On the Geometry of Differential Privacy

Moritz Hardt Kunal Talwar

Abstract

We consider the noise complexity of differentially private mechanisms in the setting where the user asks d linear queries $f\colon \mathbb{R}^N \to \mathbb{R}$ non-interactively. Here, the database is represented by a vector in $\mathbb{R}^N$ and proximity between databases is measured in the ℓ_1-metric.

We show that the noise complexity is determined by two geometric parameters associated with the set of queries. We use this connection to give tight upper and lower bounds on the noise complexity for any $d \leqslant N$. We show that for d random linear queries of sensitivity 1, it is necessary and sufficient to add ℓ_2-error $\Theta(\min\{d\sqrt{d}/\varepsilon, d\sqrt{\log(N/d)}/\varepsilon\})$ to achieve ε-differential privacy. Assuming the truth of a deep conjecture from convex geometry, known as the Hyperplane conjecture, we can extend our results to arbitrary linear queries giving nearly matching upper and lower bounds.

Our bound translates to error $O(\min\{d/\varepsilon, \sqrt{d\log(N/d)}/\varepsilon\})$ per answer. The best previous upper bound (Laplacian mechanism) gives a bound of $O(\min\{d/\varepsilon, \sqrt{N}/\varepsilon\})$ per answer, while the best known lower bound was $\Omega(\sqrt{d}/\varepsilon)$. In contrast, our lower bound is strong enough to separate ε-differential privacy from (ε, δ)-differential privacy even when δ is a negligible function of d.

1 Introduction

The problem of privacy-preserving data analysis has attracted a lot of attention in recent years. Several databases, e.g., those held by the U.S. Census Bureau or the National Institute of Health, contain private data provided by individuals; protecting the privacy of those individuals is an important concern. Differential privacy is a formal notion due to Dwork et al. [Dwo06, DMNS06] that attaches a rigorous meaning to the word *privacy*. Intuitively speaking, differential privacy gives the strong guarantee that the presence or absence of any single individual in a data set will only insignificantly affect the outcome of an analysis. More precisely, differentially private algorithms are randomized algorithms whose output distribution remains nearly unchanged even if we perturb the information of a single participant arbitrarily. Differential privacy has the appealing feature that it provides a privacy guarantee while making very mild assumptions on the background knowledge of the adversary. Differential privacy is thus highly resilient to attacks utilizing unanticipated auxiliary information. There are several other useful properties that make differential privacy a robust definition. For instance, differential privacy *composes* gracefully in the sense that the interaction of two differentially private mechanisms remains differentially private up to a small quantitative loss in the privacy guarantee (not expressed in the informal description above). We refer the reader to surveys [Dwo09, Dwo11] for additional motivation of the definition.

As differential privacy poses a strong requirement, it becomes a challenging task to design useful algorithms that satisfy differential privacy. The most basic and well-studied setting of differential privacy is the case where a trusted database curator responds to a number of queries given by an (untrusted) data analyst. We consider the following general setting: A *database* is represented by a vector $x \in \mathbb{R}^N$. The *queries* that the analyst may ask are *linear* combinations of the entries of x. More precisely, a multidimensional query is a map $F\colon \mathbb{R}^N \to \mathbb{R}^d$, and we will restrict ourselves to linear maps F with coefficients in the interval $[-1,1]$. Thus F is a $d \times N$ matrix with entries in $[-1,1]$. In this work, we assume throughout that $d \leqslant N/2$. This is without loss of generality as we may always add zero coordinates to x. A mechanism is a randomized algorithm which holds a database $x \in \mathbb{R}^N$, receives a query $F\colon \mathbb{R}^N \to \mathbb{R}^d$ and answers with some $a \in \mathbb{R}^d$. Informally, we say a mechanism satisfies differential privacy in this setting if the densities of the output distributions on inputs $x, x' \in \mathbb{R}^N$ with $\|x - x'\|_1 \leqslant 1$ are point wise within an $\exp(\varepsilon)$ multiplicative factor of each other. Here and in the following, $\varepsilon > 0$ is a parameter that measures the strength of the privacy guarantee (smaller ε being a stronger guarantee). The *error* of a mechanism is the expected Euclidean distance between the correct answer Fx and the actual answer a.

In this work, we use methods from convex geometry to determine a nearly optimal trade-off between privacy and error. We will see a lower bound on how much error any differentially private mechanism must add. And we present a mechanism whose error nearly matches this lower bound.

As mentioned, the above setup is fairly general. To illustrate it and facilitate comparison with previous work, we will describe some specific instantiations below.

Histogram view In the most common setting a database is a set $D \subseteq U$ where U is a universe of data items. We think of each data item $u \in D$ as belonging to one individual. This setting maps into our framework by putting $N = |U|$ and representing D as a histogram

$x \in \mathbb{R}^N$ in the natural way; that is, x_u counts the number of occurences of the data item u in the data base. Note that adding or removing a single individual's data from the data set translates to a unit change in the ℓ_1-norm of the corresponding histogram x.

Note that in the definition of differential privacy, we require the mechanism to be defined for all $x \in \mathbb{R}^N$ and demand that the output distributions be close whenever $\|x - x'\|_1 \leqslant 1$. This is a stronger requirement than asserting this property only for integer vectors x and x'. It only makes our upper bounds stronger. For the lower bounds, this strengthening allows us to ignore the discretization issues that would arise in the usual definition. Building on the techniques presented in this work, De [De11] showed that our lower bounds can be extended to hold for the usual definition.

As explained above, our upper bound holds for any linear query on the histogram. This includes some well-studied and natural classes of queries. For instance, *counting queries* as studied by [BLR08, DNR+09, DRV10, RR10, HR10], or *contingency tables* (see, e.g., [BCD+07, KRSU10, GHRU11]) are linear queries on the histogram.

