Differential privacy for symmetric log-concave mechanisms*

Staal A. Vinterbo[†]

Abstract

Adding random noise to database query results is an important tool for achieving privacy. A challenge is to minimize this noise while still meeting privacy requirements. Recently, a sufficient and necessary condition for (ϵ, δ)-differential privacy for Gaussian noise was published. This condition allows the computation of the minimum privacy-preserving scale for this distribution. We extend this work and provide a sufficient and necessary condition for (ϵ, δ)-differential privacy for all symmetric and log-concave noise densities. Our results allow fine-grained tailoring of the noise distribution to the dimensionality of the query result. We demonstrate that this can yield significantly lower mean squared errors than those incurred by the currently used Laplace and Gaussian mechanisms for the same ϵ and δ.

1 INTRODUCTION

The Hippocratic oath, estimated to be about 2500 years old, mentions patient privacy. Privacy is also recognized as a fundamental human right (The United Nations, 1948) and is featured in some form in the constitutions of more than 180 countries (Elkins et al., 2013). Still, privacy, and data privacy in particular, is very much an unsolved problem.

Data privacy rests on keeping the data itself confidential as well as controlling information leakage by processes that have access to the data. In the following we consider leakage from database query results. A problem is that the result of a database query can be highly revealing about an individual. For example, consider the case where all records except one are known to an adversary. Now, any result of any linear query immediately reveals the query value of the unknown record.

The definition of (ϵ, δ)-differential privacy (Dwork et al., 2006b,a) addresses the problem of revealing query responses by requiring query response algorithms to be insensitive to single record changes. This insensitivity is achieved by random perturbations in the computation of query results and is quantified by that any single record change is not allowed to change the likelihood of any output event by more than a multiplicative factor e^{ϵ} plus a small additive term δ. For small ϵ and very small δ, this provides a strong form of privacy but still allows producing high quality statistics, particularly when they are informative and the data is representative (Dwork and Lei, 2009).

For real-valued queries, the prototypical differentially private response algorithms add noise distributed according to either the Laplace or Gaussian distributions to the real query answer.

*This work has been funded in part by Innlandet Fylkeskommune, as well as Research Council of Norway grants 308904 and 309834. We thank the anonymous reviewers for helpful comments.

[†]Department of Information Security and Communication Technology, Norwegian University of Science and Technology.

These algorithms are called the Laplace and Gaussian mechanisms, respectively (see e.g., Dwork and Roth (2014)).

For these mechanisms that add symmetric noise to query results, differential privacy places a condition on the scale parameter of the noise distributions. More privacy requires larger scale. Of course, a larger scale decreases the concentration of the now noisy query result around its true value. Therefore, a goal is to find the minimal scale that meets the requirement of differential privacy.

Pursuing this goal, there exists a line of work into sufficent conditions on the scale for the (ϵ, δ)-differentially private Gaussian mechanism, starting with what could be called the standard closed form scale bound for the Gaussian mechanism noted in Dwork et al. (2006b) (and also discussed in Dwork and Roth (2014)), via an improvement in the underlying analytic sufficient condition by Le Ny and Pappas (2014), to the analytic sufficient and necessary condition by Balle and Wang (2018). To our knowledge, the Gaussian distribution is the only continuous distribution for which a necessary and sufficient condition for (ϵ, δ)-differential privacy has been described so far.

Inspired by the above, we extend the line of work by a sufficient and necessary condition for (ϵ, δ)-differential privacy for mechanisms that add noise distributed according to log-concave and symmetric densities.

Our results allow us to demonstrate the optimization of utility by not only minimizing the scale of a fixed mechanism's perturbations to meet the privacy requirements, but also simultaneously choose the mechanism itself by taking the dimensionality of the query function into account.

More background and other related works are described in Section 5.

1.1 Summary of findings

Let d be a database, q be a real-valued query function on databases, $s \geq 0$, and X be a random variable distributed according to density f. We will call an algorithm that returns a variate of the random variable $q(d) + sX$ a "mechanism" and often identify the mechanism by its random variable. When $f = e^{-\psi}$ where ψ is even and convex, we call the associated mechanism symmetric log-concave.

Our main theoretical findings are

- Lemma 1 that states a sufficient and necessary condition for (ϵ, δ)-differential privacy for symmetric log-concave mechanisms for real valued query functions, and
- Lemma 8 that states that the condition in Lemma 1 can be extended naturally to query functions with values in $\mathbb{R}^n$ for $n \geq 1$ and mechanisms that add noise vectors distributed according to spherically symmetric log-concave distributions that are spherical with respect to the norm used to define the global sensitivity of the query. In particular, this is the case for noise vectors of iid Subbotin$_p$ random variables and p-norms for $p \geq 1$ as described in Theorem 9, giving rise to an infinite family of mechanisms containing both the Laplace and the Gaussian mechanisms.

These results enable maximizing utility by not only finding the optimal member of a fixed noise scale family such as the Gaussian, but also simultaneously selecting the scale family itself depending on the query response dimensionality. In our case, given the one-to-one correspondence between vector valued Subbotin$_p$ mechanisms and the p-norm used to define the global sensitivity Δ of the query function, we select the scale family by including the parameter p into the parameter optimization process. We demonstrate this idea by empirically showing that depending on the number of columns in a real valued data table, there are different Subbotin_p mechanisms that yield the smallest l_2-error for estimating the average row under given privacy constraints. As the number of columns increases, the error of the selected mechanism becomes much smaller than if just considering the Laplace or Gaussian scale families.

In addition, we show that a minimum scale bound that fulfills the condition in Lemma 1 is linear in Δ (Lemma 2), we produce closed form necessary and sufficient conditions for (ϵ, δ)-differential privacy and scale s

$$s \geq \frac{\Delta}{\epsilon - 2\log(1-\delta)}, \text{ and}$$
$$s \geq \frac{\Delta}{2\log\left(\frac{e^{\frac{\epsilon}{2}} + \sqrt{\delta(e^{\epsilon}+\delta-1)}}{1-\delta}\right)},$$

for the Laplace and Logistic mechanisms, respectively (Theorems 3 and 4). In the case of mechanisms supported on $\mathbb{R}$, we prove that one-dimensional mechanisms that can achieve $(\epsilon, 0)$-differential privacy can exhibit arbitrarily smaller variances than those that that cannot for small enough δ (Theorem 6). We demonstrate that this δ need not be very small by numerically showing that the Laplace and Logistic mechanisms, which can achieve $(\epsilon, 0)$-differential privacy, exhibit smaller variance than the Gaussian, that cannot, for a significant range of privacy parameters, for example when $\epsilon \geq 0.05$ and $\delta \leq 0.001$.

2 PRELIMINARIES

We briefly recapitulate select definitions and known results.

2.1 Subbotin densities

Let $C(r) := 2\Gamma\left(\frac{1}{r}\right) r^{\frac{1}{r}-1}$ where Γ denotes the gamma function, and let

$$f_r(x) := \frac{e^{-\frac{|x|^r}{r}}}{C(r)}.$$

The family of densities f_r for $r > 0$ are called Subbotin (Subbotin, 1923) densities, exponential power distribution densities, or the generalized normal distribution (Nadarajah, 2005) densities. This family includes the standard Laplace density as f_1, the standard Gaussian density as f_2, and the uniform density on $(-1, 1)$ in the limit as $r \to \infty$. Let $\Gamma(s, x)$ denote the upper incomplete gamma function. Then,

$$\begin{aligned} F_r(x) &:= \int_{-\infty}^{x} f_r(w)dw \\ &= \frac{1}{2} + \text{sign}(x)\left(\frac{1}{2} - \frac{\Gamma\left(\frac{1}{r}, \frac{|x|^r}{r}\right)}{2\Gamma\left(\frac{1}{r}\right)}\right), \\ F_r^{-1}(p) &= \text{sign}\left(p - \frac{1}{2}\right) F_\Gamma^{-1}\left(2|p - \frac{1}{2}|, \frac{1}{r}, \frac{1}{r}\right)^{\frac{1}{r}}, \end{aligned}$$

where $F_{\Gamma}^{-1}(p, \alpha, \beta)$ is the inverse CDF of the Gamma distribution with density $f_{\Gamma}(x) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$. For a random variable X distributed according to f_r,

$$\mathrm{Var}(X) = r^{\frac{2}{r}} \frac{\Gamma\left(\frac{3}{r}\right)}{\Gamma\left(\frac{1}{r}\right)}.$$

We will denote the Subbotin distribution with density f_r as Subbotin_r.

The function

$$\psi_r(x) := \frac{|x|^r}{r} + \log(C(r)) = -\log(f_r(x))$$

is even as x only enters as $|x|$, and also convex for $r \geq 1$ as $\psi_r''(x) = \frac{|x|^r(r-1)}{x^2} \geq 0$ then. Consequently, the Subbotin densities are even and log-concave for $r \geq 1$.

2.2 Differential privacy

Formally, let a database d be a collection of record values from some set V. Two databases d and d' are *neighboring* if they differ in one record. Let $\mathcal{N}$ be the set of all pairs of neighboring databases.

Definition 1. The global sensitivity of a real-valued function q on databases is

$$\Delta := \sup_{(d,d') \in \mathcal{N}} |q(d) - q(d')|.$$

Definition 2 ((ϵ, δ)-differential privacy Dwork et al. (2006b,a))**.** A randomized algorithm M is called (ϵ, δ)-differentially private if for any measurable set S of possible outputs and all $(d, d') \in \mathcal{N}$

$$\Pr(M(d) \in S) \leq e^\epsilon \Pr(M(d') \in S) + \delta,$$

where the probabilities are over randomness used in M. By ϵ-differential privacy we mean $(\epsilon, 0)$-differential privacy.

We are now ready to develop our results.

3 THE ONE-DIMENSIONAL CASE

3.1 A condition for real-valued mechanisms

The following is our first main result.

Lemma 1. *Let X be a random variable distributed according to density $f(x) = e^{-\psi(x)}$ where $\psi : \mathbb{R} \to (-\infty, \infty]$ is convex and even, with f having support $(-a, a)$ for $a \in (0, \infty]$. Then for a real-valued function q on databases with global sensitivity $\Delta \geq 0$ and a database d, the mechanism returning a variate of $q(d) + sX$ is (ϵ, δ)-differentially private for $\epsilon \geq 0$ and $\delta \geq 0$ if and only if for $F(x) = \int_{-\infty}^x f(u)du$*

$$F\left(\frac{\Delta - t}{s}\right) - e^\epsilon F\left(-\frac{t}{s}\right) \leq \delta \tag{1}$$

where

$$t := \sup\left\{z < as \mid \psi\left(\frac{z}{s}\right) - \psi\left(\frac{z-\Delta}{s}\right) \leq \epsilon\right\}. \tag{2}$$

Furthermore, if the above holds for scale $s > 0$, then it also holds for scales $s' > s$.

Remark 1. Whenever $t = as$, the criterion (1) reduces to $F\left(\frac{\Delta}{s} - a\right) \leq \delta$. When in addition $a = \infty$ and therefore $t = \infty$, then the mechanism is (ϵ, δ)-differentially private for all $\delta \geq 0$.

When $a < \infty$ and $\Delta \geq 2as$, we get that $f((x-\Delta)/s)/f(x/s) = 0$ for all $x \in (-as, as)$ and therefore the mechanism is not (ϵ, δ)-differentially private for any $\delta < 1$. ◁

The smallest finite scale s for which (1) holds is monotone in ϵ, δ, and Δ. We make this precise in the following.

Lemma 2. *Let $s(\epsilon, \delta, \Delta)$ be the smallest s such that (1) holds for log-concave and even density f and real valued function q on databases with global sensitivity $\Delta > 0$. Then $s(\epsilon, \delta, \Delta) = \Delta\, s(\epsilon, \delta, 1)$, and $s(\epsilon, \delta, \Delta)$ is non-increasing in both ϵ and δ.*

Unless otherwise evident, let in the following $s(\epsilon, \delta, \Delta)$ be defined as Lemma 2.

