon,0}$, and let G denote its cdf function. Since g is the density of a symmetric random variable centered at zero, $g(x) = g(-x)$ for all $x \in \mathbb{R}$. By Awan and Vadhan [2021, Proposition 3.7], we have that $G(x + 1) = 1 - f_{\epsilon,0}(1 - G(x))$, which implies that $g(x + 1) = f'_{\epsilon,0}(1 - G(x))g(x)$. For $x > 0$, we have that $G(x) \geq 1/2 > c$, where c satisfies $f_{\epsilon,0}(1 - c) = c$. Recall that $f_{\epsilon,0}(\alpha) = \alpha e^{-\epsilon}$ for all $\alpha \leq c$. Then $f'_{\epsilon,0}(1 - G(x)) = e^{-\epsilon}$ for $x \geq 0$. We see that we can write $g(u + k) = e^{-|k|\epsilon} g(u)$ for $u \in [-1/2, 1/2]$ and $k \in \mathbb{Z}$.

We see that so far, g has the freedom to choose its values in $[-1/2, 1/2]$ and then all other values are determined by the above recurrence. Note that for the Tulap distribution, its density is the constant value of $\frac{\exp(\epsilon)-1}{\exp(\epsilon)+1}$ on $(-1/2, 1/2)$, since it is the constructed CND by Proposition 2.4 [Awan and Vadhan, 2021, Corollary 3.10]. Suppose that $g(u)$ is non constant on $(-1/2, 1/2)$. Then it must take on some values above and below $\frac{\exp(\epsilon)-1}{\exp(\epsilon)+1}$ in order to still integrate to 1. To rule out trivial cases, where $g(x)$ is equivalent to the Tulap density up to a set of measure zero, we assume that the sets

$$
\begin{aligned}
A &:= \left\{ u \in (-1/2, 1/2) \mid g(u) > \frac{\exp(-\epsilon) - 1}{\exp(\epsilon) + 1} \right\},\\
B &:= \left\{ u \in (-1/2, 1/2) \mid g(u) \leq \frac{\exp(-\epsilon) - 1}{\exp(\epsilon) + 1} \right\},
\end{aligned}
$$

both have positive Lebesgue measure. We denote λ as the Lebesgue measure.

By Lemma A.11, we know that there exists $\omega \in (-1, 1)$ such that $(B + \omega) \cap A)$ has positive Lebesgue measure. By symmetry of g about zero, there exists a positive shift $\omega \in (0, 1)$ such that $\lambda((B+\omega)\cap A) > 0$. Consider $\Delta := 1-\omega \in (0, 1)$. Let $v \in (B+\omega)\cap A$ and $u := v-\omega \in B\cap(A-\omega)$, and consider the likelihood ratio:

$$
\frac{g(1 + u - \Delta)}{g(1 + u)} = \frac{g(u + \omega)}{g(1 + u)} = \frac{g(v)}{e^{-\epsilon} g(u)} > e^{\epsilon},
$$

where we use the fact that $u \in B$, $v \in A$, and $\lambda((B+\omega) \cap A) > 0$. This means that the likelihood ratio $g(x-\Delta)/g(y)$ is not bounded by $\exp(\epsilon)$ almost everywhere, for all $\Delta \in [-1,1]$; by Awan et al. [2019, Proposition 2.3], this means that the additive mechanism with density g does not satisfy ϵ-DP. In other words, for $X \sim g$ and Δ as above, $T(X, X+\Delta)(\alpha) < f_{\epsilon,0}(\alpha)$ for some $\alpha \in (0,1)$. We conclude that g is not a CND for $f_{\epsilon,0}$. The only assumption we made about g was that it was non-constant on $(-1/2, 1/2)$ on a set of positive probability. Due to the contradiction, we conclude that g is in fact constant on $[-1/2, 1/2]$ almost everywhere, which means that it is distributed as $\mathrm{Tulap}(0, \exp(-\epsilon), 0)$. □

Theorem 3.3. *A nontrivial tradeoff function $f \in \mathscr{F}$ is log-concave if and only if it is infinitely divisible. In particular,*

1. *If f is log-concave with log-concave CND $N \sim F$, then $\{f_t \mid t \geq 0\}$ defined by $f_t = F(F^{-1}(\alpha) - t)$ satisfies the assumptions of Definition 3.2.*
2. *Let f be infinitely divisible, with monoid $\{f_t \in \mathscr{F} \mid t \geq 0\}$, as defined in Definition 3.2, such that $f = f_1$. Let F_s be any CND for f_s (such as constructed in Proposition 2.4). Then the following limit exists $F^*(t) := \lim_{s \to 0} F_s(\frac{1}{s}t)$ and $N \sim F^*$ is the unique log-concave CND for f. Furthermore, $F^*(st)$ is the unique log-concave CND for f_s, for all $s > 0$.*

Proof. 1) Let f be a log-concave tradeoff function with log-concave CND F. Define $f_s(\alpha) = F(F^{-1}(\alpha) - s)$, which is a tradeoff function since F is log-concave. Note that $f_s \circ f_t = f_{s+t}$, f_t is nontrivial except wen $t = 0$, and $f_t \to \mathrm{Id}$ as $t \to 0$.

2) For part 2, we first show that the limit exists for the specific sequence $s_n = 1/n!$, and then we will show that convergence holds for all sequences that converge to zero. By construction, $F_{s_n}(\cdot)$ is a CND for f_{s_n}. So, $F_{s_n}(\cdot)$ has values determined on $(1/2)\mathbb{Z}$, no matter the choice of CND. Furthermore, $F_{s_n}(t/s_n)$ is a CND for $f_1 = f_{s_n}^{\circ n!}$, by Lemma A.5. Then for any choice of CND for F_{s_n}, the cdf $F_{s_n}(t/s_n)$ has values determined on $(s_n/2)\mathbb{Z}$, and $F_{s_n}(t/s_n)$ is a continuous cdf (as it is a CND). Note that the sequence $(s_n)_{n=1}^{\infty}$ satisfies $(s_n/2)\mathbb{Z} \subset (s_{n+1}/2)\mathbb{Z}$ for all n, and as $n \to \infty$, we have that $\bigcup_{n=1}^{\infty}((s_n/2)\mathbb{Z}) = \mathbb{Q}$, the set of rational numbers. In words, the set of determined values of $F_{s_n}(t/s_n)$ is an increasing sequence of sets, whose limit is the rational numbers. Due to this, the sequence of CNDs $F_{s_n}(t/s_n)$ is "pinned down" at an increasing number of points, and is eventually determined at every rational number. Since every $F_{s_n}(t/s_n)$ is monotone, and $\mathbb{Q}$ is dense in $\mathbb{R}$, it follows that the limit of this sequence, F^*, is a unique monotone function.