Other settings In the setting looked at by Dinur and Nissim [DN03], the database $y \in \{0,1\}^n$ consists of one private bit for each individual and each query ask for the number of 1's amongst a (random) subset on $[n]$. Given d such queries, one can define $n \leqslant 2^d$ types of individuals, depending on the subset of the queries that ask about an individual. The vector y then maps to a histogram x in the natural way with x_i denoting the number of individuals of type i with their private bit set to 1. Our results then imply a lower bound of $\Omega(d/\varepsilon)$ per answer for any ε-differentially private mechanism. This improves on the $\Omega(\sqrt{d})$ bound for $d = n$ from [DN03] for a weaker privacy definition (blatant non-privacy). A closely related rephrasing is to imagine each individual having d private $\{0,1\}$ attributes so that $n = 2^d$. The d queries that ask for the 1-way marginals of the input naturally map to a matrix F and Theorem 1.1 implies a lower bound of $\Omega(d/\varepsilon)$ noise per marginal for such queries.

One can also look at x itself as a database where each individuals private data is in $[0,1]$; in this setting the dimension of the data n equals the number of individuals n. Our results lead to better upper bounds for this setting.

Finally, there are settings such as the work of [MM09] on private recommendation systems, where the private data is transformed with a stability guarantee so that nearby databases get mapped to vectors at ℓ_1 distance at most 1.

1.1 Our results

We relate the noise complexity of differentially private mechanisms to some geometric properties of the image of the unit ℓ_1-ball, denoted B_1^N, when applying the linear mapping F. We will denote the resulting convex polytope by $K = FB_1^N$. Our first result lower bounds the noise any ε-differentially private mechanism must add in terms of the *volume radius* of K, denoted $\mathrm{vr}(K)$. Here,

$$\mathrm{vr}(K) \overset{\mathrm{def}}{=} \left(\frac{\mathrm{Vol}(K)}{\mathrm{Vol}(B_2^d)}\right)^{1/d},$$

where B_2^d denotes the d-dimensional Euclidean ball.[1]

[1] Volume radius is typically defined as $\mathrm{Vol}(K)^{1/d}$. The different normalization we chose will be convenient for us.

Theorem 1.1. *Let $\varepsilon > 0$ and suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map. Then, every ε-private mechanism M has error at least*

$$\Omega\left(\frac{d \cdot \mathrm{vr}(K)}{\varepsilon}\right). \tag{1}$$

Recall, the term *error* refers to the expected Euclidean distance between the output of the mechanism and the correct answer to the query F.

We then describe a differentially private mechanism whose error depends on the expected ℓ_2-norm of a randomly chosen point in K. Our mechanism is an instantiation of the exponential mechanism [MT07] with the score function defined by the (negative of the) norm $\|\cdot\|_K$, that is the norm which has K as its unit ball. Hence, we will refer to this mechanism as the K-norm mechanism. Note that as the definition of this norm depends on the query F, so does the output of our mechanism. The error of this mechanism can be described in terms of the *mean radius* of K defined as

$$\mathrm{mr}(K) \overset{\mathrm{def}}{=} \mathop{\mathbb{E}}_{z \sim K} \|z\|_2$$

where z is drawn uniformly at random from K.

Theorem 1.2. *Let $\varepsilon > 0$ and suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map with $K = FB_1^N$. Then, the K-norm mechanism is ε-differentially private and has error at most*

$$O\left(\frac{d \cdot \mathrm{mr}(K)}{\varepsilon}\right). \tag{2}$$

As it turns out, when F is a random Bernoulli ± 1 matrix our upper bound matches the lower bound up to constant factors. In this case, K is a random polytope and its volume and mean radius have been determined rather recently. Specifically, we apply a volume lower bound of Litvak et al. [LPRN05], and an upper bound on the mean radius due to Klartag and Kozma [KK09]. Quantitatively, we obtain the following theorem.

Theorem 1.3. *Let $\varepsilon > 0$ and $d \leqslant N/2$. Then, for almost all matrices $F \in \{-1,1\}^{d \times N}$,*

1. *any ε-differentially private mechanism M has error $\Omega(d/\varepsilon) \cdot \min\left\{\sqrt{d}, \sqrt{\log(N/d)}\right\}$.*
2. *the K-norm mechanism is ε-differentially private with error $O(d/\varepsilon) \cdot \min\left\{\sqrt{d}, \sqrt{\log(N/d)}\right\}$.*

We remark that Litvak et al. [LPRN05] also give an explicit construction of a mapping F realizing the lower bound.

More generally, we can relate our upper and lower bounds whenever the body K is in *approximately isotropic position*. Informally, this condition implies that $\mathrm{mr}(K) \sim \mathrm{vr}(K) L_K$ where L_K denotes the so-called *isotropic constant* which is defined in Section 6.

Theorem 1.4. *Let $\varepsilon > 0$ and suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map such that $K = FB_1^N$ is in approximately isotropic position. Then, the K-norm mechanism is ε-differentially private with error at most $O(\varepsilon^{-1}\mathrm{vr}(dK))$.*

Notice that the bound in the previous theorem differs from the lower bound by a factor of L_K. A central conjecture in convex geometry, sometimes referred to as the "Hyperplane

Mechanism	**ℓ_2-error**	**privacy**	reference
Laplacian noise	$\varepsilon^{-1}d\sqrt{d}$	ε	[DMNS06]
K-norm	$\varepsilon^{-1}d\sqrt{\log(N/d)}$	ε	this paper
lower bound	$\Omega(\varepsilon^{-1}d)$	(ε,δ)	[DN03]
lower bound	$\Omega(\varepsilon^{-1}d)\min\left\{\sqrt{\log(N/d)},\sqrt{d}\right\}$	ε	this paper

Figure 1: Summary of results in comparison to best previous work for d *random* linear queries each of sensitivity 1 where $1 \leqslant d \leqslant n$. Note that informally the average per coordinate error is smaller than the stated bounds by a factor of $\sqrt{d}$. Here, (ε,δ)-differential privacy refers to a weaker approximate notion of pricacy introduced later. Our lower bound does not apply to this notion.

Conjecture" or "Slicing Conjecture" states that $L_K = O(1)$. See, e.g., [MP89, Gia03, KK09] for further information on the subject.

Unfortunately, in general the polytope K could be very far from isotropic. In this case, both our volume-based lower bound and the K-norm mechanism can be quite far from optimal. We give a recursive variant of our mechanism and a natural generalization of our volume-based lower bound which are nearly optimal even if K is non-isotropic.