3.2 Selected analytic conditions

The following closed form condition is a tightening of Proposition 1 in Dwork et al. (2006b).

Theorem 3. *Let X be a standard Laplace variable, and let q be a real valued function on databases with global sensitivity Δ. Then the mechanism that returns a variate of $q(d) + sX$ is (ϵ, δ)-differentially private if and only if $s \geq s_1(\epsilon, \delta, \Delta) := \frac{\Delta}{\epsilon - 2\log(1-\delta)}$.*

Remark 2. An example of a mechanism that adds symmetric truncated log-concave noise is the truncated Laplacian mechanism (Geng et al., 2020) which Geng et al. use to establish upper and lower error bounds for (ϵ, δ)-differential mechanisms. Given a standard Laplace variable Y, this mechanism is constructed by truncating the support of $(\frac{\Delta}{\epsilon})Y$ to $(-A, A)$. The way this is done makes this mechanism meet the criterion (1) in Lemma 1 with equality. ◁

The following closed form condition is to our knowledge new.

Theorem 4. *Let X be a standard Logistic variable having density $f_{\log}(x) := \frac{e^{-x}}{(1+e^{-x})^2}$, and let q be a real valued function on databases with global sensitivity Δ. Then the mechanism that returns a variate of $q(d) + sX$ is (ϵ, δ)-differentially private for $\delta < 1$ if and only if $s \geq s_{\log}(\epsilon, \delta, \Delta) := \frac{\Delta}{2\log\left(\frac{e^{\frac{\epsilon}{2}} + \sqrt{\delta(e^\epsilon + \delta - 1)}}{1-\delta}\right)}$.*

Remark 3. We have that $s_1(\epsilon, 0, \Delta) = s_{\log}(\epsilon, 0, \Delta) = \frac{\Delta}{\epsilon}$, $s_1(0, \delta, \Delta) = \frac{\Delta}{2\log\left(\frac{1}{1-\delta}\right)}$, and $s_{\log}(0, \delta, \Delta) = \frac{\Delta}{2\log\left(\frac{1+\delta}{1-\delta}\right)}$. The Logistic mechanism serves as a second example besides the Laplace mechanism that can achieve ϵ-differential privacy. Both can also achieve $(0, \delta)$-differential privacy. Furthermore, letting $\epsilon_1(\epsilon, \delta) := \Delta/s_1(\epsilon, \delta, \Delta) = \epsilon - 2\log(1-\delta)$, and $\epsilon_{\log}(\epsilon, \delta) := \Delta/s_{\log}(\epsilon, \delta, \Delta) = 2\log\left(\frac{e^{\frac{\epsilon}{2}} + \sqrt{\delta(e^\epsilon + \delta - 1)}}{1-\delta}\right)$, we get that for given ϵ and δ, the Laplace and Logistic mechanisms are

$(\epsilon_1(\epsilon, \delta), 0)$- and $(\epsilon_{\log}(\epsilon, \delta), 0)$-differentially private, respectively. The difference between $\epsilon_1(\epsilon, \delta)$ and $\epsilon_{\log}(\epsilon, \delta)$ can be interpreted as a mechanisms sensitivity to δ. In practice, this translates to the change in scale as δ changes. ◁

The following condition appears in Theorem 8 in Balle and Wang (2018).

Theorem 5. *Let Z be a standard Gaussian variable, and let q be a real valued function on databases with global sensitivity Δ. Then the mechanism that returns a variate of $q(d) + \sigma Z$ is (ϵ, δ)-differentially private if and only if*

$$\Phi\left(\frac{\Delta}{2\sigma} - \frac{\epsilon\sigma}{\Delta}\right) - e^{\epsilon}\Phi\left(-\frac{\Delta}{2\sigma} - \frac{\epsilon\sigma}{\Delta}\right) \leq \delta. \tag{3}$$

3.3 A benefit of achieving ϵ-differential privacy

The well known Laplace and Gaussian mechanisms are both supported on $\mathbb{R}$, i.e., the noise densities are are supported everywhere, which for us means that $\psi : \mathbb{R} \to \mathbb{R}$ as the only time $e^{-\psi} = 0$ is if $\psi = \infty$. As we will see, whether a symmetric log-concave mechanism supported everywhere can achieve ϵ-differential privacy or not has consequences for its utility.

Remark 4. In this work, we rely heavily on that for log concave densities f, the likelihood ratio $f((z - \Delta)/s)/f(z/s)$ is monotone and non-decreasing in z. A mechanism that returns a variate $q(d) + sX$ where $X \sim f$ can achieve $(\epsilon, 0)$-differential privacy when the likelihood ratio above is bounded from above as z goes to infinity. The usual proof of ϵ-differential privacy for the Laplace mechanism essentially shows that for $s = \frac{\Delta}{\epsilon}$, this bound is e^{ϵ}. ◁

We now distinguish between those that are bounded and not bounded.

Definition 3 (MLR-boundedness)**.** Let $\psi : \mathbb{R} \to \mathbb{R}$ be convex and even, and let $f(x) = e^{-\psi(x)}$. If for all $\Delta > 0$ and $s > 0$

$$\lim_{z\to\infty} \log\left(\frac{f\left(\frac{z-\Delta}{s}\right)}{f\left(\frac{z}{s}\right)}\right) = \lim_{z\to\infty} \psi\left(\frac{z}{s}\right) - \psi\left(\frac{z-\Delta}{s}\right) = \infty,$$

we say that ψ and f are monotone likelihood ratio-unbounded (MLR-unbounded). If f is not MLR-unbounded it is MLR-bounded.

We can show the following.

Theorem 6. *Let X be a random variable distributed according to density $f(x) = e^{-\psi(x)}$ where $\psi : \mathbb{R} \to \mathbb{R}$ is convex, even, and MLR-unbounded. Then, $s \to \infty$ as $\delta \to 0$ for the (ϵ, δ)-differentially private mechanism that returns a variate of $q(d) + sX$.*

Remark 5. Let $M_{\delta=0}$ and $M_{\delta>0}$ be symmetric log-concave mechanisms with noise that is supported everywhere such that $M_{\delta=0}$ can achieve ϵ-differential privacy and $M_{\delta>0}$ can only achieve (ϵ, δ)-differential privacy for $\delta > 0$. If we fix ϵ, then what Theorem 6 means is that for any utility that is unbounded and strictly decreasing in the scale of the added noise, we can choose a small enough δ such that the utility of $M_{\delta=0}$ is better by an arbitrary amount. In this sense, Theorem 6 is a utility separation theorem for log-concave mechanisms. ◁

Specifically, for the Logistic and log-concave Subbotin densities,

$$\lim_{z\to\infty} \psi\left(\frac{z}{s}\right) - \psi\left(\frac{z-\Delta}{s}\right) = \begin{cases} \frac{\Delta}{s}, & \psi \in \{\psi_{\log}, \psi_1\}, \\ \infty, & \psi \in \{\psi_r \mid r > 1\} \end{cases}$$

for $s > 0$. This means that for fixed ϵ we can always choose δ small enough such that the Laplace and Logistic mechanisms have arbitrary smaller variances than mechanisms that add Subbotin$_r$ noise for $r > 1$, which includes the Gaussian at $r = 2$.

3.4 Mechanism variance ratios

We now empirically investigate the last paragraph in the previous Section. We choose the Laplace and Logistic mechanisms as representatives of MLR-bounded mechanisms and the popular Gaussian mechanism as the representative of the MLR-unbounded mechanisms that add Subbotin$_r$ noise.

Let $X_i \sim f_i$ be a random variable distributed according to symmetric log-concave density f_i, and let $Y_i = q(d) + s_i X_i$ for $s_i = s_i(\epsilon, \delta, \Delta)$. For two mechanisms Y_a and Y_b, let $v_{a,b} := \frac{\mathrm{Var}(Y_a)}{\mathrm{Var}(Y_b)} = \left(\frac{s_a}{s_b}\right)^2 \frac{\mathrm{Var}(X_a)}{\mathrm{Var}(X_b)}$. Now, $v_{a,b} \leq 1$ if and only if

$$\rho_{a,b} := \frac{s_a}{s_b} \leq \sqrt{\frac{\mathrm{Var}(X_b)}{\mathrm{Var}(X_a)}}. \tag{4}$$

From Lemma 2 we get that for any $\Delta > 0$

$$\rho_{a,b}(\epsilon, \delta) = \frac{s_a(\epsilon, \delta, \Delta)}{s_b(\epsilon, \delta, \Delta)} = \frac{s_a(\epsilon, \delta, 1)}{s_b(\epsilon, \delta, 1)}$$

for $a, b \in \{\log\} \cup \{r \mid r \geq 1\}$. From condition (4), we note that the curve given by $\rho_{a,b}(\epsilon, \delta) = \sqrt{\frac{\mathrm{Var}(X_b)}{\mathrm{Var}(X_a)}}$, acts as a separator on the (ϵ, δ)-space where one or the other mechanism exhibit smaller variance. We will now compare the selected mechanisms pairwise.

3.4.1 Logistic vs. Laplace mechanisms

For completeness, we first compare the MLR-bounded mechanisms.

From Theorems 3 and 4 we can state $\rho_{1,\log}(\epsilon, \delta)$ in closed form. We start by exploiting this to prove some specifics about $\rho_{1,\log}(\epsilon, \delta)$.

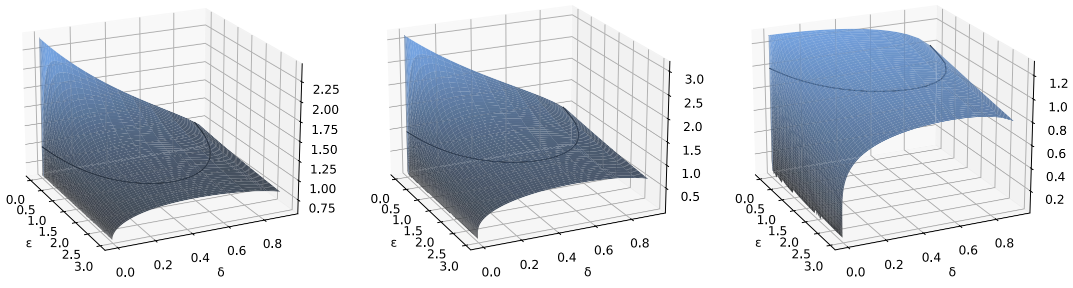
Theorem 7. *Let $\Delta > 0$, $\epsilon \geq 0$, $1 > \delta \geq 0$, and*

$$\rho_{1,\log}(\epsilon, \delta) := \frac{s_1(\epsilon, \delta, \Delta)}{s_{\log}(\epsilon, \delta, \Delta)}.$$

Then,

$$\rho_{1,\log}(\epsilon, \delta) = \frac{2\log\left(\sqrt{\delta}\sqrt{e^{\epsilon} + \delta - 1} + e^{\frac{\epsilon}{2}}\right) - 2\log(1-\delta)}{\epsilon - 2\log(1-\delta)},$$

and $2 > \rho_{1,\log}(\epsilon, \delta) \geq 1$ is sharp.



(a) $v_{1,\log}$ **(b)** $v_{1,2}$ **(c)** $v_{\log,2}$

Figure 1: For pairs of Logistic(log), Laplace (1), and Gaussian (2) distributions: plots of variance ratios $v_{a,b}$ and contours where the resulting variances are equal.

An immediate consequence is the following.

Corollary 1. *Let X_1 and $X_{\log}$ be random variables distributed according to f_1 and $f_{\log}$, respectively. Then,*

$$2.44 > \frac{24}{\pi^2} > \frac{\mathrm{Var}(s_1(\epsilon,\delta,\Delta)X_1)}{\mathrm{Var}(s_{\log}(\epsilon,\delta,\Delta)X_{\log})} \geq \frac{6}{\pi^2} > 0.6,$$

and

$$\frac{\mathrm{Var}(s_1(\epsilon,\delta,\Delta)X_1)}{\mathrm{Var}(s_{\log}(\epsilon,\delta,\Delta)X_{\log})} \leq 1$$

when $\rho_{1,\log}(\epsilon,\delta) \leq \sqrt{\frac{\pi^2}{6}}$.