Next we show that F^* is a continuous cdf. We already mentioned that F^* is non-decreasing, and it is easy to show that $\lim_{t \to +\infty} F^*(t) = 1$ and $\lim_{t \to -\infty} F^*(t) = 0$. The challenging part is to show that F^* is continuous. It suffices to show that the convergence of $(F_{s_n}(t/s_n))_{n=1}^{\infty}$ to F^* is uniform. Before we show this, we establish the following inequality: for all $t \in \mathbb{R}$, $|F_{s_n}(t/s_n) - F^*(t)| \leq \sup_t |t - f_{s_n}(t)|$. To see this, let $t \in \mathbb{R}$. Then for each n, there exists $k_n \in \mathbb{Z}$ such that $\frac{(k_n-1)s_n}{2} \leq t \leq \frac{(k_n+1)s_n}{2}$. Since F^* is a non-decreasing function, this implies that

$$
\begin{aligned}
F^*\left(\frac{(k_n-1)s_n}{2}\right) \leq F^*(t) &\leq F^*\left(\frac{(k_n+1)s_n}{2}\right) \\
F_{s_n}\left(\frac{k_n-1}{2}\right) \leq F^*(t) &\leq F_{s_n}\left(\frac{k_n+1}{2}\right) \\
f_{s_n}\left(F_{s_n}\left(\frac{k_n+1}{2}\right)\right) \leq F^*(t) &\leq F_{s_n}\left(\frac{k_n+1}{2}\right),
\end{aligned}
$$

where if $F_{s_n}(\frac{k_n+1}{2}) < 1$ the third line is equivalent to the second line by Lemma 2.3, and if $F_{s_n}(\frac{k_n+1}{2}) = 1$, then the inequality in the third line is potentially weaker. By similar reasoning, we have that $f_{s_n}(F_{s_n}(\frac{k_n+1}{2})) \leq F_{s_n}(t/s_n) \leq F_{s_n}(\frac{k_n+1}{2})$ as well. Therefore,

$$
|F^*(t) - F_{s_n}(t/s_n)| \leq \left| F_{s_n}\left(\frac{k_n+1}{2}\right) - f_{s_n}\left(F_{s_n}\left(\frac{k_n+1}{2}\right)\right)\right| \leq \sup_{t \in [0,1]} |t - f_{s_n}(t)|.
$$

We are now ready to prove uniform convergence. Let $\epsilon > 0$ be given. Let $N \in \mathbb{Z}^+$ be such that $\sup_{t \in [0,1]} |t - f_{s_n}(t)| < \epsilon$, which is possible since $f_{s_n}(t) \to t$ uniformly (Polya's theorem). Then

for all $n \geq N$, we have that $|F_{s_n}(t/s_n) - F^*(t)| \leq \sup_t |t - f_{s_n}(t)| < \epsilon$. Uniform convergence of continuous functions implies that the limit function is also continuous, so we conclude that F^* is a continuous cdf.

Next we will show that for all $t \in \mathbb{R}^+$, $f_t = F^*(F^{*-1}(\alpha) - t)$. Let $(-x, x) := F^{*-1}((0,1))$ be the support of the distribution F^*. Let $q \in \mathbb{Q}^+$ be the "shift," and let $p \in (-\infty, x-q) \cap \mathbb{Q}$ be the "threshold" in the test. Let $n \in \mathbb{Z}^+$ be such that $p = a/n!$ and $q = b/n!$ for some $a \in \mathbb{Z}$ and $b \in \mathbb{Z}^+$. As we did earlier, denote $s_n = 1/n!$. Recall that $F^*(t)$ and $F_{s_n}(t/s_n)$ agree for all $t \in \frac{s_n}{2}\mathbb{Z}$. In particular, $p \in \frac{s_n}{2}\mathbb{Z}$. Then

$$\begin{aligned}
F^*(p) &= F^*(s_n a) \\
&= F_{s_n}(a) \\
&= f_{s_n} \circ F_{s_n}(a+1) \\
&= f_{s_n}^{\circ b} \circ F_{s_n}(a+b) \\
&= f_q \circ F_{s_n}(a+b) \\
&= f_q \circ F^*(s_n(a+b)) \\
&= f_q \circ F^*(p+q),
\end{aligned}$$

where we used the fact that $p < x - q$ to establish that $F_{s_n}(a+b) = F^*(p+q) < 1$, which enabled the recurrence application of Lemma 2.3. Furthermore, since the rational numbers are dense in $\mathbb{R}$, F^* is continuous, and f_q is continuous on $[0,1)$, we have that $F^*(t) = f_q \circ F^*(t+q)$ for all t such that $t < x - q$. Now let $\alpha \in [0,1)$ and call $t_\alpha = F^{*-1}(\alpha) - q$. Note that $F^*(q + t_\alpha) = \alpha$ and that $t_\alpha < x - q$. Then we have that $F^*(F^{*-1}(\alpha) - q) = f_q(\alpha)$ for all $\alpha \in [0,1)$. Finally, we extend the result for arbitrary f_r, $r \in \mathbb{R}^+$. Let $q_n \in \mathbb{Q}^+$ be a sequence such that $q_n \to r$. Then

$$f_r = f_{r-q_n} \circ f_{q_n} = f_{r-q_n} \circ F^*(F^{*-1}(\alpha) - q_n) \to \mathrm{Id} \circ F^*(F^{*-1}(\alpha) - r),$$

since f_{r-q_n} converges to Id uniformly, and F^* is continuous. We have that $F^*(s\cdot)$ is a CND for f_s: the symmetry of F^* is obvious, and the fact that F^* satisfies DP follows by the property that $f_s = F^*(F^{*-1}(\alpha) - s)$.

Let $s > 0$. Let $N \sim F^*$. Since $N/s \sim F^*(st)$ is a CND for f_s, we have that $T(N, N+s)(\alpha) = T(N/s, N/s{+}1)(\alpha) = f_s(\alpha) = F^*(F^{*-1}(\alpha) - s)$. However, $T(N, N{+}s)(\alpha) = F^*(F^{*-1}(\alpha) - s)$ holds for all $s > 0$ if and only if F^* has a log-concave density [Dong et al., 2022, Lemma A.3]. Therefore F^* is a log-concave distribution, and $F^*(s\cdot)$ is a log-concave CND for f_s for all $s > 0$.

Finally, we will make sure the limit does not depend on the specific sequence $s_n = 1/n!$. We will use a very similar argument as when we established uniform convergence to show that for any positive sequence r_n which converges to zero, $F_{r_n}(t/r_n)$ also converges uniformly to $F^*(t)$. Let $t \in \mathbb{R}$. Then for any $n \in \mathbb{Z}^+$, there exists k_n such that $\frac{(k_n-1)r_n}{2} \leq t \leq \frac{(k+1)r_n}{2}$. Then

$$F^*\left(\frac{(k_n-1)r_n}{2}\right) \leq F^*(t) \leq F^*\left(\frac{(k_n+1)r_n}{2}\right).$$

Since $F^*(\cdot r_n)$ and $F_{r_n}(\cdot)$ are both CNDs for f_{r_n}, they agree on all half integer values. So,

$$\begin{aligned}
F_{r_n}\left(\frac{k_n-1}{2}\right) \leq F^*(t) &\leq F_{r_n}\left(\frac{k_n+1}{2}\right) \\
f_{r_n}\left(F_{r_n}\left(\frac{k+1}{2}\right)\right) \leq F^*(t) &\leq F_{r_n}\left(\frac{k_n+1}{2}\right).
\end{aligned}$$

By similar reasoning, we have that $f_{r_n}\left(F_{r_n}\left(\frac{k+1}{2}\right)\right) \leq F_{r_n}(t/r_n) \leq F_{r_n}\left(\frac{k_n+1}{2}\right)$. Then

$$|F^*(t) - F_{r_n}(t/r_n)| \leq \left|F_{r_n}\left(\frac{k_n+1}{2}\right) - f_{r_n}\left(F_{r_n}\left(\frac{k_n+1}{2}\right)\right)\right| \leq \sup_{t\in[0,1]} |t - f_{r_n}(t)|.$$

Since $r_n \to 0$, we have that $f_{r_n}(t)$ converges uniformly to t. So, we have that $F_{r_n}(t/r_n)$ converges uniformly to $F^*(t)$. □