Theorem 1.5. *Let $\varepsilon > 0$. Suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map. Further, assume the Hyperplane Conjecture. Then, the mechanism introduced in Section 7 is ε-differentially private and has error at most $O(\log^{3/2} d)\cdot \mathrm{VolLB}(K,\varepsilon)$. where $\mathrm{VolLB}(K,\varepsilon)$ is a lower bound on the error of the optimal ε-differentially private mechanism.*

While we restricted our theorems to $F \in [-1,1]^{d\times N}$, they apply more generally to any linear mapping F.

Efficient Mechanisms Our mechanism is an instantiation of the exponential mechanism and involves sampling random points from rather general high-dimensional convex bodies. This is why our mechanism is not efficient as it is. However, we can use rapidly mixing geometric random walks for the sampling step. These random walks turn out to approach the uniform distribution in a metric that is strong enough for our purposes. It will follow that both of our mechanisms can be implemented in polynomial time.

Theorem 1.6. *The mechanisms given in Theorem 1.2 and Theorem 1.5 can be implemented in time polynomial in $n, 1/\varepsilon$ such that the stated error bound remains the same up to constant factors, and the mechanism achieves ε-differential privacy.*

We note that our lower bound VolLB can also be approximated up to a constant factor. Together these results give polynomial time computable upper and lower bounds on the error of any differentially private mechanism, that are always within an $O(\log^{3/2} d)$ of each other.

Figure 1 summarizes our results. Note that we state our bounds in terms of the total ℓ_2 error, which informally is a $\sqrt{d}$ factor larger than the average per coordinate error.

1.2 Related Work

In this section we describe related previous as well as subsequent work.

Laplacian mechanism Queries of the kind described above have (total) *sensitivity* d, and hence the work of Dwork et al. [DMNS06] shows that adding Laplace noise with parameter d/ε to each entry of Fx ensures ε-differential privacy. Moreover, adding Laplace noise to the histogram x itself leads to another private mechanism. Thus such questions can be answered with noise $\min(d/\varepsilon, \sqrt{n}/\varepsilon, N)$ per entry of Fx. Some specific classes of queries can be answered with smaller error. Nissim, Raskhodnikova and Smith [NRS07] show that one can add noise proportional to a smoothed version of the *local sensitivity* of the query, which can be much smaller than the global sensitivity for some *non-linear* queries.

Mechanisms for large numbers of counting queries Blum, Ligett and Roth [BLR08] consider the problem of releasing a large number of counting queries on a database of size n over a universe of size N. As mentioned earlier counting queries are a special case of linear queries in the histogram space. They show that it is possible to release answers to d counting queries with error $O(n^{2/3}\log^{1/3}(d)\log^{1/3}(N)/\varepsilon^{1/3})$ per entry. This bound is in general incomparable to ours due to the dependence on the number of individuals n. However, when $n \ll d$, their mechanism has smaller error than ours. A drawback is that their mechanism has a running time of roughly N^n. Follow-up works [DNR$^+$09, DRV10] improved the running time to polynomial in N thus matching the running time of our mechanism. Subsequent to our work, these results were further extended to the interactive[2] setting [RR10, HR10]. The latter result achieves an error bound of $O(\varepsilon^{-1/2}\sqrt{n}\log(d)\log^{1/4}(N))$ with a running time polynomial in N. Following our work, [HR10] exploits the idea of viewing a data set as a histogram $x \in \mathbb{R}^N$. Their algorithm (based on a multiplicative weights update rule) however does not have a natural geometric interpretation.

Feldman et al. [FFKN09] construct private core sets for the k-median problem, enabling approximate computation of the k-median cost of any set of k facilities in $\mathbb{R}^d$. Private mechanisms with small error, for other classes of queries have also been studied in several other works, see e.g. [BDMN05, BCD$^+$07, MT07, CM08, GLM$^+$10].

Lower bounds Dinur and Nissim [DN03] initiated the study of lower bounds on the amount of noise private mechanisms must add. They showed that any private mechanism that answers $\widetilde{O}(n)$ random subset sum queries about a set of N people each having a private bit must add noise $\Omega(\sqrt{n})$ to avoid nearly full disclosure of the database (*blatant non-privacy*). This implies that as one answers more and more questions, the amount of error needed per answer must grow to provide any kind of privacy guarantee. These results were strengthened by Dwork, McSherry and Talwar [DMT07], and by Dwork and Yekhanin [DY08]. However all these lower bounds protect against blatant non-privacy and cannot go beyond noise larger than $\min(\sqrt{d}, \sqrt{n})$ per answer, for d queries. Kasiviswanathan, Rudelson, Smith and Ullman [KRSU10] show lower bounds of the same nature ($\min(\sqrt{d}, \sqrt{n})$ for d queries) for contingency tables (rather than random counting queries). Their lower bounds also apply to (ε, δ)-differential privacy and are tight when ε and δ are constant. For the case of $d = 1$, Ghosh, Roughgarden and Sundararajan [GRS09] show that adding Laplace noise is in fact optimal in a very general decision-theoretic framework, for any symmetric decreasing loss function. For the case that all sum queries need to be answered (i.e. all queries of the form $f_P(y) = \sum_{i=1}^n P(y_i)$ where P is a 0-1 predicate), Dwork et al. [DMNS06] show that any

[2] Here the analyst asks one query at a time and receives an answer to each query before asking the next.

differentially private mechanism must add noise $\Omega(n)$. Rastogi et. al. [RSH07] show that half of such queries must have error $\Omega(\sqrt{n})$. Blum, Ligett and Roth [BLR08] show that any differentially private mechanism answering all (real-valued) halfspace queries must add noise $\Omega(n)$.

Our lower bounds (explained in Section 3) are in short based on a *packing argument*. Packing arguments have since found several further applications in proving lower bounds in differential privacy. Subsequent to our work, De [De11] gave a discrete analog of our lower bound. Among other results, he showed that our lower bounds also hold under the weaker requirement that the mechanism is only defined on non-negative integer points $\mathbb{Z}_+^N$ rather than $\mathbb{R}^N$. A lower bound related to our work was shown independently by Beimel, Kasiviswanathan and Nissim [BKN10]. They show a lower bound of $\Omega(\log d/\varepsilon)$ on the accuracy of any ε-differentially private mechanism for a specific set of d counting queries over a universe of size $N = d$.