Remark 6. Combining Theorem 7 with Remark 3, we see that $\epsilon_1(\epsilon,\delta) \leq \epsilon_{\log}(\epsilon,\delta)$. ◁

A plot of $v_{1,\log}(\epsilon,\delta)$ can be seen in Figure 1a. The plot shows the contour line where $v_{1,\log}(\epsilon,\delta) = 1$. In the ϵ and δ region where $v_{1,\log}(\epsilon,\delta) < 1$, the Laplace mechanism exhibits smaller variance than the Logistic mechanism. From Corollary 1, we have that this region is determined by $\rho(\epsilon,\delta) < \sqrt{\frac{\pi^2}{6}}$. As an example, for $\delta \leq 10^{-4}$, the Laplace mechanism variance is smaller as long as $\epsilon \geq 0.0047$.

3.4.2 Laplace and Logistic vs. Gaussian mechanisms

When comparing Laplace and Logistic mechanisms above, we were able to take advantage of s_1 and $s_{\log}$ having closed forms, and consequently also $\rho_{1,\log}$ having a closed form. Unfortunately, we only have access to values for the Gaussian scale s_2 through numeric computations. The plots of $v_{1,2}(\epsilon,\delta)$ and $v_{\log,2}(\epsilon,\delta)$ with the respective unit contours are given in Figure 1b and Figure 1c, respectively. The regions where the Laplace and Logistic exhibit lower variance are given by $\rho_{1,2}(\epsilon,\delta) < \sqrt{\frac{1}{2}}$ and $\rho_{\log,2}(\epsilon,\delta) < \frac{\sqrt{3}}{\pi}$, respectively. In the plots, these are the regions where the respective surfaces have values smaller than 1.

Numeric computations yield that for $\epsilon \geq 0.05$, the Laplace mechanism yields smaller mean squared error than both the Logistic and the Gaussian mechanisms as long as $\delta \leq 0.001$, and the Logistic mechanism yields smaller squared error than the Gaussian mechanism as long as $\delta \leq 0.002$.

The variance is unbounded and strictly increasing in the scale for the Logistic distribution as well as all Subbotin$_{r\geq 1}$ distributions. From Section 3.3 we have that since both the Laplace and Logistic noise is MLR-bounded, while the Gaussian is MLR-unbounded, the ratios $v_{1,2}$ and $v_{\log,2}$ go to 0 as $\delta \to 0$ for any ϵ. Furthermore, for all symmetric densities, we have that the mechanism that add noise according to these exhibit a mean squared error (MSE) that is equivalent to the variance. Consequently, the variance comparisons above hold for the MSE as well. This is particularly relevant for histogram queries, since each individual histogram count can be computed using a one-dimensional mechanism.

4 THE MULTIDIMENSIONAL CASE

So far, we discussed (ϵ, δ)-differential privacy for real-valued database queries q, i.e., $q(d) \in \mathbb{R}$ for database d. In order to generalize our results regarding one-dimensional mechanisms to queries in $\mathbb{R}^n$, we generalize a symmetric log-concave density f to a $\|\cdot\|$-spherically symmetric log-concave density $Cf(\|x\|)$ for a normalization constant $C \in \mathbb{R}$. For this generalization, the isodensity surfaces are spherical with respect to the norm $\|\cdot\|$. If $n = 1$, this density coincides with f and we will simply write f also for $n > 1$.

Furthermore, in our above discussion of adding real-valued noise, a critical parameter was the global sensitivity Δ. This maximal change was defined in terms of the absolute value $|\cdot|$. In the multidimensional case we generalize this as

$$\Delta := \max_{(d,d')\in\mathcal{N}} \|q(d) - q(d')\|$$

where $\|\cdot\|$ is a norm in $\mathbb{R}^n$.

Our second main result comes in two parts. The first of these, Lemma 8 below, essentially extends Lemma 1 to the case of $\|\cdot\|$-spherically symmetric log-concave densities when the global sensitivity is defined using the same norm $\|\cdot\|$. The second part, Theorem 9 below, in turn specializes Lemma 8 to a case where the noise can be represented by a vector of iid random variables.

Lemma 8. *Let $\mathbf{X} \in \mathbb{R}^n$ for $n \geq 1$ be distributed according to $\|\cdot\|$-spherically symmetric log-concave density f. Then, if q and w are $\mathbb{R}^n$ and $\mathbb{R}$ valued functions on databases, respectively, both with global sensitivity $\Delta = \max_{(d,d')\in\mathcal{N}} \|q(d) - q(d')\| = \max_{(d,d')\in\mathcal{N}} |w(d) - w(d')|$, then for $s > 0$, the mechanism $q(d) + s\mathbf{X}$ is (ϵ, δ)-differentially private if and only if the mechanism $w(q) + sY$ for $Y \sim f$ is.*

Let $\|(v_1, v_2, \ldots, v_n)\|_p := (\sum_{i=1}^n |v_i|^p)^{1/p}$ for $p \in \mathbb{R}$, $p \geq 1$ denote the p-norm on $\mathbb{R}^n$. Using particulars of the multivariate Gaussian density, Balle and Wang (2018) show that for Δ defined in terms of the Euclidean norm $\|\cdot\|_2$, the condition (3) is sufficient and necessary for the mechanism that adds $s(Z_1, Z_2, \ldots, Z_n)$ where the Z_i are independent standard Gaussian variables. Osiewalski and Steel (1993) show that when the independent $X_i \sim f_p$, where f_p is the Subbotin$_p$ density, the variable $(X_1, X_2, \ldots, X_n)$ has a $\|\cdot\|_p$-spherical distribution. Using this result, we specialize Lemma 8 to pairings of Subbotin$_p$ variables and p-norms for $p \geq 1$ in the following Theorem.

Theorem 9. *Let $n \in \mathbb{N}$, $n \geq 1$, $p \geq 1$, $X_1, \ldots, X_n$ be independent Subbotin$_p$ random variables, and $\mathbf{X} = (X_1, X_2, \ldots, X_n)$. Then for a $\mathbb{R}^n$-valued function q on databases with global sensitivity $\Delta = \max_{(d,d') \in \mathcal{N}} \|q(d) - q(d')\|_p$ and a database d, the mechanism returning a variate of $q(d) + s\mathbf{X}$ is (ϵ, δ)-differentially private if and only if*

$$F_p\left(\frac{\Delta - t}{s}\right) - e^{\epsilon} F_p\left(-\frac{t}{s}\right) \leq \delta$$

where

$$t := \sup\left\{z \in \mathbb{R} \mid \psi_p\left(\frac{z}{s}\right) - \psi_p\left(\frac{z - \Delta}{s}\right) \leq \epsilon\right\}.$$

Furthermore, if the above holds for scale $s > 0$, then it also holds for scales $s' > s$.

We will call the mechanism from Theorem 9 the *Subbotin$_r$ mechanism* and let $s_r(\epsilon, \delta, \Delta)$ denote the minimum scale s for which the condition in Theorem 9 holds for the Subbotin$_r$ mechanism.

Remark 7. Combining Theorem 9 with Theorem 3, we get that the vector Laplace mechanism $q(d) + (sX_1, sX_2, \ldots, sX_n)$ where the independent $X_i \sim f_1$ and the global sensitivity Δ is defined in terms of $\|\cdot\|_1$ is (ϵ, δ)-differentially private if and only if

$$s \geq \frac{\Delta}{\epsilon - 2\log(1 - \delta)}.$$ ◁

4.1 Optimizing query responses using subbotin noise

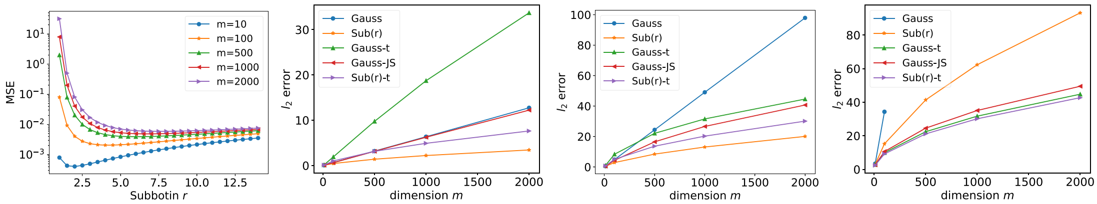
Let R be a set of record values let $q : R^n \to \mathbb{R}^m$ be a query function on length n databases d. An important problem is to minimize the MSE $m^{-1} \operatorname{E}[\|M(d) - q(d)\|_2^2]$ for multidimensional mechanisms M. Given a choice between different mechanisms, a question is then which one minimizes the MSE. We now look at this question for the family of Subbotin$_r$ mechanisms M_r. Let Δ_r denote the global sensitivity of q defined using the r-norm. Then, using definitions from Theorem 9, $m^{-1} \operatorname{E}[\|M_r(d) - q(d)\|_2^2] = m^{-1} \operatorname{E}[\|s\mathbf{X}\|_2^2] = s^2 \operatorname{Var}(X)$ for $X \sim f_r$ and $s = s_r(\epsilon, \delta, \Delta_r)$. We can minimize this over the choices of r by letting

$$r = \arg\min_{p \geq 1} s_p(\epsilon, \delta, \Delta_p)^2 \left(p^{(2/p)} \frac{\Gamma(3/p)}{\Gamma(1/p)}\right). \tag{5}$$

The optimization (5) requires determining Δ_r. This can be difficult. For particular classes of queries, we can make it a little easier. We call query function q *linear* if there exist linear function $f : R \to \mathbb{R}^m$ and function $\nu : \mathbb{N} \to \mathbb{R}$ such that $q(r_1, r_2, \ldots, r_n) = \nu(n) \sum_{i=1}^n f(r_i)$. To clarify which f and ν we associate with linear q, we write $q_{\nu,f}$. Let $\|S\|_\infty := \sup\{\|x - y\|_\infty \mid x, y \in S\}$ denote the l_∞ diameter of a set $S \subseteq \mathbb{R}^m$. Then, if $\|f(R)\|_\infty$ is easier to determine or bound than Δ_r the following can help.

Theorem 10. *Let $q_{\nu,f} : R^n \to \mathbb{R}^m$ be a linear database query function for some integers $m > 0$ and $n > 0$. Then $\Delta_p \leq m^{1/p} |\nu(n)| \, \|f(R)\|_\infty$ with equality if $f(R)$ is an l_∞-ball.*

A natural evaluation of optimizing (5) is to compare with the go-to Gaussian mechanism. We therefore essentially replicate the estimation of the mean m-dimensional vector experiment performed by Balle and Wang (2018). We also chose this experiment because Balle and Wang use this to evaluate two de-noising post-processing methods for the Gaussian mechanism that we also want to compare with.



(a) Mean squared errors for $\epsilon = 1$ **(b)** l_2 errors for $\epsilon = 1$ **(c)** l_2 errors for $\epsilon = 0.1$ **(d)** l_2 errors for $\epsilon = 0.01$

Figure 2: Graphical summary of the mean estimation experiments.

Let the mean vector query be the linear query $q = q_{\nu,f}$ for $f(x) := x$ and $\nu(n) := 1/n$. Suppose y is the output of the Gaussian mechanism with scale s for query q. Then, the first of the de-noising methods is the adaptive James-Stein estimator $y_{JS} := \left(1 - \frac{(m-2)s^2}{\|y\|_2^2}\right)$, with the underlying assumption that the query response is a random Gaussian vector with iid dimensions. The second is the soft-thresholding proposed by Donoho et al. (1994; 1995) given by $y_{th} := \text{sign}(y)\max(0, |y| - t)$ where $t := s\sqrt{2\log(m)}$. For this method, ideally t is chosen such that it defines an ∞-ball containing the noise. We recognize t as a standard upper bound for the expected maximum of m iid scale s Gaussian random variables, so the defined ∞-ball contains the noise in expectation. We adapt the soft thresholding to scale s Subbotin$_r$ mechanism output by letting $t = s/300 \sum_{i=1}^{300} \max(x_i)$ where the m elements in $x_i \in \mathbb{R}^m$ are independent standard Subbotin$_r$ variates.