Lemma A.12. *Let f and g be tradeoff functions.*

1. *If f and g are piece-wise linear with k and ℓ break points (respectively), and f satisfies $f(x) = 0$ implies $x = 0$, then $f \circ g$ is piece-wise linear with at most $k + \ell$ break points, and at least $\max\{k, \ell\}$ break points.*

2. *If f is piece-wise linear with $k \geq 1$ break points, and $f(x) = 0$ implies that $x = 0$, then $f^{\circ n}$ has at least $k + (n - 1)$ break points.*

3. *If $f \circ g$ is piece-wise linear, then g is piece-wise linear on $[0, 1]$ and f is piece-wise linear on $[0, g(1)]$. (note that f can be arbitrary on $(g(1), 1]$ and it does not affect $f \circ g$)*

Proof. 1. The composition of linear functions is linear. So, it is clear that $f \circ g$ is piece-wise linear. Let B_g be the set of break points of g and B_f be the set of break points of f. Then the break points of $f \circ g$ are $g^{-1}(B_f) \bigcup B_g$, since f is invertible, which has at most $|B_f| + |B_g| = k + \ell$ elements, and at least $\max\{k, \ell\}$ elements.

2. Let B_f be the set of break points of f. Then the set of break points of $f^{\circ n}$ is $B_f \cup f^{-1}(B_f) \cup f^{-1}(f^{-1})(B_f) \cup \cdots \cup (f^{-1})^{\circ(n-1)}(B_f)$. The number of break points of $f^{\circ n}$ is then lower bounded by $|B_f \cup f(B_f) \cup f^{\circ 2}(B_f) \cup \cdots \cup f^{\circ(n-1)}(B_f)|$, by applying $f^{\circ(n-1)}$ to each of the sets (since applying a function to a set cannot increase the cardinality). We know that $|B_f| = k$. Because $f(x) = 0$ implies that $x = 0$, we have that f is strictly increasing on $(0, 1)$; so we have that $|f(B_f)| = k$ as well. Furthermore, for each $x \in B_f$, $f(x) < x$ as f is nontrivial ($k \geq 1$ implies nontrivial). Let x_m be the minimum element in B_f. Then $f(x_m) \in f(B_f)$ and $f(x_m) \notin B_f$. So, $|B_f \cup f(B_f)| \geq |B_f| + 1 = k + 1$. Continuing this process, we get that the number of break points of $f^{\circ n}$ is at least $k + n - 1$.

3. Since $f \circ g$ is piece-wise linear, $\frac{d^2}{d\alpha^2} f \circ g(\alpha) = 0$ except at finitely many values. Then

$$\begin{aligned} 0 &= \frac{d^2}{d\alpha^2} f \circ g(\alpha) \\ &= \frac{d}{d\alpha} \left(f'(g(\alpha))g'(\alpha)\right) \\ &= f''(g(\alpha))(g'(\alpha))^2 + f'(g(\alpha))g''(\alpha), \end{aligned}$$

except at finitely many values. Note that since f and g are non-decreasing and convex, the following quantities are non-negative (whenever they are well-defined): f', g', f'', and g''. So, the above equation implies that for all but finitely many α, either $f''(g(\alpha)) = 0$ or $g'(\alpha) = 0$ and either $f'(g(\alpha)) = 0$ or $g''(\alpha) = 0$. Note that g is zero on $\{\alpha \mid g'(\alpha) = 0\}$ and f is zero on $\{g(\alpha) \mid f'(g(\alpha)) = 0\}$. Furthermore, g is piece-wise linear on $\{\alpha \mid g''(\alpha) = 0\}$ and f is piece-wise linear on $\{g(\alpha) \mid f''(g(\alpha)) = 0\}$. We see that on $(0, 1)$, g is either zero or piece-wise linear, and so it is piece-wise linear on $[0, 1]$. Similarly on $(g(0), g(1)) = (0, g(1))$, f is either zero or piece-wise linear, and so it is piece-wise linear on $[0, g(1)]$.

□

Proposition 3.6. *1. Let f be a nontrivial piece-wise linear tradeoff function with $k \geq 1$ breakpoints and such that $f(x) = 0$ implies that $x = 0$. Then there is no tradeoff function g such that $g^{\circ(k+1)} = f$.*

2. *Let $\epsilon > 0$. There does not exist nontrivial symmetric tradeoff functions f_1 and f_2 such that $f_{\epsilon,0} = f_1 \circ f_2$.*

3. *Let f be the tradeoff function obtained by an arbitrary sequence of mechanism compositions, functional compositions, or subsampling (without replacement) of $f_{\epsilon,0}$ (could be different ϵ values for each). Then f is not infinitely divisible and so does not have a log-concave CND.*

Proof. 1. By part 3 of Lemma A.12, if f could be written as $g^{\circ(k+1)} = f$, then g must also be piece-wise linear. Since $f(x) = 0$ implies that $x = 0$, g also satisfy $g(x) = 0$ implies $x = 0$. If g is a nontrivial piece-wise linear tradeoff function (with $j \geq 1$ breakpoints), then by part 2 of Lemma A.12, $g^{\circ(k+1)}$ has $j + (k + 1) - 1 > k$ break points. This contradicts that $g^{\circ(k+1)} = f$.

2. Suppose that $f_{\epsilon,0} = g \circ h$, where both g and h are non-trivial. By part 3 of Lemma A.12, we know that h is piece-wise linear. Since we are assuming that h is non-trivial, it must have at least one break point. Since $f_{\epsilon,0}(x) = 0$ implies that $x = 0$, g must have this property as well. By Lemma part 1 of A.12, this implies that h must have a single break point. To agree with $f_{\epsilon,0}$, the breakpoint of h must be at $1 - c$, where $c = 1/(1 + \exp(\epsilon))$ is the solution to $f_{\epsilon,0}(1 - c) = c$, since this is where the unique breakpoint of $f_{\epsilon,0}$ lies. However, since h is a symmetric piece-wise linear function with a unique breakpoint at $1 - c$, the only possibility is that $h = f_{\epsilon,0}$.

3. Each application of composition, functional composition, and subsampling without replacement preserves the piece-wise property of the tradeoff function, as well as the property that $f(x) = 0$ implies that $x = 0$. The result follows from part (a).

□

Proposition 4.2. *Suppose that $f = f_1 \otimes f_2 \otimes \cdots \otimes f_k$ all be nontrivial and symmetric tradeoff functions, and let $F_1, F_2, \ldots, F_k$ be CNDs for $f_1, \ldots, f_k$ respectively. Let $N = (N_1, \ldots, N_k)$ be the random vector where $N_i \sim F_i$ are independent. Then N is a CND for f with respect to $\|\cdot\|_\infty$.*

Proof. For property 1 of Definition 4.1, let v be such that $\|v\|_\infty \leq 1$. Then $v = (v_1, \ldots, v_k)$ is such that $|v_i| \leq 1$. Then

$$\begin{aligned} T(N, N + v) &= T((N_1, \ldots, N_k), (N_1 + v_1, \ldots, N_k + v_k)) \\ &= T(N_1, N_1 + v_1) \otimes \cdots \otimes T(N_k, N_k + v_k) \\ &\geq f_1 \otimes \cdots \otimes f_k. \end{aligned}$$

If we set $v_i^* = 1$ for all i, then repeating the above gives equality in the last step, proving property 2 of Definition 4.1. Call $g(x_1, \ldots, x_n) := F_1'(x_1) \cdots F_k'(x_k)$ the density of N. For property 4 of Definition 4.1, since F_i' is symmetric about zero we have that g is also symmetric about zero. For property 3, let $w = (w_1, \ldots, w_k)^\top$ be any vector and $v^* = (1, 1, \ldots, 1)^\top$. Then,

$$\frac{g(w + tv^* - v^*)}{g(w + tv^*)} = \frac{F_1'(w_1 + (t-1)) \cdots F_k'(w_k + (t-1))}{F_1'(w_1 + t) \cdots F_k'(w_k + t)},$$

which is increasing in t, since each of the factors is increasing in t, by property 3 of Definition 2.2.