1.3 Overview and organization of the paper

In this section we will give a broad overview of our proof and outline the remainder of the paper.

Section 2 contains some preliminary facts and definitions. Specifically, we describe a linear program that defines the optimal mechanism for any set of queries. This linear program (also studied in [GRS09] for the one-dimensional case) is exponential in size, but in principle, given any query and error function, can be used to compute the best mechanism for the given set of queries. Moreover, dual solutions to this linear program can be used to prove lower bounds on the error. However, the asymptotic behavior of the optimum value of these programs for multi-dimensional queries was not understood prior to this work. Our lower bounds can be reinterpreted as dual solutions to the linear program. The upper bounds give near optimal primal solutions. Also, our results lead to a polynomial-time approximation algorithm for the optimum when F is linear.

We prove our lower bound in Section 3. Given a query $F\colon \mathbb{R}^N \to \mathbb{R}^d$, our lower bound depends on the d-dimensional volume of $K = FB_1^N$. If the volume of K is large, then a packing argument shows that we can pack exponentially many points inside K so that each pair of points is far from each other. We then scale up K by a suitable factor λ. By linearity, all points within λK have preimages under F that are still λ-close in ℓ_1-distance. Hence, the definition of ε-differential privacy (by transitivity) enforces some constraint between these preimages. We can combine these observations so as to show that any differentially private mechanism M will have to put significant probability mass in exponentially many disjoint balls. This forces the mechanism to have large expected error.

We then introduce the K-norm mechanism in Section 4. Our mechanism computes Fx and then adds a noise vector to Fx. The key point here is that the noise vector is not independent of F as in previous works. Instead, informally speaking, the noise is tailored to the exact shape of $K = FB_1^N$. This is accomplished by picking a particular noise vector a with probability proportional to $\exp(-\varepsilon\|Fx - a\|_K)$. Here, $\|\cdot\|_K$ denotes the (Minkowski) norm defined by K. While our mechanism depends upon the query F, it does *not* depend on the particular database x. We can analyze our mechanism in terms of the expected Euclidean distance from the origin of a random point in K, i.e., $\mathbb{E}_{z\in K}\|z\|_2 = \mathrm{mr}(K)$. Arguing optimality of our mechanism hence boils down to relating $\mathrm{mr}(K)$ to the volume of K.

Indeed, using several results from convex geometry, we observe that our lower and upper bounds match up to constant factors when F is drawn at random from $\{-1,1\}^{d\times N}$. As it turns out the polytope K can be interpreted as the symmetric convex hull of the row vectors of F. When F is a random matrix, K is a well-studied random polytope. Some recent results on random polytopes give us suitable lower bounds on the volume and upper bounds on the average Euclidean norm. More generally, our bounds are tight whenever K is in isotropic position (as pointed out in Section 6). This condition intuitively gives a relation between volume and average distance from the origin. Our bounds are actually only tight up to a factor of L_K, the isotropic constant of K. A well-known conjecture from convex geometry, known as the Hyperplane Conjecture or Slicing Conjecture, implies that $L_K = O(1)$.

The problem is that when F is not drawn at random, K could be very far from isotropic. In this case, the K-norm mechanism by itself might actually perform poorly. We thus give a recursive variant of the K-norm mechanism in Section 7 which can handle non-isotropic bodies. Our approach is based on analyzing the covariance matrix of K in order to partition K into parts on which our earlier mechanism performs well. Assuming the Hyperplane conjecture, we derive bounds on the error of our mechanism that are optimal to within polylogarithmic factors.

The costly step in both of our mechanisms is sampling uniformly from high-dimensional convex bodies such as $K = FB_1^N$. To implement the sampling step efficiently, we will use geometric random walks. It can be shown that these random walks approach the uniform distribution over K in polynomial time. We will actually need convergence bounds in a metric strong enough to entail guarantees about differential privacy (i.e., a multiplicative rather than additive guarantee on the probability density).

Some complications arise, since we need to repeat the privacy and optimality analysis of our mechanisms in the presence of approximation errors (such as an approximate covariance matrix and an approximate separation oracle for K). The details can be found in Section 8.

Acknowledgments

We would like to thank Frank McSherry, Aaron Roth, Katrina Ligett, Indraneel Mukherjee, Nikhil Srivastava for several useful discussions, and Adam Smith for discussions and comments on a previous version of the paper.

2 Preliminaries

Notation We will write B_p^d to denote the unit ball of the p-norm in $\mathbb{R}^d$. When $K \subseteq \mathbb{R}^d$ is a centrally symmetric convex set, we write $\|\cdot\|_K$ for the (Minkowski) norm defined by K (i.e. $\|x\|_K = \inf\{r\colon x \in rK\}$). The ℓ_p-norms are denoted by $\|\cdot\|_p$, but we use $\|\cdot\|$ as a shorthand for the Euclidean norm $\|\cdot\|_2$. Given a function $F : \mathbb{R}^{d_1} \to \mathbb{R}^{d_2}$ and a set $K \in \mathbb{R}^{d_1}$, FK denotes the set $\{F(x) : x \in K\}$.

2.1 Differential Privacy

Definition 2.1. A *mechanism* M is a family of probability measures $M = \{\mu_x\colon x \in \mathbb{R}^N\}$ where each measure μ_x is defined on $\mathbb{R}^d$. A mechanism is called *ε-differentially private*, if for all

$x, y \in \mathbb{R}^N$ such that $\|x-y\|_1 \leqslant 1$, we have $\sup_{S \subseteq \mathbb{R}^d} \frac{\mu_x(S)}{\mu_y(S)} \leqslant \exp(\varepsilon)$, where the supremum runs over all measurable subsets $S \subseteq \mathbb{R}^d$.

A common weakening of ε-differential privacy is the following notion of *approximate* privacy.

Definition 2.2. A mechanism satisfies (ε,δ)*-differential privacy*, if for all $x, y \in \mathbb{R}^N$ we have $\mu_x(S) \leqslant \exp(\varepsilon)\mu_y(S) + \delta$ for all measurable subsets $S \subseteq \mathbb{R}^N$ whenever$\|x-y\|_1 \leqslant 1$. When $\delta = 0$, we will say the mechanism is ε-differentially private.