Like Balle and Wang, we require that records lie in the unit l_∞-ball $R = v + [-1/2, 1/2]$, where the v_i in $v = (v_1, v_2, \ldots, v_m)$ are standard normal variates, let databases be length $n = 500$ sequences of elements from R chosen uniformly at random, and performed our experiments with fixed $\delta = 10^{-4}$ and errors measured by the l_2-error. For each pair (ϵ, m) in $\{0.01, 0.1, 1\} \times \{10, 100, 500, 1000, 2000\}$ we numerically optimized (5) on a regular grid $(1, 1.5, 2, \ldots, 14)$ yielding $r_{\epsilon,m}$ and generated 100 databases. For each of the databases from we then computed the l_2-error of the Gaussian mechanism (Gauss), the Subbotin$_{r_{\epsilon,m}}$ mechanism (Sub(r)), their soft-thresholded versions (Gauss-t, Sub(r)-t), the James-Stein denoised Gaussian mechanism (Gauss-JS), after which we computed the average for each of the five mechanisms for the 100 databases. These averages can be seen in Figures 2b – 2d. In Figure 2d the Gauss errors were much larger than all the others and were truncated to allow details to be seen for the other error series. Figure 2a shows the mean squared errors $s^2\,\text{Var}(X)$ computed during the optimization of (5) for $\epsilon = 1$ and all values m.

Since the optimal scale for the Subbotin$_r$ mechanisms is generally not available analytically, neither is the optimal r as a function of m. Therefore, we optimized r numerically on a grid of r values as described above. The maximum grid value $r = 14$ was chosen to contain a local optimum for all m, which are listed below. The grid points are shown as dots in Figure 2a.

For $\epsilon = 1$, $\epsilon = 0.1$, and $\epsilon = 0.01$, the selected $r_{\epsilon,m}$ values were (2,4,6,7,7.5), (2.5, 5, 7.5, 8.5, 9), and (3.5, 7, 10.5, 11.5, 13), respectively. The corresponding Subbotin$_r$ scales rounded to two decimal places were (0.02, 0.06, 0.08, 0.09, 0.10), (0.16, 0.37, 0.52, 0.58, 0.63), and (1.14, 2.07, 2.63, 2.84, 3.04), respectively. The corresponding Gaussian mechanism scales were (0.02, 0.06, 0.14, 0.20, 0.28), (0.16, 0.49, 1.10, 1.55, 2.19), and (1.09, 3.45, 7.72, 10.91, 15.44), respectively.

We make the following observations. Optimizing (5) is relevant for minimizing the l_2-error because of the monotone relationship between the MSE and the l_2-error. Doing so consistently produced smaller l_2-errors over the Gaussian mechanism and its denoised versions. It is well known that $\frac{m}{\sqrt{m+1}} \leq \mathrm{E}[\|\mathbf{Z}\|_2] \leq \sqrt{m}$ for a length m vector $\mathbf{Z}$ of iid standard Gaussian random variables. Also knowing that $\Delta \propto \sqrt{m}$, we can apply Lemma 2 to explain why the l_2-error of the Gaussian mechanism in our experiments appears linear in the query dimension m. This application of Lemma 2 may be of independent interest. The James-Stein estimator always improved on the unprocessed Gaussian mechanism (we refer the reader to Balle and Wang (2018) for an analysis). The soft-thresholded mechanisms were better than their unprocessed counterparts when the scale was large, while the opposite was true when the scale was smaller. Empirically, for Gaussian mechanism in our experimental setting, the changeover happens close to scale $s = 1$, for which the v and noise distributions are equal. When the scale becomes large as ϵ becomes small and m large, almost all of the entries in the thresholded estimates are 0, which suggests why the errors of the thresholded Gaussian and Subbotin$_{r_{\epsilon,m}}$ mechanisms for $\epsilon = 0.01$ are very similar and their comparison might be less informative. Also, when considering the results, it might be helpful to remember that we chose the experimental setup to favor the James-Stein estimator for Gaussian noise.

5 BACKGROUND AND OTHER RELATED WORKS

Adding randomness to the computation of query results in order to protect privacy has a long history (see e.g., Dalenius (1978), Section 13 or Denning (1980)). The definitions of privacy in these works differed and were often concerned with disallowing the variance in the query answers resulting from the randomness to be reduced too much by repeated queries. Following work by Dinur and Nissim (2003), noting that a "succinct catch-all definition of privacy is very elusive", Dwork et al. published (2006b) a definition of privacy that would become known as ϵ-differential privacy. This definition has three important properties: quantification of a strong form of privacy, can be composed across queries types and databases, and is immune to post-processing. The latter meaning that no amount of processing after the fact will change the guarantees. The definition of ϵ-differential privacy was soon extended to (ϵ, δ)-differential privacy (Dwork et al., 2006a), which allowed the use of, among others, output perturbation by Gaussian noise, and turned out to improve the compositional properties (Dwork et al., 2010; Kairouz et al., 2017) and thereby the utility–privacy trade-off over multiple queries. However, the extension to (ϵ, δ)-differential privacy represents a weakening of the privacy guaranteed over pure ϵ-differential privacy that is not uncontroversial (McSherry, 2017). In a machine learning context, compositional properties received much attention for establishing end-to-end privacy, particularly when using tools for private optimization (Bassily et al., 2014; Abadi et al., 2016). Partially as a result of this, several similar relaxations of ϵ-differential privacy have been proposed that improve compositional properties over their predecessors. These include Concentrated differential privacy (Dwork and Rothblum, 2016; Bun and Steinke, 2016), Rényi differential privacy (Mironov, 2017), and Gaussian differential privacy (Dong et al., 2019). Similarly to the work of Balle et al. (2018), our work can be seen as complementary to these efforts in that we concentrate on optimizing privacy parameters for *single* applications of a particular class of mechanisms under (ϵ, δ)-differential privacy. Liu (2019) defines the Generalized Gaussian mechanism that is effectively equivalent to the Subbotin$_r$ mechanism. However, only a sufficient criterion for (ϵ, δ)-differential privacy for this family is developed. For the Gaussian case, this criterion reduces to a criterion inferior to the one described by Le Ny and Pappas (2014). They also do not consider a systematic search among members in this family for optimizing utility. Optimality with respect to different utilities has been shown for ϵ-differentially private mechanisms such as the staircase mechanism (Geng et al., 2015) and randomized response in a multi-party setting (Kairouz et al., 2015).

6 SUMMARY AND DISCUSSION

We provided a necessary and sufficient condition for (ϵ, δ)-differential privacy for mechanisms that add noise from a symmetric and log-concave distribution (Lemma 1). This condition is given directly in terms of the standard distribution function, the needed scale of the distribution, ϵ, δ, and the global sensitivity Δ of the query in question. Previously, such a condition was only known for the Gaussian distribution (Balle and Wang, 2018) and was shown to be essentially identical for the discrete Gaussian distribution (Canonne et al., 2020).

Importantly, we show that our necessary and sufficient condition extends naturally to multidimensional query responses. Key to this extension is that the noise distribution is spherically symmetric log-concave and is spherical with respect to the norm used to define the query global sensitivity. We are able to show that this holds for mechanisms where the noise vector consists of independent Subbotin$_r$ random variables and the query sensitivity is assessed using the corresponding l_r-norm for $r \geq 1$ (Theorem 9). This infinite family of Subbotin$_r$ mechanisms contains the prototypical Laplace and Gaussian mechanisms with $r = 1$ and $r = 2$, respectively.

In the one-dimensional query case, we show that for symmetric log-concave mechanisms supported everywhere that cannot achieve pure ϵ-differential privacy, the scale grows to infinity as δ goes to 0 (Theorem 6). Since this is not the case when the mechanism can achieve ϵ-differential privacy, we achieve a separation of these two types of symmetric log-concave mechanism for any utility that is decreasing and unbounded in the noise scale. We demonstrate this utility separation by showing that the Laplace and Logistic mechanisms exhibits smaller variance than the Gaussian mechanism for a large and relevant range of values for ϵ and δ.

We further showed that the optimal scale for log-concave noise mechanisms is proportional to the query sensitivity Δ (Lemma 2). This means that the optimal noise scale is proportional to the behaviour of the norm used for the query sensitivity. As a consequence of this and having access to a Subbotin$_r$ mechanism for every l_r-norm for $r \geq 1$, we were able to demonstrate that optimizing the choice of r for the multidimensional case can significantly improve the mean squared error and any similar utility such as the l_2-error over any fixed choice of mechanism such as the Gaussian.

Consider that a high-dimensional random vector $\mathbf{X}$ with iid components appears to be concentrated on a sphere (see e.g., Biau and Mason (2013)), a phenomenon that Dong et al. (2021) state as that the mechanism adding $\mathbf{X}$ essentially behaves like a Gaussian mechanism. So, for a fixed high-dimensional query and given privacy parameters ϵ and δ, two different Subbotin mechanisms can essentially behave like two Gaussian mechanisms that differ in the noise scale. This suggests a "Gaussian lens" through which to interpret a choice of norm. However, such an interpretation cannot be strict: even though the Laplace mechanism essentially behaves like a Gaussian mechanism in high dimensions, it differs in the type of privacy guarantee it affords since the Gaussian mechanism cannot achieve $(\epsilon, 0)$-differential privacy for any ϵ.

Dwork and Roth (2014) warn that claims of differential privacy should be seen in relation to the granularity of data, essentially reflected in the neighborhood relation used to define privacy. Similarly, the choice of norm affects sensitivity and the warning could be considered to apply here as well, particularly in light of the preceding paragraph. We recommend taking this into account when, for example, considering the range of r to optimize Subbotin mechanisms over.

As mentioned in Section 5, other members in the family of differential privacy definitions exhibit better composition properties and can be attractive alternatives to (ϵ, δ)-differential privacy. They have in common that the query sensitivity is defined by some norm and that an increase in the sensitivity yields an increase in the scale of the noise. For example, in the Concentrated differential privacy case, the sensitivity enters the scale of the Gaussian mechanism as a multiplicative factor for a given level of privacy. Furthermore, the extension of a scale bound for univariate unimodal and symmetric distributions to the multidimensional spherical case appears quite general. Therefore, we consider our results regarding Subbotin$_r$ mechanisms and optimizing the choice of distribution in terms of query response dimensionality as an example of how to implement a general method under (ϵ, δ)-differential privacy.

References

Abadi, M., Chu, A., Goodfellow, I., McMahan, H. B., Mironov, I., Talwar, K., and Zhang, L. (2016). Deep Learning with Differential Privacy. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS'16*, pages 308–318.

Anderson, G., Vamanamurthy, M., and Vuorinen, M. (2006). Monotonicity rules in calculus. *The American Mathematical Monthly*, 113(9):805–816.

Balle, B. and Wang, Y.-X. (2018). Improving the Gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. volume 80 of *Proceedings of Machine Learning Research*, pages 394–403, Stockholmsmässan, Stockholm Sweden. PMLR.

Bassily, R., Smith, A., and Thakurta, A. (2014). Private empirical risk minimization: Efficient algorithms and tight error bounds. In *Proceedings of the 2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, FOCS '14, pages 464–473, USA. IEEE Computer Society.

Biau, G. and Mason, D. D. (2013). High-dimensional p-norms.

Bun, M. and Steinke, T. (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Proceedings, Part i, of the 14th International Conference on Theory of Cryptography - Volume 9985*, pages 635–658, Berlin, Heidelberg. Springer-Verlag.

Canonne, C. L., Kamath, G., and Steinke, T. (2020). The discrete gaussian for differential privacy. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. F., and Lin, H., editors, *Advances in Neural Information Processing Systems*, volume 33, pages 15676–15688. Curran Associates, Inc.

Dalenius, T. (1978). *Information Privacy and Statistics : A Topical Bibliography / by Tore Dalenius*. Washington, D.C. : U.S. Dept. of Commerce, Bureau of the Census.