□

Theorem 4.3. *Let f be a nontrivial and symmetric log-concave tradeoff function with log-concave CND F. Let $N = (N_1, \ldots, N_k)$ be the random vector where $N_i \sim F$ are independent. Then N is a (log-concave) CND for f with respect to $\|\cdot\|_1$.*

Proof. Since f is a nontrivial log-concave tradeoff function, by Theorem 3.3 there exists a monoid of log-concave tradeoff functions $\{f_t \in \mathscr{F} \mid t \geq 0\}$ satisfying $f_t \circ f_s = f_{t+s}$ such that $f_1 = f$ and $f_t = F(F^{-1}(\alpha) - t)$ for all $t > 0$. Note that for any $t \in \mathbb{R}$, $T(N_i, N_i + t) = f_{|t|}$.

For property 1 of Definition 4.1, let x be such that $\|x\|_1 \leq 1$. Note that $|x_i| < 1$ for all $i = 1, \ldots, k$. Then

$$\begin{aligned} T(N, N + x) &= f_{|x_1|} \otimes f_{|x_2|} \otimes \cdots \otimes f_{|x_k|} \\ &\geq f_{|x_1|} \circ f_{|x_2|} \circ \cdots \circ f_{|x_k|} \\ &= f_{\sum_{i=1}^k |x_i|} \\ &= f_{\|x_i\|_1} \\ &\geq f_1, \end{aligned}$$

where in the first line, we use the property that $T(N_i, N_i + x_i) = f_{|x_i|}$ which uses log-concavity, the second line uses Lemma A.2, and the third line uses the property that $f_s \circ f_t$ within the monoid. Note that for $v = (1, 0, 0, \ldots, 0)$, $T(N, N + v) = T(N_1, N_1 + 1) = f$, proving property 2 of Definition 4.1. Since N is constructed by independent 1-d CNDs, the same arguments used in the proof of Proposition 4.2 can be used to prove properties 3 and 4 of Definition 4.1. Note that N is log-concave, since it is a product distribution with log-concave components. □

Proposition 4.4. *Let Σ be a $d \times d$ positive definite matrix. Let $v^* \in \operatorname{argmax}_{\|u\| \le 1} \|\Sigma^{-1/2} u\|_2$. Then $N(0, \Sigma)$ is a d-dimensional CND for $\|\Sigma^{-1/2} v^*\|_2$-GDP with respect to the norm $\|\cdot\|$.*

Proof. Let $N \sim N(0, \Sigma)$. Note that $\Sigma^{-1/2} N \sim N(0, I_d)$. Let u be such that $\|u\| \le 1$. Then

$$\begin{aligned} T(N, N+u) &= T\left(\Sigma^{-1/2} N, \Sigma^{-1/2} N + \Sigma^{-1/2} u\right) \\ &= T\left(N(0, I_d), N(0, I_d) + \Sigma^{-1/2} u\right) \\ &= T\left(N(0,1), N(\|\Sigma^{-1/2} u\|_2, 1)\right) \\ &= G_{\|\Sigma^{-1/2} u\|_2} \\ &\ge G_{\|\Sigma^{-1/2} v^*\|_2}, \end{aligned}$$

where for the third line, we use the rotational invariance of the multivariate Gaussian distribution. Note that setting $u = v^*$ gives equality. This establishes properties 1 and 2 of Definition 4.1. Property 4 of Definition 4.1 holds since the density of $N(0, \Sigma)$ is symmetric about zero. For property 3, let $w \in \mathbb{R}^k$ be any vector, and call g the density of $N(0, \Sigma)$. Call $a = (w + tv^*)$ and $b = -v^*$. Then,

$$\begin{aligned} \log \frac{g(w + tv^* - v^*)}{g(w + tv^*)} &= \log \frac{\exp(-\frac{1}{2}(w + (t-1)v^*)^\top \Sigma^{-1} (w + (t-1)v^*))}{\exp(-\frac{1}{2}(w + tv^*)^\top \Sigma^{-1} (w + tv^*))} \\ &= -\frac{1}{2}\left[(w + tv^*) - v^*)^\top \Sigma^{-1} ((w + tv^*) - v^*) + (w + tv^*)^\top \Sigma^{-1} (w + tv^*)\right] \\ &= -\frac{1}{2}\left[(a+b)^\top \Sigma^{-1} (a+b) - a^\top \Sigma^{-1} a\right] \\ &= -\frac{1}{2}\left[a^\top \Sigma^{-1} a + 2a^\top \Sigma^{-1} b + b^\top \Sigma^{-1} b - a^\top \Sigma^{-1} a\right] \\ &= -\frac{1}{2}\left[2a^\top \Sigma^{-1} b + b^\top \Sigma^{-1} b\right] \\ &= -\frac{1}{2}\left[2(w + tv^*)^\top \Sigma^{-1} (-v^*) + (v^*)^\top \Sigma^{-1} v^*\right] \\ &= t(v^*)^\top \Sigma^{-1} v^* + w^\top \Sigma^{-1} v^* - \frac{1}{2}(v^*)^\top \Sigma^{-1} v^*, \end{aligned}$$

which is increasing in t, since Σ is positive definite, which verifies property 3 of Definition 4.1. □

Proposition 4.6. *Let $0 < \delta \le 1$, $d \ge 1$, and $\|\cdot\|$ be a norm on $\mathbb{R}^d$. Call $v^* \in \underset{\|v\| \le 1}{\arg\min} \prod_{i=1}^d (1 - \delta |v_i|)$ and $A = \prod_{i=1}^d (1 - \delta |v_i^*|)$. Then $U(\frac{-1}{2\delta}, \frac{1}{2\delta})^n$ is a CND for $f_{0, 1-A}$ under $\|\cdot\|$-sensitivity. In the special case of $\|\cdot\| = \|\cdot\|_\infty$, this simplifies to $A = (1-\delta)^d$.*

Proof. Let $X \sim U(\frac{-1}{2\delta}, \frac{1}{2\delta})^n$, and let v be such that $\|v\| \le 1$. We need a lower bound on $T(X, X+v)$. Since this is the testing of shifted uniforms, $T(X, X+v) = f_{0, \mathrm{TV}(X, X+v)}$.