The definition of privacy is transitive in the following sense.

Fact 2.3. *If M is an ε-differentially private mechanism and $x, y \in \mathbb{R}^N$ satisfy $\|x-y\|_1 \leqslant k$, then for measurable $S \subseteq \mathbb{R}^d$ we have $\frac{\mu_x(S)}{\mu_y(S)} \leqslant \exp(\varepsilon k)$.*

Definition 2.4 (Error)**.** Let $F\colon \mathbb{R}^N \to \mathbb{R}^d$ and $\ell\colon \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$. We define the *$\ell$-error* of a mechanism M as $\mathrm{err}_\ell(M,F) = \sup_{x \in \mathbb{R}^N} \mathbb{E}_{a \sim \mu_x} \ell(a, Fx)$. Unless otherwise specified, we take ℓ to be the Euclidean distance.

Definition 2.5 (Sensitivity)**.** We will consider mappings F which possess the Lipschitz property, $\sup_{x \in B_1^N} \|Fx\|_1 \leqslant d$. In this case, we will say that F has *sensitivity d*.

Our goal is to show trade-offs between privacy and error. The following standard upper bound, usually called the Laplacian mechanism, is known.

Theorem 2.6 ([DMNS06])**.** *For any mapping $F\colon \mathbb{R}^N \to \mathbb{R}^d$ of sensitivity d and any $\varepsilon > 0$, there exists an ε-differentially private mechanism M with* $\mathrm{err}(M,F) = O(d\sqrt{d}/\varepsilon)$.

There is also the Gaussian mechanism which trades a dependence on $1/\delta$ for a better dependence on d.

Theorem 2.7 ([DKM+06])**.** *Let $\varepsilon, \delta > 0$. Then, for any mapping $F\colon \mathbb{R}^N \to \mathbb{R}^d$ of sensitivity d there exists a an (ε,δ)-differentially private mechanism M with* $\mathrm{err}(M,F) = O(d\sqrt{\log(1/\delta)}/\varepsilon)$.

2.2 Isotropic Position

Definition 2.8 (Isotropic Position)**.** We say a convex body $K \subseteq \mathbb{R}^d$ is in *isotropic position* with isotropic constant L_K if for every unit vector $v \in \mathbb{R}^d$,

$$\frac{1}{\mathrm{Vol}(K)} \int_K |\langle z, v\rangle|^2 dz = L_K^2 \mathrm{Vol}(K)^{2/d}. \tag{3}$$

Fact 2.9. *For every convex body $K \subseteq \mathbb{R}^d$, there is a volume-preserving linear transformation T such that TK is in isotropic position.*

For an arbitrary convex body K, its isotropic constant L_K can then be defined to be L_{TK} where T brings L to isotropic position. It is known (e.g. [MP89]) that T is unique up to an orthogonal transformation and thus this is well-defined. We refer the reader to the paper of Milman and Pajor [MP89], as well as the extensive survey of Giannopoulos [Gia03] for a proof of this fact and other facts regarding the isotropic constant.

2.3 Gamma Distribution

The *Gamma distribution* with shape parameter $k > 0$ and scale $\theta > 0$, denoted $\mathrm{Gamma}(k,\theta)$, is given by the probability density function

$$f(r;k,\theta) = r^{k-1}\frac{e^{-r/\theta}}{\Gamma(k)\theta^k}.$$

Here, $\Gamma(k) = \int e^{-r}r^{k-1}\,\mathrm{d}r$ denotes the Gamma function. We will need an expression for the moments of the Gamma distribution.

Fact 2.10. *Let $r \sim \mathrm{Gamma}(k,\theta)$. Then,*

$$\mathbb{E}[r^m] = \frac{\theta^m\Gamma(k+m)}{\Gamma(k)}. \tag{4}$$

Proof.

$$\begin{aligned}\mathbb{E}[r^m] = \int_{\mathbb{R}} r^{k+m-1}\frac{e^{-r/\theta}}{\Gamma(k)\theta^k}\,\mathrm{d}r &= \frac{1}{\Gamma(k)\theta^k}\int_{\mathbb{R}}(\theta r)^{k+m-1}e^{-r}\,\mathrm{d}\theta r \\ &= \frac{\Gamma(k+m)\theta^{k+m}}{\Gamma(k)\theta^k} = \frac{\Gamma(k+m)\theta^m}{\Gamma(k)}\end{aligned}$$

□

2.4 Linear Programming Characterization

Suppose that the set of databases is given by some set $\mathcal{D}$, and let $\sigma : \mathcal{D}\times\mathcal{D} \to \mathbb{R}_0$ be a distance function on $\mathcal{D}$. A query q is specified by an error function $\mathrm{err} : \mathcal{D}\times\mathcal{R} \to \mathbb{R}$. For example $\mathcal{D}$ could be the Hamming cube $\{0,1\}^N$ with σ being the Hamming distance. Given a query $F\colon \{0,1\}^N \to \mathbb{R}^d$, the error function could be $\mathrm{err}(x,a) = \|a - F(x)\|_2$ if we wish to compute $F(x)$ up to a small ℓ_2 error.

A mechanism is specified by a distribution μ_x on $\mathcal{R}$ for every $x \in \mathcal{D}$. Assume for simplicity that $\mathcal{D}$ and $\mathcal{R}$ are both finite. Thus a mechanism is fully defined by real numbers $\mu(x,a)$, where $\mu(x,a)$ is the probability that the mechanism outputs answer $a \in \mathcal{R}$ on databases $x \in \mathcal{D}$. The constraints on μ for an ε-differentially private mechanism are given by

$$\begin{aligned}&\sum_{a\in\mathcal{R}}\mu(x,a) = 1 && \forall x \in \mathcal{D}\\ &\mu(x,a) \geqslant 0 && \forall x \in \mathcal{D}, a \in \mathcal{R}\\ &\mu(x,a) \leqslant \exp(\varepsilon\sigma(x,x'))\mu(x',a) && \forall x,x' \in \mathcal{D}, a \in \mathcal{R}\end{aligned}$$

The expected error (under any given prior over databases) is then a linear function of the variables $\mu(x,a)$ and can be optimized. Similarly, the worse case (over databases) expected error can be minimized, and we will concentrate on this measure for the rest of the paper. However these linear programs can be prohibitive in size. Moreover, it is not a priori clear how one can use this formulation to understand the asymptotic behavior of the error of the optimum mechanism.