Denning, D. E. (1980). Secure Statistical Databases with Random Sample Queries. *ACM Transactions on Database Systems*, 5(3):291–315.

Dinur, I. and Nissim, K. (2003). Revealing information while preserving privacy. In *PODS '03: Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems*, pages 202–210, New York, NY, USA. ACM.

Dong, J., Roth, A., and Su, W. J. (2019). Gaussian differential privacy. *CoRR*, abs/1905.02383.

Dong, J., Su, W., and Zhang, L. (2021). A central limit theorem for differentially private query answering. In *Advances in Neural Information Processing Systems*, volume 34.

Donoho, D. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3):613–627.

Donoho, D. L. and Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3):425–455.

Dwork, C., Kenthapadi, K., McSherry, F., Mironov, I., and Naor, M. (2006a). Our Data, Ourselves: Privacy Via Distributed Noise Generation. In *Advances in Cryptology (EUROCRYPT 2006)*, volume 4004, pages 486–503, Saint Petersburg, Russia. Springer Verlag.

Dwork, C. and Lei, J. (2009). Differential privacy and robust statistics. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, STOC '09, pages 371–380, New York, NY, USA. Association for Computing Machinery.

Dwork, C., McSherry, F., Nissim, K., and Smith, A. (2006b). Calibrating Noise to Sensitivity in Private Data Analysis. In *Proceedings of the Conference on Theory of Cryptography*.

Dwork, C. and Roth, A. (2014). The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407.

Dwork, C. and Rothblum, G. N. (2016). Concentrated differential privacy. *CoRR*, abs/1603.01887.

Dwork, C., Rothblum, G. N., and Vadhan, S. (2010). Boosting and Differential Privacy. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 51–60.

Elkins, Z., Ginsburg, T., and Melton, J. (2013). Constitute: The worlds constitutions to read, search, and compare.

Geng, Q., Ding, W., Guo, R., and Kumar, S. (2020). Tight analysis of privacy and utility tradeoff in approximate differential privacy. In Chiappa, S. and Calandra, R., editors, *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics*, volume 108 of *Proceedings of Machine Learning Research*, pages 89–99. PMLR.

Geng, Q., Kairouz, P., Oh, S., and Viswanath, P. (2015). The staircase mechanism in differential privacy. *IEEE Journal of Selected Topics in Signal Processing*, 9(7):1176–1184.

Kairouz, P., Oh, S., and Viswanath, P. (2015). Secure Multi-party Differential Privacy. In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems 28*, pages 2008–2016. Curran Associates, Inc.

Kairouz, P., Oh, S., and Viswanath, P. (2017). The composition theorem for differential privacy. *IEEE Transactions on Information Theory*, 63(6):4037–4049.

Le Ny, J. and Pappas, G. J. (2014). Differentially private filtering. *IEEE Transactions on Automatic Control*, 59(2):341–354.

Liu, F. (2019). Generalized gaussian mechanism for differential privacy. *IEEE Transactions on Knowledge and Data Engineering*, 31(4):747–756.

McSherry, F. (2017). How many secrets do you have?

Mironov, I. (2017). Renyi Differential Privacy. *2017 IEEE 30th Computer Security Foundations Symposium (CSF)*, pages 263–275.

Nadarajah, S. (2005). A generalized normal distribution. *Journal of Applied Statistics*, 32(7):685–694.

Osiewalski, J. and Steel, M. F. J. (1993). Robust bayesian inference in lq-Spherical models. *Biometrika*, 80(2):456–460.

Saumard, A. and Wellner, J. A. (2014). Log-concavity and strong log-concavity: A review.

Subbotin, M. (1923). On the Law of Frequency of Error. *Mat. Sb.*, 31(2):296–301.

The United Nations (1948). *Universal Declaration of Human Rights*.

A POSTPONED PROOFS

After recapitulating some known material, we present postponed proofs section-wise.

A.1 Important known results we use

For completeness, we include the proofs of the following known results we rely on. Their presentation is tailored to our needs.

We will be using the following result repeatedly. The formulation and proof is taken from Saumard and Wellner (2014).

Lemma 11 (Monotone likelihood ratio)**.** *A density f on $\mathbb{R}$ is log-concave if and only if the translation family $\{f(\cdot - \theta) \mid \theta \in \mathbb{R}\}$ has monotone likelihood ratios: i.e., for every $\theta \leq \theta'$ the ratio $f(x-\theta')/f(x-\theta)$ is a monotone nondecreasing function of x.*

Proof of Lemma 11. First note that for all $x < x'$ and $\theta < \theta'$

$$\frac{f(x-\theta')}{f(x-\theta)} \leq \frac{f(x'-\theta')}{f(x'-\theta)} \tag{6}$$

iff

$$\log f(x-\theta') + \log f(x'-\theta) \leq \log f(x'-\theta') + \log f(x-\theta). \tag{7}$$

Let $t = (x'-x)/(x'-x+\theta'-\theta)$ and note that

$$\begin{aligned} x-\theta &= t(x-\theta') + (1-t)(x'-\theta) \\ x'-\theta' &= (1-t)(x-\theta') + t(x'-\theta). \end{aligned}$$

Log-concavity of f implies that

$$\begin{aligned} \log f(x-\theta) &\geq t\log f(x-\theta') + (1-t)\log f(x'-\theta) \\ \log f(x'-\theta') &\geq (1-t)\log f(x-\theta') + t\log f(x'-\theta). \end{aligned}$$

Adding these yields (7), and we can conclude that f being log-concave implies (6).

Now, suppose that (6) holds. Then, (7) holds, and does so in particular if x, x', θ, θ' satisfy $x-\theta' = a < b = x'-\theta$ and $t = (x'-x)/(x'-x+\theta'-\theta) = 1/2$, so that $x-\theta = (a+b)/2 = x'-\theta'$. Then (7) becomes

$$\log f(a) + \log f(b) \leq 2\log f((a+b)/2).$$

This, together with measurability of f, implies that f is log-concave. □

Lemma 12 (Translation invariance)**.** *Let X be a random variable taking values in $\mathbb{R}^n$ for natural number $n \geq 1$. For arbitrary but fixed $x, y, z \in \mathbb{R}^n$ we have that $\Pr(X + (x-z) \in S) \leq e^\epsilon \Pr(X + (y-z) \in S) + \delta$ for all measurable $S \subseteq \mathbb{R}^n$ implies $\Pr(X+x) \in S) \leq e^\epsilon \Pr(X+y \in S)+\delta$ for all measurable $S \subseteq \mathbb{R}^n$.*

Proof of Lemma 12. Follows directly from that if $S \subseteq \mathbb{R}^n$ is measurable, so is $r + S$ for any $r \in \mathbb{R}^n$, including $r = -z$. □

Lemma 13 (Symmetry of distance)**.** *Let f be an even density and let $X \sim f$, let $\mathcal{S}$ be the measureable sets $S \subseteq \mathbb{R}^n$ for some positive integer n, $d \geq 0$, $\epsilon \geq 0$, $\delta \geq 0$. Then*

$$(\forall S \in \mathcal{S}) \Pr(d + X \in S) - e^{\epsilon} \Pr(X \in S) \leq \delta$$
$$\Longleftrightarrow$$
$$(\forall S \in \mathcal{S}) \Pr(-d + X \in S) - e^{\epsilon} \Pr(X \in S) \leq \delta.$$

Proof of Lemma 13. Note that for $S \in \mathcal{S}$

$$S \in \mathcal{S} \iff -S \in \mathcal{S} \tag{8}$$
$$x \in S \iff -x \in -S \tag{9}$$
$$X \stackrel{d}{=} -X \tag{10}$$

where $\stackrel{d}{=}$ means equal in distribution. Now,

$$(\forall S \in \mathcal{S}) \Pr(d + X \in S) - e^{\epsilon} \Pr(X \in S) \leq \delta$$
$$\stackrel{\text{(by (8))}}{\iff}$$
$$(\forall S \in \mathcal{S}) \Pr(d + X \in -S) - e^{\epsilon} \Pr(X \in -S) \leq \delta$$
$$\stackrel{\text{(by (9))}}{\iff}$$
$$(\forall S \in \mathcal{S}) \Pr(-(d + X) \in S) - e^{\epsilon} \Pr(-X \in S) \leq \delta$$
$$\stackrel{\text{(by (10))}}{\iff}$$
$$(\forall S \in \mathcal{S}) \Pr(-d + X \in S) - e^{\epsilon} \Pr(X \in S) \leq \delta. \qquad \square$$

The result below appears in the proof of Theorem 5 in Balle and Wang (2018).

Lemma 14 (Removal of quantifier)**.** *Let random variables Y_+ and Y_- taking values from $\mathbb{R}^n$ be distributed according to densities f_+ and f_-, respectively. Let $\mathcal{S}$ denote the collection of all measurable sets of elements in V, and let $A^c = \{x \mid f_+(x) > e^{\epsilon} f_-(x)\}$. Then,*

$$\forall S \in \mathcal{S} \;\; \Pr(Y_+ \in S) - e^{\epsilon} \Pr(Y_- \in S) \leq \delta$$
$$\iff$$
$$\Pr(Y_+ \in A^c) - e^{\epsilon} \Pr(Y_- \in A^c) \leq \delta.$$

Proof of Lemma 14. Let $A = \{x \mid f_+(x) \leq e^{\epsilon} f_-(x)\}$ and note that A^c is the complement of A. Then,

$$\begin{aligned}
\Pr(Y_+ \in S \cap A) &= \int_{S \cap A} f_+(x) dx \\
&\leq \int_{S \cap A} e^{\epsilon} f_-(x) dx \\
&= e^{\epsilon} \Pr(Y_- \in S \cap A).
\end{aligned}$$

Using this, we get

$$
\begin{aligned}
&\Pr(Y_+ \in S) - e^\epsilon \Pr(Y_- \in S) \\
&\quad = \Pr(Y_+ \in S \cap A) + \Pr(Y_+ \in S \cap A^c) \\
&\qquad - e^\epsilon \left(\Pr(Y_- \in S \cap A) + \Pr(Y_- \in S \cap A^c)\right) \\
&\quad \le \Pr(Y_+ \in S \cap A^c) - e^\epsilon \Pr(Y_- \in S \cap A^c) \\
&\quad \le \int_{S \cap A^c} \left(f_+(x) - e^\epsilon f_-(x)\right) dx \\
&\quad \le \int_{A^c} \left(f_+(x) - e^\epsilon f_-(x)\right) dx \\
&\quad = \Pr(Y_+ \in A^c) - e^\epsilon \Pr(Y_- \in A^c).
\end{aligned}
$$

The last inequality comes from $f_+(x) > e^\epsilon f_-(x)$ for $x \in A^c$, so $f_+(x) - e^\epsilon f_-(x) \ge 0$ for $x \in A^c$.

Since $\Pr(Y_+ \in S) - e^\epsilon \Pr(Y_- \in S) \le \Pr(Y_+ \in A^c) - e^\epsilon \Pr(Y_- \in A^c)$ the Lemma follows from $\Pr(Y_+ \in A^c) - e^\epsilon \Pr(Y_- \in A^c) \le \delta$ implies $\Pr(Y_+ \in S) - e^\epsilon \Pr(Y_- \in S) \le \delta$. □

A.2 Proofs from Section 3.1

This section contains the proof of the first of our main results given in Lemma 1 regarding the privacy of the mechnism $M(d) = q(d) + sX$.

The proof can be outlined as follows. Letting $\mathcal{S}$ denote the measureable sets, the first step in the proof is the removal of the quantifier over $\mathcal{S}$ in the privacy criterion

$$
(\forall S \in \mathcal{S}) \Pr(M(d) \in S) - e^\epsilon \Pr(M(d') \in S) \le \delta \tag{11}
$$

given by Lemma 14 above. This removal shows that that there exists a partition of the sample space into regions A and A^c such that

$$
\Pr(M(d) \in S) > e^\epsilon \Pr(M(d') \in S)
$$

if and only if $S \cap A^c \neq \emptyset$. The next step uses log-concavity of the density f of X, particularly the monotonicity of likelihood ratios in the translation family of f given in Lemma 11, to show in Lemma 17 that we can treat A^c as an interval (t, ∞) where threshold t depends on ϵ, s and $d = |q(d) - q(d')|$. This allows stating the criterion (11) above in terms of

$$
\Pr(d + sX > t) - e^\epsilon \Pr(sX > t) \le \delta,
$$

which is equivalent to the statement in Lemma 1. The proof of Lemma 1 is completed by showing monotonicity of the right hand side in the above display in Lemma 18.