$$\begin{aligned} \mathrm{TV}(X, X+v) &= 1 - \frac{\prod_{i=1}^d \left(\frac{1}{\delta} - |v_i|\right)}{\delta^{-d}} \\ &= 1 - \prod_{i=1}^d (1 - \delta |v_i|) \\ &\ge 1 - A, \end{aligned}$$

which establishes property 1 of Definition 4.1. Note that using v^* as defined above, we get that $\mathrm{TV}(X, X+v^*) = 1 - A$, giving property 2 of Definition 4.1. Property 4 of Definition 4.1 is obvious, since each uniform is centered at zero. For property 3, let $w \in \mathbb{R}^d$. The likelihood ratio is

$$\frac{g(w + (t-1)v^*)}{g(w + tv^*)} = \prod_{i=1}^d \frac{I(\frac{-1}{2\delta} \le w_i + (t-1)v_i^* \le \frac{1}{2\delta})}{I(\frac{-1}{2\delta} \le w_i + tv_i^* \le \frac{1}{2\delta})},$$

and we see that each of these factors can take the possible values:

$$\begin{cases} \text{undefined} & \text{when } w_i + (t-1)v_i^* \notin [\frac{-1}{2\delta}, \frac{1}{2\delta}] \text{ and } w_i + tv_i^* \notin [\frac{-1}{2\delta}, \frac{1}{2\delta}] \\ 0 & \text{when } w_i + (t-1)v_i^* \notin [\frac{-1}{2\delta}, \frac{1}{2\delta}] \text{ and } w_i + tv_i^* \in [\frac{-1}{2\delta}, \frac{1}{2\delta}] \\ 1 & \text{when } w_i + (t-1)v_i^* \in [\frac{-1}{2\delta}, \frac{1}{2\delta}] \text{ and } w_i + tv_i^* \in [\frac{-1}{2\delta}, \frac{1}{2\delta}] \\ +\infty & \text{when } w_i + (t-1)v_i^* \in [\frac{-1}{2\delta}, \frac{1}{2\delta}] \text{ and } w_i + tv_i^* \notin [\frac{-1}{2\delta}, \frac{1}{2\delta}] \end{cases}$$

If $w_i + tv_i^* \in [\frac{-1}{2\delta}, \frac{1}{2\delta}]$ for some t, then we have that as t progresses from $-\infty$ to ∞, the value of each factor goes from undefined to 0 to 1 to $+\infty$ to undefined, which is a non-decreasing sequence. If $w_i + tv_i^* \notin [\frac{-1}{2\delta}, \frac{1}{2\delta}]$ for every t, then the likelihood ratio is always undefined, which is also trivially non-decreasing. We see that property 3 of Definition 4.1 holds. □

The privacy loss random variable is a concept that appears in all major definitions of differential privacy. In fact, Zhu et al. [2022] showed that the privacy loss random variables can be losslessly converted back and forth to the corresponding tradeoff function. For part of the proof of Proposition 4.7, it will be easier to work with the privacy loss random variables than directly with the tradeoff functions. First, we give a formal definition and a few basic properties of privacy loss random variables. While similar results appeared in Zhu et al. [2022], we include them here for completeness.

Definition A.13 (Privacy Loss Random Variable)**.** Let X and Y be two random variables on $\mathbb{R}^d$, with densities p and q, respectively. The *privacy loss random variable* is $\mathrm{PLRV}(X|Y) := \log \frac{p(X)}{q(X)}$, where $X \sim p$.

Lemma A.14 (Privacy Loss RV is Sufficient)**.** *Let $X \sim p$ and $Y \sim q$ be two random variables on $\mathbb{R}^d$ with densities p and q, respectively. Define $L(x) : \mathscr{X} \to \mathbb{R}$ by $L(x) = \log[q(x)/p(x)]$. Note that $L(X) \overset{d}{=} -\mathrm{PLRV}(X|Y)$ and $L(Y) \overset{d}{=} \mathrm{PLRV}(Y|X)$. Then $T(X,Y) = T(L(X), L(Y)) = T(-\mathrm{PLRV}(X|Y), \mathrm{PLRV}(Y|X))$.*

Proof. First, by postprocessing, we have that $T(X,Y) \leq T(L(X), L(Y))$ [Dong et al., 2022, Lemma 2.9]. For the other direction, note that by the Neyman Pearson Lemma, the optimal test for $H_0 : X$ versus $H_1 : Y$ at size α is of the form

$$\phi(x) = \begin{cases} 1 & L(x) > t \\ c & L(x) = t \\ 0 & L(x) < t, \end{cases}$$

where L is defined in the Lemma statement, and the values of t and c are uniquely chosen such that $\mathbb{E}_{X\sim p}\phi(X) = \alpha$.

For a specified t and c, the type I error is

$$\begin{aligned} \text{type I} = \mathbb{E}_{X\sim p}\phi &= \mathbb{E}_{X\sim p}[I(L(X) > t) + cI(L(X) = t)] \\ &= P_{X\sim p}(L(X) > t) + cP_{X\sim p}(L(X) = t), \end{aligned}$$

which we see only depends on the statistic $L(X)$. On the other hand,

$$\begin{aligned} \text{type II} = 1 - \mathbb{E}_{Y\sim q}\phi(Q) &= 1 - \mathbb{E}_{Y\sim q}[I(L(Y) > t) + cI(L(Y) = t)] \\ &= P_{Y\sim q}(L(Y) \leq t) - cP_{Y\sim q}(L(Y) = t), \end{aligned}$$

which we see only depends on the statistic $L(Y)$.

So, when testing $H_0 : L(X)$ versus $H_1 : L(Y)$, using the particular test $\psi(L) = I(L > t) + cI(L = t)$, where the values of c and t are chosen as above, we recover the type I and type II errors of $T(X,Y)$. We conclude that $T(L(X), L(Y)) \leq T(X,Y)$. Combining the inequalities, we have that $T(X,Y) = T(L(X), L(Y))$. The fact that $L(X) = -\mathrm{PLRV}(X|Y)$ and $L(Y) = \mathrm{PLRV}(Y|X)$ follows from the definition of privacy loss random variables. □

Lemma A.15. *Let $X \in \mathbb{R}^d$ be a continuous random vector with density g, which is symmetric about zero. Then for any $v \in \mathbb{R}^d$, $\mathrm{PLRV}(X|X+v) = \mathrm{PLRV}(X+v|X)$. It follows that*

1. *$T(X, X+v) = T(\mathrm{PLRV(X|X+v)}, -\mathrm{PLRV}(X|X+v))$, and*

2. *Let $Y \in \mathbb{R}^p$ be another continuous random vector symmetric about zero, and let $w \in \mathbb{R}^p$. Then if* $\mathrm{PLRV}(X|X+v) \stackrel{d}{=} \mathrm{PLRV}(Y|Y+w)$ *then* $T(X, X+v) = T(Y, Y+w)$.

Proof. First note that $\mathrm{PLRV}(X|X+v) = \log \frac{g(X)}{g(X-v)}$, where $X \sim g$. Setting $Z = v - X$, we can write

$$\begin{aligned}
\mathrm{PLRV}(X+v|X) &= \log \frac{g(X-v)}{g(X)}, \text{ where } (X-v) \sim g \\
&= \log \frac{g(-Z)}{g(-Z+v)}, \text{ where } -Z \sim g \\
&= \log \frac{g(Z)}{g(Z-v)}, \text{ where } Z \sim g \quad \text{(by symmetry of } g) \\
&\stackrel{d}{=} \mathrm{PLRV}(X|X+v).
\end{aligned}$$

Combining the above work with Lemma A.14, we get $T(-\mathrm{PLRV}(X|X+v), \mathrm{PLRV}(X|X+v))$, which is equivalent to statement 1, since the tradeoff function is symmetric. For part two, if $\mathrm{PLRV}(X|X+v) \stackrel{d}{=} \mathrm{PLRV}(Y|Y+w)$, then clearly $T(\mathrm{PLRV}(X|X+v), -\mathrm{PLRV}(X|X+v)) = T(\mathrm{PLRV}(Y|Y+w), -\mathrm{PLRV}(Y|Y+w))$, which is equivalent to the statement in part 2, by part 1. □

Proposition 4.7. *Let $\epsilon > 0$, and $d \geq 1$. Let X be a d-dimensional random vector with density $g(x) = \frac{\exp(-\epsilon\|x\|_\infty)}{d!(2/\epsilon)^d}$. Then X is a CND for the tradeoff function L_ϵ with respect to $\|\cdot\|_\infty$.*

Proof. Note that for any vector $s \in \{-1, 1\}^d$, $sX \stackrel{d}{=} X$ (entry-wise multiplication). Because of this, it suffices to consider $T(X, X+v)$ for $v \geq 0$ (all entries non-negative).