Our work leads to a constant approximation to the optimum of this linear program when F is a random in $\{-1,+1\}^{d\times N}$ and an $O(\log^{3/2} d)$-approximation otherwise.

3 Lower bounds via volume estimates

In this section we show that lower bounds on the volume of the convex body $FB_1^N \subseteq \mathbb{R}^d$ give rise to lower bounds on the error that any private mechanism must have with respect to F.

Definition 3.1. A set of points $Y \subseteq \mathbb{R}^d$ is called a *r-packing* if $\|y - y'\|_2 \geqslant r$ for any $y, y' \in Y, y \neq y'$.

Lemma 3.2. *Let $K \subseteq \mathbb{R}^d$ be a measurable set. Then, K contains an $\frac{1}{4}\mathrm{vr}(K)$-packing of size $\exp(d)$.*

Proof. By the definition of $\mathrm{vr}(K)$, the set K has the same volume as a ball of radius $\mathrm{vr}(K)$. Hence, any maximal $\frac{\mathrm{vr}(K)}{4}$-packing then has the desired property. □

Theorem 3.3. *Let $\varepsilon > 0$ and suppose $F : \mathbb{R}^N \to \mathbb{R}^d$ is a linear map and let $K = FB_1^N$. Then, every ε-differentially private mechanism M must have*

$$\mathrm{err}(M, F) \geqslant \Omega\left(\frac{d}{\varepsilon} \cdot \mathrm{vr}(K)\right). \tag{5}$$

Proof. Put $\lambda = d/2\varepsilon$. By Lemma 3.2, λK contains an $\frac{1}{4}\lambda\mathrm{vr}(K)$-packing Y of size $\exp(d)$. Let $X \subseteq \mathbb{R}^N$ be a set of arbitrarily chosen preimages of $y \in Y$ so that $|X| = |Y|$ and $FX = Y$. By linearity, $\lambda K = F(\lambda B_1^N)$ and hence we may assume that every $x \in X$ satisfies $\|x\|_1 \leqslant \lambda$.

We will now assume that $M = \{\mu_x : x \in \mathbb{R}^N\}$ is an ε-differentially private mechanism with error $\frac{d}{16\varepsilon}\mathrm{vr}(K)$ and lead this to a contradiction. By the assumption on the error, Markov's inequality implies that for all $x \in X$, we have $\mu_x(B_x) \geqslant 1/2$, where B_x is a ball of radius $\frac{d}{8\varepsilon}\mathrm{vr}(K) = \frac{\lambda}{4}\mathrm{vr}(K)$ centered at Fx. Since $Y = FX$ is an $\frac{\lambda}{4}\mathrm{vr}(K)$-packing, the balls $\{B_x : x \in X\}$ are disjoint. Since $\|x\|_1 \leqslant \lambda$, it follows from ε-differential privacy with Fact 2.3 that

$$\mu_0(B_x) \geqslant \exp(-\varepsilon\lambda)\mu_x(B_x) \geqslant \frac{\exp(-d/2)}{2}.$$

Since the balls B_x are pairwise disjoint,

$$1 \geqslant \mu_0\left(\bigcup_{x \in X} B_x\right) = \sum_{x \in X} \mu_0(B_x) \geqslant \frac{\exp(d)\exp(-d/2)}{2} > 1 \tag{6}$$

for $d \geqslant 2$. We have thus obtained a contradiction. □

We will later need the following generalization of the previous argument which gives a lower bound in the case where K is close to a lower dimensional subspace and hence the volume inside this subspace may give a stronger lower bound.

Corollary 3.4. *Let $\varepsilon > 0$ and suppose $F : \mathbb{R}^N \to \mathbb{R}^d$ is a linear map and let $K = FB_1^N$. Furthermore, let P denote the orthogonal projection operator of a k-dimensional subspace of $\mathbb{R}^d$ for some $1 \leqslant k \leqslant d$. Then, every ε-differentially private mechanism M must have*

$$\mathrm{err}(M, F) \geqslant \Omega\left(\frac{k \cdot \mathrm{vr}_k(PK)}{\varepsilon}\right) \tag{7}$$

where

$$\mathrm{vr}_k \stackrel{\mathrm{def}}{=} \frac{\mathrm{Vol}_k(PK)^{1/k}}{\mathrm{Vol}(B_2^k)}.$$

Proof. Note that a differentially private answer a to Fx can be projected down to a (differentially private) answer Pa to PFx. Since P has operator norm $\|P\| \leqslant 1$ this does not increase the error, i.e.,

$$\|Pa - PFx\| \leqslant \|P(a - Fx)\| \leqslant \|P\| \cdot \|a - Fx\| \leqslant \|a - Fx\|.$$

□

We will denote by $\mathrm{VolLB}(F, \varepsilon)$ the best lower bound obtainable in this manner, i.e.,

$$\mathrm{VolLB}(F, \varepsilon) = \sup_{k,P} \frac{k \cdot \mathrm{vr}_k(PFB_1^N)}{\varepsilon}$$

where the supremum is taken over all $k \in \{1, \dots, d\}$ and all k-dimensional orthogonal projections P.

3.1 Lower bounds for small number of queries

As shown previously, the task of proving lower bounds on the error of private mechanisms reduces to analyzing the volume of FB_1^N. When $d \leqslant \log N$ this is a straightforward task.

Fact 3.5. *Let $d \leqslant \log N$. Then, for all matrices $F \in [-1,1]^{d \times N}$, $\mathrm{Vol}(FB_1^N)^{1/d} \leqslant O(1)$. Furthermore, there is an explicit matrix F such that FB_1^N has maximum volume.*

Proof. Clearly, FB_1^N is always contained in B_∞^d and $\mathrm{Vol}(B_\infty^d)^{1/d} = 2$. On the other hand, since $n \geqslant 2^d$, we may take F to contain all points of the hypercube $H = \{\pm 1\}^d$ as its columns. In this case, $FB_1^N \supseteq B_\infty^d$. □

This lower bound shows that the standard upper bound from Theorem 2.6 is, in fact, optimal when $d \leqslant \log N$.