The above steps are shown using elementary means and their demonstrations are relatively self-contained, which we consider a benefit. Alternatively, machinery based on conjugate duality and hypothesis testing as employed by Dong et al. (2019) (possibly using a combination of Propositions A.3 and 2.12 as a starting point) could be employed. Our elementary analysis, particularly the monotonicity Lemma 18, suggests a strategy for numerically computing scales for given privacy parameters.[1] Furthermore, it allows for the simple reduction to the one dimensional case employed in the proof of the extension to the multidimensional case in Lemma 8.

[1]Implemented in https://github.com/laats/SubbotinMechanism

Remark 8 (Setup). Throughout this section, we let $f(x) = e^{-\psi(x)}$ be a probability density where $\psi : \mathbb{R} \to (-\infty, \infty]$ is convex and even, which we can take as a definition of log-concavity of f. We will adopt the convention that f lower semi-continuous, i.e., the set $\{x \in \mathbb{R} \mid f(x) > t\}$ for some threshold t is open, and the set $\{x \in \mathbb{R} \mid f(x) \leq t\}$ is closed. Now, $\text{Supp}(f) = (-a, a)$ for some $a \in (0, \infty]$ where $\text{Supp}(f) := \{x \in \mathbb{R} \mid f(x) > 0\}$ denotes the support of f. ◁

We first prove a small technical result we rely on later.

Lemma 15. *Let f with support $(-a, a)$ be as in the setup Remark 8, $\infty > d \geq 0$, and let $f_+(x) := f(x - d)$. Then*

$$d \geq 2a \implies \text{Supp}(f) \cap \text{Supp}(f_+) = \emptyset$$
$$d < 2a \implies \left[\frac{d}{2}, a\right) \in \text{Supp}(f) \cap \text{Supp}(f_+)$$

both hold.

Proof of Lemma 15. Note that $\text{Supp}(f) = (-a, a)$ and $\text{Supp}(f_+) = (-a + d, a + d)$. When $d \geq 2a$, then $-a + d \geq a$, so $\text{Supp}(f) \cap \text{Supp}(f_+) = \emptyset$. On the other hand, when $d < 2a$, then $d/2 < a$, so $\{-d/2, d/2\} \subseteq \text{Supp}(f)$. From $f(-d/2) = f(d/2 - d) = f_+(d/2)$, we get that $d/2 \in \text{Supp}(f) \cap \text{Supp}(f_+)$. From the overlap of supports we can conlude $[d/2, a) \in \text{Supp}(f) \cap \text{Supp}(f_+)$. □

We now follow the proof outline and show that the set A^c can be expressed in terms of a threshold.

Lemma 16 (Thresholding). *Let f with support $(-a, a)$ be as in the setup Remark 8, $\infty > d \geq 0$, $\epsilon \geq 0$, and let $f_+(x) := f(x - d)$. Then for*

$$A^c := \{z \in \mathbb{R} \mid f_+(z) > e^{\epsilon} f(z)\}$$
$$A^c_- := A^c \cap \text{Supp}(f)$$
$$A^c_+ := \text{Supp}(f_+) - \text{Supp}(f)$$

each of the following hold.

1. $A^c_- \cap A^c_+ = \emptyset$
2. $A^c = A^c_- \cup A^c_+$
3. *if* $d \geq 2a$ *then* $A^c_- = \emptyset$ *and* $A^c_+ = \text{Supp}(f_+)$
4. *if* $d < 2a$ *then,*

$$A^c_- = \emptyset \implies A^c = \{z \in \text{Supp}(f_+) \mid z \geq a\},$$
$$A^c_- \neq \emptyset \implies A^c = \left\{z \in \left(\frac{d}{2}, a\right) \mid z > z^*\right\}$$

where

$$z^* = \max\left\{z \in \left(\frac{d}{2}, a\right) \mid f(z - d) = e^{\epsilon} f(z)\right\}.$$

Proof of Lemma 16. Note

a. Necessarily, $A^c_+ \cap \text{Supp}(f) = \emptyset$.

b. $z \in A^c \implies z \in \mathrm{Supp}(f_+)$ since we need $f_+(z) > 0$ for $f_+(z) > e^\epsilon f(z)$.
c. $z \in A_+^c \implies z \in A^c$ since here $f(z) = 0$ and $f_+(z) > 0$.

We conclude 1. from a. and $A_-^c \subseteq \mathrm{Supp}(f)$.

We conclude 2. from b. and c. together with 1.

When $d \geq 2a$, we have from Lemma 15 that $\mathrm{Supp}(f) \cap \mathrm{Supp}(f_+) = \emptyset$. From b. and c. we then conclude 3.

When $d < 2a$, the first implication of 4. follows from 2. Now assume $A_-^c \neq \emptyset$. We can then write $A_-^c = \left\{z \in \mathrm{Supp}(f) \mid \frac{f_+(z)}{f(z)} > e^\epsilon\right\}$. From Lemma 15 we have that $d/2 \in \mathrm{Supp}(f) \cap \mathrm{Supp}(f_+)$. Since $\frac{f_+(d/2)}{f(d/2)} = 1 \leq e^\epsilon$, we have that $d/2 \notin A_-^c$. By Lemma 11 we then get $x \in A_-^c \implies x > d/2$. Since $A_-^c \neq \emptyset$, there exists a $z \in (d/2, a)$ such that $f_+(z) > e^\epsilon f(z)$. Therefore, by lower semi-continuity of f we have $\sup A = \max A$ for $A = \{z \in (d/2, a) \mid f_+(z) = e^\epsilon f(z)\}$. From this, we conclude the second implication in 4. □

Lemma 17 (Privacy criterion)**.** *Let f with support $(-a, a)$ be as in the setup Remark 8, $\infty > d \geq 0$, $\epsilon \geq 0$, and let $f_+(x) := f(x - d)$. Furthermore, let $A^c = \{z \in \mathbb{R} \mid f_+(z) > e^\epsilon f(z)\}$, and let $Y_+ \sim f_+$ and $Y \sim f$ be two random variables. Then both the following hold.*

1. *If $d \geq 2a$, then*
$$P(Y_+ \in A^c) - e^\epsilon P(Y \in A^c) = P(Y_+ \in A^c) = 1.$$
2. *If $d < 2a$, then for $\delta \geq 0$*
$$P(Y_+ \in A^c) - e^\epsilon P(Y \in A^c) \leq \delta$$
$$\iff$$
$$P(Y_+ > t) - e^\epsilon P(Y > t) \leq \delta$$

   where
$$t := \sup\{z < a \mid f_+(z) \leq e^\epsilon f(z)\}.$$

Proof of Lemma 17. Let A_-^c and A_+^c be defined as in Lemma 16.

Point 1. follows directly from points 2. and 3. in Lemma 16.

From points 1. and 3. in Lemma 16 we have that

$$P(Y \in A^c) = P(Y \in A_-^c) + P(Y \in A_+^c) = P(Y \in A_-^c).$$

Also, note that $A_-^c = \emptyset \iff \left(\forall z \in (\tfrac{d}{2}, a)\right) f_+(z) \leq e^\epsilon f(z)$.

If $A_-^c = \emptyset$ then, by the above,

$$\begin{aligned} &P(Y_+ \in A^c) - e^\epsilon P(Y \in A^c) \leq \delta \\ &\iff \\ &P(Y_+ \in A^c) \leq \delta \\ &\iff \\ &P(Y_+ \geq a) \leq \delta \end{aligned}$$

where the last equivalence is due to Lemma 16 point 4. Since $P(Y_- \geq a) = 0$ always, we conclude that if $A_-^c = \emptyset$, then

$$P(Y_+ \in A^c) - e^\epsilon P(Y \in A^c) \leq \delta$$
$$\Longleftrightarrow$$
$$P(Y_+ \geq a) - e^\epsilon P(Y \geq a) \leq \delta.$$

If $A_-^c \neq \emptyset$, then by Lemma 16 point 4.

$$P(Y_+ \in A^c) = P(Y_+ > z^*)$$
$$P(Y_- \in A^c) = P(Y_- > z^*)$$

for $z^* = \max\left\{z \in \left(\frac{d}{2}, a\right) \mid f_+(z) \leq e^\epsilon f(z)\right\}$. We note that in both cases above, the thresholds a and z^*, respectively, can be expressed as $\sup\{z < a \mid f_+(z) \leq e^\epsilon f(z)\}$.

We conclude the proof by noting that $P(Y_+ \geq a) = P(Y_+ > a)$. □

Lemma 18 (Monotonicity of privacy criterion)**.** *Let f with support $(-a, a)$ be as in the setup Remark 8, $\infty > d > 0$, $\epsilon \geq 0$, $s > 0$. Also, let $F(x) = \int_{-\infty}^x f(y)dy$. Then, the function*

$$g(\epsilon, d, s) := F\left(\frac{d-t}{s}\right) - e^\epsilon F\left(-\frac{t}{s}\right)$$

for

$$t := \sup\left\{z < as \mid f\left(\frac{z-d}{s}\right) \leq e^\epsilon f\left(\frac{z}{s}\right)\right\}$$

is

a. increasing in d and strictly so if $t < as$
b. decreasing in s and strictly so if $t < as$
c. decreasing ϵ and strictly so if $t < as$
d. non-negative, and
e. zero if and only if $t = \infty$.

Proof of Lemma 18. We can partition the parameter region $W \subseteq [0, \infty)^2 \times (0, \infty)$ into two disjoint regions: $W_=$ when $t = as$ and $W_<$ when $t < as$.

First assume $(\epsilon, d, s) \in W_=$. If $a = \infty$, then $g = 0$, and a. − e. hold. Now, assume $a < \infty$. Then $g(\epsilon, d, s) = F\left(\frac{d-as}{s}\right)$ and does not depend on ϵ, i.e., g is increasing (but not strictly) in ϵ. Note $F(z)$ is increasing in z. Therefore, since $\frac{d-as}{s} = \frac{d}{s} - a$ is increasing in d and decreasing in s, we have that g is increasing in d and decreasing in s. Since F is non-negative, g is also non-negative, and also zero only if $d = 0$. Consequently, a. − e. all hold in this case.

Now assume $(\epsilon, d, s) \in W_<$. Let $\alpha = \frac{t-d}{s}$ and $\beta = \frac{t}{s}$. Then viewing $t = t(\epsilon, d, s) < as$ and applying the chain rule for differentiation and rearranging, we get

$$\begin{aligned}
\frac{\partial}{\partial d}g =& -\frac{1}{s}\left(f\left(\alpha\right)-e^{\epsilon}f\left(\beta\right)\right)\left(\frac{\partial}{\partial d}t\left(\epsilon,d,s\right)\right) \\
&+\frac{f\left(\alpha\right)}{2s}+\frac{e^{\epsilon}f\left(\beta\right)}{2s} \\
\frac{\partial}{\partial s}g =& \frac{1}{s^2}\left(f\left(\alpha\right)-e^{\epsilon}f\left(\beta\right)\right)t\left(\epsilon,d,s\right) \\
&-\frac{1}{s}\left(f\left(\alpha\right)-e^{\epsilon}f\left(\beta\right)\right)\left(\frac{\partial}{\partial s}t\left(\epsilon,d,s\right)\right) \\
&-\frac{f\left(\alpha\right)d}{2s^2}-\frac{e^{\epsilon}f\left(\beta\right)d}{2s^2}, \\
\frac{\partial}{\partial \epsilon}g =& \, e^{\epsilon}(F(\beta)-1)-\frac{1}{s}\left(\frac{\partial}{\partial \epsilon}t(\epsilon,d,s)\right)(f(\alpha)-e^{\epsilon}f(\beta))
\end{aligned}$$

Since f is lower semicontinous, we have that $f(\alpha)-e^{\epsilon}f(\beta)=0$ when $t<as$. Furthermore, since $t<as \implies t\in(-as,as)$, we get that $\beta=t/s\in(-a,a)$ and consequently $f(\beta)>0$ and $F(\beta)<1$. Then, since $d>0$, $s>0$, $f\geq 0$, and $F\leq 1$,

$$\begin{aligned}
\frac{\partial}{\partial d}g &= \frac{f\left(\alpha\right)}{2s}+\frac{e^{\epsilon}f\left(\beta\right)}{2s}>0 \\
\frac{\partial}{\partial s}g &= -\left(\frac{f\left(\alpha\right)d}{2s^2}+\frac{e^{\epsilon}f\left(\beta\right)d}{2s^2}\right)<0 \\
\frac{\partial}{\partial \epsilon}g &= e^{\epsilon}(F(\beta)-1)<0.
\end{aligned}$$

We have now proven parts a., b., c. and the "if" part of e. We now turn to part d.