First we will show that $T(X, X+1) = T(L, L+1)$, where X is the ℓ_∞-mech, and $L \sim \mathrm{Laplace}(0, 1/\epsilon)$ which has density $\frac{\epsilon}{2}\exp(-\epsilon|x|)$. We will do this using privacy loss random variables, applying part 2 of Lemma A.15. Note that since X and L are both symmetric random variables, it suffices to equate the privacy loss random variables $\mathrm{PLRV}(X|X+1)$ and $\mathrm{PLRV}(L|L+1)$. We can easily derive that $\mathrm{PLRV}(L|L+1) = -\epsilon|L| + \epsilon|L-1| = \epsilon[1-2L]_{-1}^{1}$, where $L \sim \mathrm{Laplace}(0, 1/\epsilon)$ and $[x]_a^b := \min\{\max\{x, a\}, b\}$ is the clamping function. Note that $\mathrm{PLRV}(L|L+1) \stackrel{d}{=} \epsilon[1-L_2]_{-1}^{1}$, where $L_2 \sim \mathrm{Laplace}(2/\epsilon)$.

Now for $T(X, X+1)$, the privacy loss random variable is $\mathrm{PLRV}(X|X+1) = -\epsilon\|X\|_\infty + \epsilon\|X-1\|_\infty$, where $X \sim g(x)$. We can simplify this expression as follows, using the notation $\max(X) = \max_{1\leq i\leq d} X_i$ and $\min(X) = \min_{1\leq i\leq d} X_i$:

$$\begin{aligned}
&\mathrm{PLRV}(X|X+1) \\
&= -\epsilon\|X\|_\infty + \epsilon\|X-1\|_\infty \\
&= \begin{cases} -\epsilon\max(X) + \epsilon(1-\min(X)) & \text{if } \max(X) \geq -\min(X) \;\&\; 1-\min(X) \geq \max(X)-1 \\ -\epsilon\max(X) + \epsilon(\max(X)-1) & \text{if } \max(X)-1 > 1-\min(X) \\ -\epsilon(-\min(X)) + \epsilon(1-\min(X)) & \text{if } -\min(X) > \max(X) \end{cases} \\
&= \begin{cases} -\epsilon\max(X) + \epsilon(1-\min(X)) & \text{if } -1 \leq [1-(\max(X)+\min(X))] \leq 1 \\ -\epsilon & \text{if } [1-(\max(X)+\min(X))] < -1 \\ \epsilon & \text{if } [1-(\max(X)+\min(X))] > 1 \end{cases} \\
&= \epsilon[1-(\max(X)+\min(X))]_{-1}^{1}.
\end{aligned}$$

Comparing this expression with $\mathrm{PLRV}(L|L+1)$, we see that it suffices to show $\max(X)+\min(X) \stackrel{d}{=} \mathrm{Laplace}(2/\epsilon)$. Recall that $X \stackrel{d}{=} RU$, where $R \sim \mathrm{Gamma}(d+1, \epsilon)$, using the shape, rate parameterization, and $U_i \stackrel{\text{iid}}{\sim} U(-1, 1)$ for $i = 1, \ldots, d$ [Hardt and Talwar, 2010, Remark 4.2]. By factoring out R, we get

$$\max(X) + \min(X) \stackrel{d}{=} R(\max(U) + \min(U)).$$

So, we will determine the distribution of $\max(U) + \min(U)$ first. We can easily compute the joint distribution of $\max(U)$ and $\min(U)$, as these are the minimum and maximum order statistics:

$$f_{\min(U),\max(U)}(x, y) = \frac{d(d-1)}{4}\left(\frac{y-x}{2}\right)^{d-2} I(-1 \le x \le y \le 1).$$

Now consider the change of variables $m = x$ and $w = x + y$. Applying change of variables, we have

$$\begin{aligned} f_{\min(U),\max(U)+\min(U)}(m, w) &= \frac{d(d-1)}{4}\left(\frac{(w-m)-m}{2}\right)^{d-2} I(-1 \le m \le w-m \le 1) \\ &= \frac{d(d-1)}{2^d}(w-2m)^{d-2} I(-1 \le m, m \le w/2, w-1 \le m). \end{aligned}$$

To get the distribution of $W = \max(U) + \min(U)$, we marginalize out $M = \min(U)$:

$$\begin{aligned} f_{\max(U)+\min(U)}(w) &= \int_{\max\{-1,w-1\}}^{w/2} \frac{d(d-1)}{2^d}(w-2m)^{d-2}\, dm \\ &= \frac{d(d-1)}{2^{d+1}} \frac{-(w-2m)^{d-1}}{d-1}\Big|_{\max\{-1,w-1\}}^{w/2} \\ &= \frac{d}{2^{d+1}}\left[(w - 2\max\{-1, w-1\})^{d-1} - \cancel{(w-2(w/2))^{d-1}}\right] \\ &= \frac{d}{2^{d+1}}\begin{cases}(w+2)^{d-1} & -2 \le w \le 0 \\ (2-w)^{d-1} & 0 \le w \le 2\end{cases} \\ &= \frac{d}{2^{d+1}}(2-|w|)^{d-1} I(-2 \le w \le 2). \end{aligned}$$

Since the distribution of $W = \max(U) + \min(U)$ is symmetric about zero, $W \overset{d}{=} (-1)^B|W|$ where $B \sim \mathrm{Bern}(1/2)$. So, our goal is to show $(-1)^B RW \overset{d}{=} (-1)^B \mathrm{Exp}(\epsilon/2)$, since the left side is equal in distribution to $\max(X) + \min(X)$ and the right side is equal in distribution to $\mathrm{Laplace}(2/\epsilon)$. The pdf of $Y \overset{d}{:=} |W|$ is $f(y) = \frac{d}{2^d}(2-y)^{d-1} I(0 \le y \le 2)$. It suffices to show that $RY \overset{d}{=} \mathrm{Exp}(\epsilon/2)$. Let ϕ_R be the characteristic function of R and ϕ_{RY} be the characteristic function of RY. Then,

$$\begin{aligned} \phi_{RY}(t) &= \mathbb{E}_Y \phi_R(tY) \\ &= \mathbb{E}_Y\left(1 - \frac{ity}{\epsilon}\right)^{-(d+1)} \\ &= \frac{d}{2^d}\int_0^2 \left(1 - \frac{ity}{\epsilon}\right)^{-(d+1)} (2-y)^{d-1}\, dy \\ &= \frac{\epsilon}{\epsilon - 2it} \\ &= \frac{\epsilon/2}{\epsilon/2 - it}, \end{aligned}$$

which we identify as the characteristic function of $\mathrm{Exp}(\epsilon/2)$, establishing that $RY \overset{d}{=} \mathrm{Exp}(\epsilon/2)$. By part 2 of Lemma A.15, this completes the proof that $T(X, X+1) = T(L, L+1)$, establishing property 2 of Definition 4.1. Note that property 4 of Definition 4.1 is obvious, and property 3 holds since the likelihood ratio $g(x-1)/g(x)$ is an increasing function in $\max(x) + \min(x)$, which itself is an increasing function of t when $x = w + t$ for every vector $w \in \mathbb{R}^d$. It remains to verify property 1 of Definition 4.1.