4 The K-norm mechanism

In this section we describe a new differentially private mechanism, which we call the K-norm mechanism.

Definition 4.1 (K-norm mechanism)**.** Given a linear map $F\colon \mathbb{R}^N \to \mathbb{R}^d$ and $\varepsilon > 0$, we let $K = FB_1^N$ and define the mechanism $\mathrm{KM}(F, d, \varepsilon) = \{\mu_x \colon x \in \mathbb{R}^N\}$ so that each measure μ_x is given by the probability density function

$$f(a) = Z^{-1} \exp(-\varepsilon \|Fx - a\|_K) \tag{8}$$

defined over $\mathbb{R}^d$. Here Z denotes the normalization constant

$$Z = \int_{\mathbb{R}^d} \exp(-\varepsilon \|Fx - a\|_K)\,\mathrm{d}a = \Gamma(d+1)\mathrm{Vol}(\varepsilon^{-1}K).$$

A more concrete view of the mechanism is provided by Figure 2 and justified in the next remark.

Remark 4.2. *We can sample from the distribution μ_x as follows:*

```
Input: Query $F \in \mathbb{R}^{d\times N}$, histogram $x \in \mathbb{R}^N$, privacy parameter $\varepsilon > 0$
    1. Sample $z \in \mathbb{R}^d$ uniformly at random from $K = FB_1^N$
    2. Sample $r \in \mathbb{R}$ from Gamma$(d+1, \varepsilon^{-1})$
Output: $Fx + rz$
```

Figure 2: Description of the d-dimensional K-norm mechanism.

1. *Sample r from the Gamma distribution with parameter $d+1$ and scale ε^{-1}, denoted* Gamma$(d+1,\varepsilon^{-1})$. *That is, r is distributed as*

$$\mathbb{P}\mathrm{r}(r > R) = \frac{1}{\varepsilon^{-(d+1)}\Gamma(d+1)} \int_R^\infty e^{-\varepsilon t} t^d \,\mathrm{d}t.$$

2. *Sample a uniformly from $Fx + rK$.*

Indeed, if $\|a - Fx\|_K = R$, then the distribution of a as above follows the probability density function

$$g(a) = \frac{1}{\varepsilon^{-(d+1)}\Gamma(d+1)} \int_R^\infty \frac{e^{-\varepsilon t} t^d}{\mathrm{Vol}(tK)} \,\mathrm{d}t = \frac{\int_R^\infty e^{-\varepsilon t}\,\mathrm{d}t}{\varepsilon^{-1}\Gamma(d+1)\mathrm{Vol}(\varepsilon^{-1}K)} = \frac{e^{-\varepsilon R}}{\Gamma(d+1)\mathrm{Vol}(\varepsilon^{-1}K)}, \tag{9}$$

where we used the fact that $\int_0^\infty e^{-\varepsilon t}\,\mathrm{d}t = \varepsilon^{-1}$. We thus see that this calculation is in agreement with (8). That is, $g(a) = f(a)$.

The next theorem shows that the K-norm mechanism is indeed differentially private. Moreover, we can express its error in terms of the *expected distance from the origin* of a random point in K.

Theorem 4.3. *Let $\varepsilon > 0$. Suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map and put $K = FB_1^N$. Then, the mechanism $KM(F,d,\varepsilon)$ is ε-differentially private, and for every $p > 0$ achieves the error bound*

$$\mathop{\mathbb{E}}_{a\sim\mu_x} \|Fx - a\|_2^p \leqslant \frac{\Gamma(d+1+p)}{\varepsilon^p\Gamma(d)} \mathop{\mathbb{E}}_{z\in K} \|z\|_2^p. \tag{10}$$

In particular, the ℓ_2-error is at most

$$\frac{d+1}{\varepsilon} \mathop{\mathbb{E}}_{z\in K} \|z\|_2 = \frac{d+1}{\varepsilon} \cdot \mathrm{mr}(K).$$

Proof. To argue the error bound, we will follow Remark 4.2. Let $D = \mathrm{Gamma}(d+1, 1/\varepsilon)$. For all $x \in \mathbb{R}^N$,

$$\begin{aligned}\mathop{\mathbb{E}}_{a\sim\mu_x} \|Fx - a\|^p = \mathop{\mathbb{E}}_{a\sim\mu_0} \|a\|^p = \mathop{\mathbb{E}}_{r\sim D} \mathop{\mathbb{E}}_{a\in rK} \|a\|^p &= \left[\mathop{\mathbb{E}}_{r\sim D} r^p\right] \mathop{\mathbb{E}}_{z\in K} \|z\|^p \\ &= \frac{\Gamma(d+1+p)}{\varepsilon^p\Gamma(d+1)} \mathop{\mathbb{E}}_{z\in K} \|z\|^p. \qquad \text{(by Fact (2.10))}\end{aligned}$$

When $p = 1$, $\frac{\Gamma(d+1+p)}{\Gamma(d+1)} = d + 1$.

Privacy follows from the fact that the mechanism is a special case of the exponential mechanism [MT07]. For completeness, we repeat the argument.

Suppose that $\|x\|_1 \leqslant 1$. It suffices to show that for all $a \in \mathbb{R}^d$, the densities of μ_0 and μ_x are within multiplicative $\exp(\varepsilon)$, i.e.,

$$\frac{Z^{-1}e^{-\varepsilon\|a\|_K}}{Z^{-1}e^{-\varepsilon\|Fx-a\|_K}} = e^{\varepsilon(\|Fx-a\|_K-\|a\|_K)} \leqslant e^{\varepsilon\|Fx\|_K} \leqslant e^{\varepsilon}.$$

where in the first inequality we used the triangle inequality for $\|\cdot\|_K$. In the second step we used that $x \in B_1^N$ and hence $Fx \in FB_1^N = K$ which means $\|Fx\|_K \leqslant 1$. Hence, the mechanism satisfies ε-differential privacy. □

5 Optimality for random queries and isotropic bodies

In this section, we will show that our upper bound matches our lower bound when F is a random query. A key observation is that FB_1^N is the *symmetric* convex hull of N (random) points $\{v_1,\ldots,v_n\} \subseteq \mathbb{R}^d$, i.e., the convex hull of $\{\pm v_1,\ldots,\pm v_n\}$, where $v_i \in \mathbb{R}^d$ is the ith column of F. The symmetric convex hull of random points has been studied extensively in the theory of random polytopes. A recent result of Litvak, Pajor, Rudelson and Tomczak-Jaegermann [LPRN05] gives the following lower bound on the volume of the convex hull. For convenience, we state their result in terms of volume radius.