We first note that we can write g as $F(-\alpha)-e^{\epsilon}F(-\beta)$. In order for $g\geq 0$, we must have that

$$\frac{F(-\alpha)}{F(-\beta)}\geq e^{\epsilon}.$$

Now, let $a(x)=(x-d)/s$, $b(x)=x/s$, and $\tau=t(\epsilon,d,s)$. Then we have that $\alpha=a(\tau)$ and $\beta=b(\tau)$. Using that $F(-x)=\int_x^{\infty}f(w)dw$ and integration by substitution we get

$$\begin{aligned}
F(-\alpha) &= F(-a(\tau))=s^{-1}\int_{\tau}^{\infty}f(a(x))dx \\
F(-\beta) &= F(-b(\tau))=s^{-1}\int_{\tau}^{\infty}f(b(x))dx.
\end{aligned}$$

Recalling that the likelihood ratio $r(x)=f(a(x))/f(b(x))$ is non-decreasing and that $r(\tau)=e^{\epsilon}$, we have that $f(a(x))\geq e^{\epsilon}f(b(x))$ for $x\geq\tau$. Consequently, we can conclude d. as

$$\frac{F(-\alpha)}{F(-\beta)}\geq e^{\epsilon}.$$

Now assume that $f(a(x))=e^{\epsilon}f(b(x))$ for all $x\geq\tau$. Then $t(\epsilon,d,s)=\sup\{z\mid r(z)\leq e^{\epsilon}\}=\infty$. Consequently, when $t(\epsilon,d,s)<\infty$ we must have that $\int_{\tau}^{\infty}f(a(x))>\int_{\tau}^{\infty}e^{\epsilon}f(b(x))$. This in turn means that $\frac{F(-\alpha)}{F(-\beta)}>e^{\epsilon}$ which implies $g>0$ when $t(\epsilon,d,s)<\infty$. This concludes the proof of the "only if" part of e., and the proof of the lemma. □

Proof of Lemma 1. Let $x = q(d)$ and $y = q(d')$, $d = x - y$, $Y_+ = sX + d$, $Y_- = sX$, $f_+(x) = s^{-1}f(s^{-1}(x-d))$, and $f_-(x) = s^{-1}f(s^{-1}x))$. Then, variables Y_+ and Y_- are distributed according to densities f_+ and f_-, respectively.

By Lemma 12, in order to show $\Pr(sX + x \in S) \leq e^\epsilon \Pr(sX + y \in S) + \delta$ for all measurable sets S, it suffices to show that $\Pr(Y_+ \in S) \leq e^\epsilon \Pr(Y_- \in S) + \delta$ for all measurable sets S. The latter is equivalent to showing that for all measurable S, $\Pr(Y_+ \in S) - e^\epsilon \Pr(Y_- \in S) \leq \delta$. Since f is even, we can apply Lemma 13 to let $d \geq 0$ without loss of generality for the remainder of this proof.

Let $A^c = \{x \mid f_+(x) > e^\epsilon f_-(x)\}$ Then, from Lemma 14 we have that

$$\Pr(Y_+ \in A^c) - e^\epsilon \Pr(Y_- \in A^c) \leq \delta \tag{12}$$

is a sufficient and necessary condition for $\Pr(sX + x \in S) \leq e^\epsilon \Pr(sX + y \in S) + \delta$ for all measurable sets S.

Since f is even, we have that $F(-x) = 1 - F(x)$. This allows us to write $\Pr(Y_+ > t) = F((d-t)/s)$ and $\Pr(Y_- > t) = F(-t/s)$. Furthermore, $f((z-d)/s) \leq e^\epsilon f(z) \iff \psi(z/s) - \psi((z-d)/s) \leq \epsilon$. This, together with the chaining of Lemma 17 and Lemma 18 conclude the proof. □

Proof of Lemma 2. From Inspecting (1), it suffices to show that for $\Delta t(\epsilon, 1, s) = t(\epsilon, \Delta, s\Delta)$ for $\Delta > 0$. Let $t_z := \sup\left\{z \mid \frac{f\left(\frac{z-1}{s}\right)}{f\left(\frac{z}{s}\right)} \leq e^\epsilon\right\}$. Substituting $u = z\Delta$ into t_z, we obtain $t_u = \sup\left\{u \mid \frac{f\left(\frac{u-\Delta}{s\Delta}\right)}{f\left(\frac{u}{s\Delta}\right)} \leq e^\epsilon\right\}$. From this we see that $\Delta t_z = t_u$. That $s(\epsilon, \delta, \Delta)$ is non-increasing in both ϵ and δ is due to Lemma 18. □

A.3 Proofs from Section 3.2

Proof of Theorem 3. For the standard Laplace distribution, $\psi(x) = \psi_1 = \log(2) + |x|$, which is convex and even. Then

$$\psi\left(\frac{x}{s}\right) - \psi\left(\frac{x - \Delta}{s}\right) = \left|\frac{x}{s}\right| - \left|\frac{\Delta - x}{s}\right| = \epsilon$$

has solutions

$$S = \begin{cases} \{x \mid x \geq \Delta\}, & \Delta = \epsilon s \\ \left\{\frac{\epsilon s}{2} + \frac{\Delta}{2}\right\}, & \Delta > \epsilon s \\ \emptyset, & \text{otherwise.} \end{cases}$$

When $\Delta = \epsilon s$, we have that $t = \infty$, and applying Lemma 1, we have $(\epsilon, 0)$-differential privacy when $s \geq \Delta/\epsilon$. Now, let $\Delta > \epsilon s$. Letting $F(x) = F_1(x)$ (the standard Laplace cdf) and substituting $\frac{\epsilon s}{2} + \frac{\Delta}{2}$ for t in (1) and solving for s yields

$$s \geq \frac{\Delta}{\epsilon - 2\log(1 - \delta)}.$$

For $\delta = 0$, the above coincides with $s \geq \Delta/\epsilon$. The theorem follows from Lemma 1. □

Proof of Theorem 4. We have that $f_{\log}(x) = e^{-\psi(x)}$ for $\psi(x) = \psi_{\log} = x + 2\log\left(1+e^{-x}\right)$. Since $\psi''(x) = \frac{2e^{-x}}{(1+e^{-x})^2} \geq 0$, we have that ψ is convex. Evenness can be seen from $e^{\psi(x)} = 2 + e^{-x} + e^{x}$ and the fact that e^x is a strictly monotone function. Then, the equation

$$\begin{aligned}\psi\left(\frac{x}{s}\right) - \psi\left(\frac{x-\Delta}{s}\right) \\ &= \frac{2\log\left(1+e^{-\frac{x}{s}}\right)s - 2\log\left(1+e^{\frac{-x+\Delta}{s}}\right)s + \Delta}{s} \\ &= \epsilon\end{aligned}$$

solved for x has solution

$$x^* = -\log\left(A-1\right)s, \qquad \text{where}$$
$$A = e^{-\frac{-2s\log\left(e^{\frac{\Delta}{s}}-1\right)+2s\log\left(-1+e^{\frac{\epsilon s+\Delta}{2s}}\right)-\epsilon s+\Delta}{2s}}.$$

Let $t = x^*$, and let $F(x) = \int_{-\infty}^{x} f_{\log}(w)dw$. Then, manipulation yields

$$\begin{aligned}F\left(\frac{\Delta-t}{s}\right) - e^{\epsilon}F\left(-\frac{t}{s}\right) \\ &= \frac{e^{\frac{\epsilon s+\Delta}{s}} + e^{\frac{2\Delta}{s}} + 2e^{\frac{\epsilon s+\Delta}{2s}} - 2e^{\frac{\epsilon s+3\Delta}{2s}} - e^{\frac{\Delta}{s}} - e^{\epsilon}}{\left(e^{\frac{\Delta}{s}}-1\right)^2}.\end{aligned}$$

Solving

$$\frac{e^{\frac{\epsilon s+\Delta}{s}} + e^{\frac{2\Delta}{s}} + 2e^{\frac{\epsilon s+\Delta}{2s}} - 2e^{\frac{\epsilon s+3\Delta}{2s}} - e^{\frac{\Delta}{s}} - e^{\epsilon}}{\left(e^{\frac{\Delta}{s}}-1\right)^2} = \delta$$

for s yields two solutions

$$\frac{\Delta}{2\log\left(\frac{e^{\frac{\epsilon}{2}} \pm \sqrt{\delta(e^{\epsilon}+\delta-1)}}{1-\delta}\right)}$$

of which the larger drops out due to monotonicity of the left hand side of (1). The smaller produces the stated bound. The theorem then follows from Lemma 1. ☐

Proof of Theorem 5. The standard Gaussian density is $f_2(x) = e^{-\psi(x)}$ for convex and even $\psi(x) = \psi_2(x) = \frac{\log(2\pi)}{2} + \frac{x^2}{2}$. Since $\psi\left(\frac{z}{\sigma}\right) - \psi\left(\frac{z-d}{\sigma}\right) = \left(\frac{\Delta}{\sigma^2}\right)z - \frac{\Delta^2}{2\sigma^2}$, we see that ψ is MLR-unbounded. The equation $\psi\left(\frac{z}{\sigma}\right) - \psi\left(\frac{z-d}{\sigma}\right) = \epsilon$ has unique solution $z^* = \frac{2\epsilon\sigma^2+\Delta^2}{2\Delta}$ for $\sigma > 0$ and $\Delta > 0$. The theorem then follows from Lemma 1. ☐

A.4 Proofs from Section 3.3

Lemma 19. *Let $\psi : \mathbb{R} \to \mathbb{R}$ be even, convex, and non-constant. For $d > 0$, $l(z, s) = \psi(z/s) - \psi((z-d)/s)$ is increasing in z and strictly decreasing in $s > 0$ for $z > d/2$.*

Proof of Lemma 19. For $d > 0$, l is increasing in z by Lemma 11. For l to be decreasing in s, we need that if $s < s'$ then $l(z, s) - l(z, s') \geq 0$. We have

$$l(z, s) - l(z, s') = a - b$$

for $a = \psi\left(\frac{z}{s}\right) - \psi\left(\frac{z-d}{s}\right)$ and $b = \psi\left(\frac{z}{s'}\right) - \psi\left(\frac{z-d}{s'}\right)$. Now assume that $z > d/2$, which by non-constancy, convexity and eveness of ψ ensures that $\psi(z/s) > \psi((z-d)/s)$ for any $s > 0$, so $a > 0$ and $b > 0$. We now show that $a > b$ to round out the proof. Let $x_{s1} = (z-d)/s$, $x_{s2} = z/s$, $x_{s'1} = (z-d)/s'$, and $x_{s'2} = z/s'$. Then, $x_{s2} - x_{s1} = d/s > d/s' = x_{s'2} - x_{s'1}$, $x_{s'1} < x_{s1}$, and $x_{s'2} < x_{s2}$. Consequently (by $z > d/2$, evenness ,convexity, and non-constancy) we have that the slope of the line passing through $(x_{s'1}, \psi(x_{s'1}))$ and $(x_{s'2}, \psi(x_{s'2}))$ is positive and at most that of the line passing through $(x_{s1}, \psi(x_{s1}))$ and $(x_{s2}, \psi(x_{s2}))$. Then since $x_{s2} - x_{s1} = d/s > d/s' = x_{s'2} - x_{s'1}$, we must have that $a = \psi(x_{s2}) - \psi(x_{s1}) > b = \psi(x_{s'2}) - \psi(x_{s'1})$. □