Next we will show that for $v \in (0, 1]^d$, $T(X, X+1) \le T(X, X+v)$ (this proof strategy is based on the proof of Lemma 3.5 in Dong et al. [2021]). We will separately address the cases that some of $v_i = 0$ at the end of the proof. Call $f_1 = T(X, X+1)$ and $f_v = T(X, X+v) = T(X/v, X/v+1)$. Define the two linear maps $r : x \mapsto x/v$ and $r^{-1} : x \mapsto vx$ (entry-wise multiplication and division), which are inverse maps. Note that $r(X) = X/v$ has density proportional to $\exp(-\epsilon\|vt\|_\infty)$. Let

$\alpha \in [0,1]$ be given. Let A be the optimal rejection region for $T(X/v, X/v+1)$ at type I error α. By our earlier work, we know that

$$A = \{x \mid \epsilon(\max(vx) + \min(vx)) \geq t\},$$

for some $t \in \mathbb{R}$, and it satisfies $P(r(X) \in A) = \alpha$ and $P(r(X) + 1 \notin A) = f_v(1-\alpha)$. We can now consider $r^{-1}(A)$ as a possible rejection region for testing $T(X, X+1)$, which is at best suboptimal. We compute the type I error as

$$P(X \in r^{-1}(A_v)) = P(r(X) \in A) = \alpha.$$

Suboptimality of the rejection region implies that

$$\begin{aligned} f_1(1-\alpha) &\leq P(X+1 \notin r^{-1}(A)) \\ &= P(r(X) + r(1) \notin A) \\ &= P(r(X) + 1/v \notin A) \\ &= P(r(X) \notin A - 1/v) \\ &\leq P(r(X) \notin A - 1) \\ &= f_v(1-\alpha), \end{aligned}$$

where we used the fact that r is a linear map, and $r(1) = 1/v$; the key step is the final inequality, which we justify as follows: it suffices to show that $(A - 1/v)^c \subset (A-1)^c$ or equivalently $A - (1/v - 1) \supset A$. We verify this by inspecting the definition of A:

$$\begin{aligned} A &= \{x \mid \max(vx) + \min(vx) \geq t\} \\ &\subset \{x \mid \max(vx + (1/v-1)) + \min(vx + (1/v-1)) \geq t\} \\ &= \{x - (1/v-1) \mid \max(vx) + \min(vx) \geq t\} \\ &= A - (1/v - 1), \end{aligned}$$

where in the inclusion step, we used the fact that $(1/v - 1) \geq 0$ implies that $\max(vx + (1/v-1)) \geq \max(vx)$ and $\min(vx + (1/v-1)) \geq \min(vx)$. This completes the argument that for $v \in (0,1]^d$, $f_1 = T(X, X+1) \leq T(X, X+v) = f_v$.

Finally, let $v \in [0,1]^d$, where the entries may possibly be zero. Let $v_n \in (0,1]^d$ be a sequence of vectors converging to v. Notice that $X + v_n \overset{TV}{\to} X + v$, since X has a continuous density. Since $T(X, X+v_n) \geq T(X, X+1)$ by our above work, by Corollary A.9 we have $T(X, X+v) \geq T(X, X+1)$ as well. □

Theorem 4.8. *Let $d \geq 2$ and let $\|\cdot\|$ be any norm on $\mathbb{R}^d$. Then for any $\epsilon > 0$, there is no random vector satisfying properties 1 and 2 of Definition 4.1 for $f_{\epsilon,0}$ with respect to the norm $\|\cdot\|$. In particular, there is no multivariate CND for $f_{\epsilon,0}$.*

The proof strategy of Theorem 4.8 is as follows: 1) observe that property of Definition 4.1 enforces constraints on the likelihood ratio $\log \frac{g(x-v)}{g(x)}$, 2) establish that the measure induced by g is equivalent to Lebesgue measure, which simplifies some measure theory details, 3) show that we can construct a vector w such that $\|w\| < 1$, $\|w+v\| < 1$, and $w \notin \mathrm{Span}(v)$, 4) based on the properties of w and v, show that by taking integral combinations of w and v, we can find an arbitrarily long sequence of points each sufficiently far from each other such that the value of g is bounded below by a common constant, and 5) show that point 4 implies that g is not integrable. Because densities are only well defined up to sets of Lebesgue measure zero, the details of the proof are more complicated to ensure that we are careful about the measure theoretical details.

Proof. Suppose to the contrary that there exists a CND for $f_{\epsilon,0}$ with respect to $\|\cdot\|$, which has density g. We will denote μ_g as the measure induced by g: $\mu_g(S) = \int_S g(x)\, dx$, and use λ to denote Lebesgue measure.

By property 2 of Definition 4.1, there exists $v \in \mathbb{R}^d$ such that $\|v\| \leq 1$ and $T(N, v+N) = f_{\epsilon,0}$, where $N \sim g$. This implies that $\log \frac{g(x-v)}{g(x)} = \pm\epsilon$ almost everywhere (μ_g) for all $x \in \mathbb{R}^d$ (if P and Q are two distributions satisfying $T(P,Q) = f_{\epsilon,0}$, then the privacy loss random variable is a binary

random variable, taking values in $\{-\epsilon, \epsilon\}$). Furthermore, by property 1, for any other vector $w \in \mathbb{R}^d$ such that $\|w\| \leq 1$, we have that $\log \frac{g(x-w)}{g(x)} \in [-\epsilon, \epsilon]$ for almost every $x \in \mathbb{R}^d$ (μ_g).

Before we begin our main argument, we will show that (if such a g exists,) μ_g must be equivalent to Lebesgue measure. We know that Lebesgue measure dominates μ_g, so we only need to show that $\mu_g = 0$ implies $\lambda = 0$. Suppose to the contrary that there exists $S \subset \mathbb{R}^d$ such that $\lambda(S) > 0$ but $\mu_g(S) = 0$ (which implies that $g(x) = 0$ a.e. on S). We claim that there exists such an S such that for some $\|t\| \leq 1$, $\mu_g(S + t) > 0$. We prove this as follows: begin with any S such that $\lambda(S) > 0$ but $\mu_g(S) = 0$. If $\mu_g(S + t) = 0$ for all $\|t\| \leq 1$, then set $S' = \bigcup_{\|t\|\leq 1}(S + t)$, which is strictly larger than S. If S' still does not have the desired property, repeat the process iteratively. Note that in the limit, this process results in $\mathbb{R}^d$, but $\mu_g(\mathbb{R}^d) = 1$. So, the process must terminate, giving us the desired set with the properties $\lambda(S) > 0$, $\mu_g(S) = 0$ and there exists some $\|t\| \leq 1$ such that $\mu_g(S + t) > 0$. Then there exists $P \subset S + t$ such that $g(x) > 0$ on P a.e., and note that $P - t \subset S$ and $g(x) = 0$ on $P - t$ a.e.. However, this implies that $\log \frac{g(x-t)}{g(x)} = \infty \notin [-\epsilon, \epsilon]$ on the set P, which has positive probability $\mu_g(P) > 0$. This contradicts property 1 of Definition 4.1, as discussed above. We conclude that μ_g and λ are equivalent measures. So, we will interchangeably use statements about Lebesgue measure and μ_g measure.

Let $r > 0$ be the largest value such that $\|x\|_2 \leq r$ implies that $\|x\| \leq 1$ (possible by the equivalence of norms on $\mathbb{R}^d$). Consider three sets

$$\begin{aligned} A &= \{w \mid \|\mathrm{Proj}_{v^\perp} w\|_2 < r/4 \;\&\; \|w\| < 1\}, \\ B &= \{w \mid \|\mathrm{Proj}_{v^\perp} w\|_2 \leq r/8 \;\&\; \|w\| < 1\}, \\ C &= \{w \mid \|w + v\| < 1\}. \end{aligned}$$

Note that $(A \setminus B) \cap C$ is an open set; we will demonstrate that it is non-empty, which implies that it has non-zero Lebesgue measure. First note that $A \setminus B \neq \emptyset$, since $d \geq 2$ implies that $\emptyset \subsetneq B \subsetneq A$. We will construct a vector $w \in (A \setminus B) \cap C$ as follows: let $y \in A \setminus B$, and call $z = \mathrm{Proj}_{v^\perp} y$. Then $r/8 < \|z\|_2 < r/4$. We set $w = z - v/2$. First we will check that $w \in A \setminus B$: since $z = \mathrm{Proj}_{v^\perp} w = \mathrm{Proj}_{v^\perp} y$, we have that $r/8 < \|\mathrm{Proj}_{v^\perp} w\| < r/4$. We also need to check that

$$\|w\| = \|z - v/2\| \leq \|z\| + \frac{1}{2}\|v\| \leq \frac{1}{4} + \frac{1}{2} < 1,$$

since $\|z\|_2 \leq r/4$ implies that $\|z\| \leq 1/4$. Next, we will check that $w \in C$:

$$\|w + v\| = \|z - v/2 + v\| = \|z + v/2\| \leq \|z\| + \frac{1}{2}\|v\| \leq \frac{1}{4} + \frac{1}{2} < 1,$$

using again the fact that $\|z\|_2 \leq r/4$ implies that $\|z\| \leq 1/4$. We conclude that $(A \setminus B) \cap C$ is a non-empty open set, which implies that it has non-zero Lebesgue measure.

Let $c := \operatorname{esssup} g$. Note that $c < \infty$, as otherwise, this would violate the log-likelihood ratio property discussed earlier. Then $g(x) \leq c$ holds with probability one. So, since $(A \setminus B) \cap C$ has positive probability, we can find a vector $w \in (A \setminus B) \cap C$ which satisfies $g(w) \leq c$.

Let $\eta \in (0, c)$ be given. Then the set $\{x \mid c - \eta \leq g(x) \leq c\}$ has positive measure. For v as defined above, and an arbitrary vector $u \in \mathbb{R}^d$, consider two more sets:

$$\begin{aligned} G_v &= \left\{ x \middle| \log \frac{g(x+v)}{g(x)} \in \{-\epsilon, \epsilon\} \right\}, \\ F_u &= \left\{ u \middle| \log \frac{g(x+u)}{g(x)} \in [-\epsilon, \epsilon] \right\}, \end{aligned}$$

which both hold with probability one whenever $\|u\| \leq 1$.

Let K be a positive integer such that $K > \left(e^{-\epsilon}(c-\eta)\frac{\pi^{d/2}}{\Gamma(1+d/2)}(r/8)^d\right)$. Then there exists $\xi \in \mathbb{R}^d$ such that for every $0 \le j \le k \le K$,

$$\begin{aligned}
\xi &\in G_v - (kw + jv), \\
\xi &\in F_w - (kw + jv), \\
\xi &\in F_{w+v} - (kw + jv), \\
c - \eta \le g(\xi) &\le c, \\
\|\xi\|_2 \le b &:= \operatorname*{esssup}_{c-\eta \le g(x) \le c} \|x\|_2,
\end{aligned}$$

since the first three lines hold with probability one, the fourth holds with positive probability as discussed earlier, and the last holds with probability one. Note that $b < \infty$ as otherwise, we would have an unbounded region with positive probability such that $g \ge c - \eta > 0$, which would imply that g is not integrable.

Since $\xi \in F_w$, we have that $g(\xi + w) \in [e^{-\epsilon}g(\xi), c] \subset [e^{-\epsilon}(c-\eta), c]$, since $\|w\| \le 1$. Similarly, as $\xi \in F_{w+v}$, we have that $g(\xi + w + v) \in [e^{-\epsilon}(c-\eta), c]$, since $\|w + v\| \le 1$. However, since $\xi \in G_v - w$, we have that $\frac{g(\xi+w+v)}{g(\xi+w)} = e^{\pm\epsilon}$. The only possibility to satisfy all of these constraints is for either $g(\xi + w) \ge c - \eta$ or for $g(\xi + w + v) \ge c - \eta$. Repeating the previous argument, starting with $g(\xi + w) \ge c - \eta$ gives either $g(\xi + 2w) \ge c - \eta$ or $g(\xi + 2w + v) \ge c - \eta$. If instead, we start with $g(\xi + w + v) \ge c - \eta$, then either $g(\xi + 2w + v) \ge c - \eta$ or $g(\xi + 2w + 2v) \ge c - \eta$. We see that after k steps of this procedure, we have that $g(\xi + kw + jv) \ge c - \eta$ for some $j \in \{0, 1, 2, \ldots, k\}$. We denote by $j(k)$ the value of j obtained by this procedure at the kth step.

For each $0 \le k \le K$, define

$$A_k = \{x \mid \|x - (\xi + kw + j(k)v)\|_2 < r/8\}.$$

Note that since $w \in A \setminus B$ from above, we know that $\|\mathrm{Proj}_{v^\perp} w\|_2 \ge r/8$. This implies that each A_k is disjoint from the others, since the A_k are of radius $r/8$, and the distance between each set is at least $r/8$. Furthermore, notice that on each A_k, $g \ge e^{-\epsilon}(c - \eta)$.

Finally, consider the integral of g, which we lower bound:

$$\begin{aligned}
\int_{\mathbb{R}^d} g(x)\, dx &\ge \sum_{k=0}^{K} \int_{A_k} g(x)\, dx \\
&\ge \sum_{k=0}^{K} \int_{A_k} e^{-\epsilon}(c-\eta)\, dx \\
&= \sum_{k=0}^{K} e^{-\epsilon}(c-\eta)\mathrm{Vol}(A_k) \\
&= (K+1)e^{-\epsilon}(c-\eta)\frac{\pi^{d/2}}{\Gamma(1+d/2)}(r/8)^d \\
&> 1,
\end{aligned}$$

where we used the formula for a d-dimensional sphere of radius $r/8$ to evaluate $\mathrm{Vol}(A_k)$, and in the last line, we used the fact that $K > \left(e^{-\epsilon}(c-\eta)\frac{\pi^{d/2}}{\Gamma(1+d/2)}(r/8)^d\right)$. We see that g cannot integrate to one, which contradicts our assumption that it is a multivariate CND. In fact, g is not even integrable, as K could have been chosen arbitrarily high.

□

Corollary 4.9. *Let $\epsilon > 0$ be given. There does not exist nontrivial symmetric tradeoff functions f_1 and f_2 such that $f_{\epsilon,0} = f_1 \otimes f_2$.*

Proof. Suppose to the contrary that there did exist a nontrivial decomposition $f_{\epsilon,0} = f_1 \otimes f_2$. Then Proposition 4.2 gives a construction for a 2-dimensional CND of $f_{\epsilon,0}$. However, we know from Proposition 4.8 that there is no 2-dimensional CND for $f_{\epsilon,0}$. □