Proof of Theorem 6. Let $l(z, s) = \psi(z/s) - \psi((z-d)/s)$. Since ψ is even and convex, l is increasing in z from Lemma 11. Since f is MLR-unbounded, there will always exist z such that $\psi(z/s) - \psi((z-d)/s) > \epsilon$ for any $\epsilon > 0$. This, in turn means that $t \in \mathrm{Supp}(f)$. From Lemma 18 we then have that $F((\Delta - t)/s) - e^\epsilon F(-t/s)$ is strictly decreasing in s, non-negative, and 0 if and only if $t = \sup\{z \mid l(z, s) \leq \epsilon\} = \infty$. Noting that we only need to consider $z > d/2$ (From the proof of Lemma 17), applying Lemma 19, we can make t arbitrarily large by increasing s. Consequently, $\lim_{s\to\infty} F((\Delta - t)/s) - e^\epsilon F(-t/s) = 0$, and we conclude the Lemma. □

A.5 Proofs from Section 3.4.1

Proof of Theorem 7. Applying Theorem 3 and Theorem 4, we see that

$$\rho(\epsilon, \delta) = \frac{2\log\left(\sqrt{\delta}\sqrt{e^\epsilon + \delta - 1} + e^{\frac{\epsilon}{2}}\right) - 2\log(1-\delta)}{\epsilon - 2\log(1-\delta)}. \tag{13}$$

For $1 > \delta \geq 0$ we get that

$$2\log(2) + \epsilon > 2\log\left(\sqrt{\delta}\sqrt{e^\epsilon + \delta - 1} + e^{\frac{\epsilon}{2}}\right) \geq \epsilon.$$

Substituting, the lower bound into (13) we get

$$\rho(\epsilon, \delta) \geq 1.$$

Renaming ϵ as x for readability, we now write

$$\begin{aligned}
\rho(x, \delta) &= \frac{f(x)}{g(x)} \text{ where} \\
f(x) &= -2\log\left(\sqrt{\delta}\sqrt{e^x + \delta - 1} + e^{\frac{x}{2}}\right) \\
&\quad + 2\log(1-\delta) \leq 0 \\
g(x) &= -x + 2\log(1-\delta) \leq 0, \text{with} \\
f'(x) &= -\frac{2\left(\frac{\sqrt{\delta}e^x}{2\sqrt{e^x+\delta-1}} + \frac{e^{\frac{x}{2}}}{2}\right)}{\sqrt{\delta}\sqrt{e^x + \delta - 1} + e^{\frac{x}{2}}} \leq 0 \\
g'(x) &= -1 \leq 0,
\end{aligned}$$

with all the inequalities strict for $x > 0$. For

$$H(x) = \frac{g(x)^2 (f/g)'(x)}{|g'(x)|}$$

we now get

$$\lim_{x\to 0} H(x) = h(\delta) \quad \text{for}$$
$$h(\delta) = \frac{(2\delta - 2)\log(1-\delta) - 2\log(\delta+1)(\delta+1)}{\delta+1},$$

and

$$h'(\delta) = \frac{4\log(1-\delta)}{(\delta+1)^2} < 0,$$

which means that for $1 > \delta \geq 0$, $h(\delta) \leq 0$.

We have that

$$(f'/g')'(x) =$$
$$q(x)\frac{\sqrt{\delta}e^{\frac{x}{2}}}{2\left(e^x + \delta - 1\right)^{\frac{3}{2}}\left(\sqrt{\delta}\sqrt{e^x+\delta-1} + e^{\frac{x}{2}}\right)^2},$$

for

$$q(x) = \left(-2e^{\frac{x}{2}}\left(\sqrt{\delta} - \delta^{\frac{3}{2}}\right)\sqrt{e^x+\delta-1}\right) + (\delta-1)(\delta+2e^x-1),$$

and recognize that $(f'/g')'(x) \leq 0$ if $q(x) \leq 0$. Since $\sqrt{\delta} - \delta^{\frac{3}{2}} \geq 0$ for $1 > \delta \geq 0$, we also recognize that $q(x) \geq 0$.

A combination of Theorems 4 and 5 in Anderson et al.(Anderson et al., 2006), yields: If g' never vanishes on an open interval $(0, b) \in \mathbb{R}$ and $gg' > 0$ on $(0, b)$, $\lim_{x\to 0} H(x) \leq 0$, and f'/g' is decreasing on $(0, b)$, then $(f/g)' < 0$ on $(0, b)$. By the above, all the conditions are met and we can conclude that $\rho(\epsilon, \delta)$ is decreasing in ϵ for $\epsilon \in (0, b)$ for arbitrary $b > 0$.

Now,

$$\lim_{\epsilon\to 0} \rho(\epsilon, \delta) = \rho(0, \delta) = \frac{-\log(\delta+1) + \log(1-\delta)}{\log(1-\delta)}$$

and

$$\frac{d}{d\delta}\rho(0, \delta) = \frac{\log\left(\frac{1-\delta}{\delta+1}\right)\delta + \log\left(\frac{1}{(\delta+1)(1-\delta)}\right)}{\log(1-\delta)^2(1-\delta)(\delta+1)},$$

which we recognize as being non-positive, meaning that the ratio decreases in δ. Consequently,

$$\rho(\epsilon, \delta) < \lim_{\delta\to 0} \rho(0, \delta) = 2,$$

which concludes the proof. □

Proof of Corollary 1. Follows from $\mathrm{Var}(X_1) = 2$, $\mathrm{Var}(X_{\log}) = \frac{\pi^2}{3}$,

$$\frac{\mathrm{Var}(s_1(\epsilon, \delta, \Delta)X_1)}{\mathrm{Var}(s_{\log}(\epsilon, \delta, \Delta)X_{\log})} = (\rho_{1,\log}(\epsilon, \delta))^2 \frac{\mathrm{Var}(X_1)}{\mathrm{Var}(X_{\log})},$$

and Theorem 7. □

A.6 Proofs from Section 4

Proof of Lemma 8. Let $d > 0$, and let

$$\begin{aligned}
D_d^{(n)} &= \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| = d\}, \\
A_{\mathbf{d}}^{(n)} &= \{\mathbf{x} \in \mathbb{R}^n \mid f(\|\mathbf{x} - \mathbf{d}\|) \leq e^{\epsilon} f(\|\mathbf{x}\|)\}, \\
A_d^{(n)} &= \bigcup_{\mathbf{d} \in D_d^{(n)}} A_{\mathbf{d}}^{(n)}.
\end{aligned}$$

We note that for $\mathbf{d} \in D_d^1$ we have

$$\begin{aligned}
A_{\mathbf{d}}^{(1)} &= \{x \in \mathbb{R} \mid f(|x - d|) \leq e^{\epsilon} f(|x|)\} \\
&= \{x \in \mathbb{R} \mid f(x - d) \leq e^{\epsilon} f(x)\}.
\end{aligned}$$

The first equality in the above display is due to isomorphy and the second is due to f being even. Therefore,

$$A_d^{(1)} = \{x \in \mathbb{R} \mid f(x - d) \leq e^{\epsilon} f(x)\}.$$

If we are able to show that $\mathbf{x} \in A_d^{(n)} \iff \|\mathbf{x}\| \in A_d^{(1)}$, then the proof follows by the argument in the proof of Lemma 1 for $A^c = \left(A_d^{(1)}\right)^c$.

We start with ($\implies$). First note $\mathbf{x} \in A_d^{(n)} \implies \forall \mathbf{d} \in D_d^{(n)}(f(\|\mathbf{x} - \mathbf{d}\|) \leq e^{\epsilon} f(\|\mathbf{x}\|))$, and in particular for $\mathbf{y} = \arg\min_{\mathbf{d} \in D_d^{(n)}} \|\mathbf{x} - \mathbf{d}\| = c\mathbf{x}$ for $c = \frac{d}{\|\mathbf{x}\|}$. We also have that

$$f(\|\mathbf{x} - \mathbf{y}\|) \geq \max_{\mathbf{d} \in D_d^{(n)}} f(\|\mathbf{x} - \mathbf{d}\|), \tag{14}$$

since f is non-increasing on $\mathbb{R}_{\geq 0}$. Now,

$$\begin{aligned}
\|\mathbf{x} - \mathbf{y}\| &= \|\mathbf{x} - c\mathbf{x}\| = \|(1 - c)\mathbf{x}\| \\
&= |1 - c|\|\mathbf{x}\| \\
&= \begin{cases} \|\mathbf{x}\| - d & \text{if } d \leq \|\mathbf{x}\|, \\ d - \|\mathbf{x}\| & \text{if } d > \|\mathbf{x}\|. \end{cases}
\end{aligned}$$

Since f is even, $f(\|\mathbf{x}\| - d) = f(d - \|\mathbf{x}\|)$, and consequently

$$f(\|\mathbf{x} - \mathbf{y}\|) \leq e^{\epsilon} f(\|\mathbf{x}\|) \implies f(\|\mathbf{x}\| - d) \leq e^{\epsilon} f(\|\mathbf{x}\|).$$

By (14) we conclude the ($\implies$) case.

We now turn to ($\impliedby$). By the reverse triangle inequality we have $\|\mathbf{x} - \mathbf{d}\| \geq \|\mathbf{x}\| - \|\mathbf{d}\|$ and since f is non-increasing on $\mathbb{R}_{\geq 0}$, $f(\|\mathbf{x} - \mathbf{d}\|) \leq f(\|\mathbf{x}\| - \|\mathbf{d}\|)$. From this we conclude the ($\impliedby$) case. □

Proof of Theorem 9. Applying a reparameterization of $\tau = (2/q)^{1/q}$ in (12) in Osiewalski and Steel (1993), they show that when the independent $X_i \sim f_p$, where f_p is the Subbotin$_p$ density, the variable $\mathbf{X}$ has a $\|\cdot\|_p$-spherical distribution. Since the distribution of $\mathbf{X}$ is $\|\cdot\|_p$-spherical, $\mathbf{X}$ is distributed according to density $g(\mathbf{x}) = h(\|\mathbf{x}\|_p)$ for a function h. Since the marginals $X_i \sim f_p$, we must have that $g(\mathbf{x}) = C^{-1} f_p(\|\mathbf{x}\|_p)$ for a normalizing constant C. The theorem then follows from Lemma 8. □

A.7 Proofs from Section 4.1

Proof of Theorem 10. Suppose $\|f(R)\|_\infty = 0$, in which case it is an l_∞-ball and $\Delta_p = 0$. Now, suppose $\|f(R)\|_\infty > 0$ and let $(v, w) \in \arg\sup_{(x,y)\in f(R)^2} \|x - y\|_\infty$. Since $\|f(R)\|_\infty > 0$, there exists distinct $a, b \in R$ such that $f(a) = v$ and $f(b) = w$. We can now let $d_1 \in \{a\}^n$ and $d_2 \in \{a\}^{n-1} \times \{b\}$. Then, $\|q(d_1) - q(d_2)\|_p = |\nu(n)| \left(\sum_{i=1}^m |v_i - w_i|^p\right)^{1/p} \leq m^{1/p}|\nu(n)|\|v - w\|_\infty = m^{1/p} |\nu(n)| \|f(R)\|_\infty$. By linearity of f, we have that $\|q(d_1) - q(d_2)\|_p \geq \|q(d) - q(d')\|_p$ for any neighboring $d, d' \in R^n$. If $f(R)$ is a l_∞-ball, we can choose v and w such that $|v_i - w_i| = \|v - w\|_\infty$ for all i. Then we get the claimed equality. □