Theorem 5.1 ([LPRN05])**.** *Let $2d \leqslant n \leqslant 2^d$ and let F denote a random $d \times N$ Bernoulli matrix. Then,*

$$\mathrm{vr}\left(FB_1^N\right) \geqslant \Omega(1)\sqrt{\log\left(\frac{N}{d}\right)}, \tag{11}$$

with probability $1-\exp(-\Omega(d^\beta n^{1-\beta}))$ for any $\beta \in (0,\frac{1}{2})$. Furthermore, there is an explicit construction of n points in $\{-1,1\}^d$ whose convex hull achieves the same volume.

We are mostly interested in the range where $N \gg d\log d$ in which case the theorem was already proved by Giannopoulos and Hartzoulaki [GH02] (up to a weaker bound in the probability and without the explicit construction).

The bound in (11) is tight up to constant factors. A well known result [BF88] shows that if K is the convex hull of any N points on the sphere in $\mathbb{R}^d$ of radius $\sqrt{d}$, then

$$\mathrm{vr}(K) \leqslant O(1)\sqrt{\log\left(\frac{N}{d}\right)}. \tag{12}$$

Notice, that in our case $K = FB_1^N \subseteq B_\infty^d \subseteq \sqrt{d}B_2^d$ and in fact the vertices of K are points on the $(d-1)$-dimensional sphere of radius $\sqrt{d}$. However, equation (11) states that the normalized volume of the random polytope K will be proportional to the volume of the Euclidean ball of radius $\sqrt{\log(N/d)}$ rather than $\sqrt{d}$. When $d \gg \log n$, this means that the volume of K will be tiny compared to the volume of the infinity ball B_∞^d. By combining the volume lower bound with Theorem 3.3, we get the following lower bound on the error of private mechanisms.

Theorem 5.2. *Let $\varepsilon > 0$ and $0 < d \leqslant N/2$. Then, for almost all matrices $F \in \{-1,1\}^{d\times N}$, every ε-differentially private mechanism M must have*

$$\operatorname{err}(M,F) \geqslant \Omega(d/\varepsilon)\cdot \min\left\{\sqrt{d}, \sqrt{\log\left(\frac{N}{d}\right)}\right\}. \tag{13}$$

5.1 A separation result

We use this paragraph to point out that our lower bound immediately implies a separation between (ε,δ)-differential privacy (see Definition 2.2) and $(\varepsilon,0)$-differential privacy. The Gaussian mechanism (see Theorem 2.7) gives (ε,δ)-differential privacy with error $o\left(\varepsilon^{-1}\sqrt{\log(N/d)}\right)$ as long as $\delta \geqslant 1/n^{o(1)}$. Our lower bound in Theorem 5.2 on the other hand states that the error of any ε-differentially private mechanism must be $\Omega\left(\varepsilon^{-1}\sqrt{\log(N/d)}\right)$ (assuming $d \gg \log(n)$). We get the strongest separation when $d \leqslant \log(n)$ and δ is constant. In this case, our lower bound is a factor $\sqrt{d}$ larger than the upper bound for approximate differential privacy.

5.2 Upper bound on average Euclidean norm

Klartag and Kozma [KK09] recently gave a bound on the quantity $\mathbb{E}_{z\sim K}\|z\|$ when $K = FB_1^N$ for random F.

Theorem 5.3 ([KK09])**.** *Let F be a random $d\times N$ Bernoulli matrix and put $K = FB_1^N$. Then, there is a constant $C > 0$ so that with probability greater than $1 - Ce^{-O(n)}$,*

$$\frac{1}{\operatorname{Vol}(K)}\int_{z\in K}\|z\|^2\,\mathrm{d}z \leqslant C\log\left(\frac{N}{d}\right). \tag{14}$$

An application of Jensen's inequality thus gives us the following corollary.

Corollary 5.4. *Let $\varepsilon > 0$ and $0 < d \leqslant N/2$. Then, for almost all matrices $F \in \{-1,1\}^{d\times N}$, the mechanism $KM(F,d,\varepsilon)$ is ε-differentially private with error at most*

$$O\left(\frac{d}{\varepsilon}\right)\cdot \min\left\{\sqrt{d}, \sqrt{\log\left(\frac{N}{d}\right)}\right\}. \tag{15}$$

6 Approximately isotropic bodies

The following definition is a relaxation of isotropic position used in literature (e.g., [KLS97])

Definition 6.1 (Approximately Isotropic Position)**.** We say a convex body $K \subseteq \mathbb{R}^d$ is in *c-approximately isotropic position* if for every unit vector $v \in \mathbb{R}^d$,

$$\frac{1}{\operatorname{Vol}(K)}\int_K |\langle z,v\rangle|^2\,\mathrm{d}z \leqslant c^2 L_K^2 \operatorname{Vol}(K)^{\frac{2}{d}}. \tag{16}$$

The results of Klartag and Kozma [KK09] referred to in the previous section show that the symmetric convex hull n random points from the d-dimensional hypercube are in $O(1)$-approximately isotropic position and have $L_K = O(1)$. More generally, the K-norm mechanism can be shown to be approximately optimal whenever K is nearly isotropic.

Theorem 6.2 (Theorem 1.2 restated)**.** *Let $\varepsilon > 0$. Suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map such that $K = FB_1^N$ is in c-approximately isotropic position. Then, the K-norm mechanism is ε-differentially private and has error at most $O(cL_K) \cdot \frac{d\mathrm{vr}(K)}{\varepsilon}$*

Proof. By Theorem 4.3, the K-norm mechanism is ε-differentially private and has error $\frac{d+1}{\varepsilon}\mathbb{E}_{z\sim K}\|z\|$. By the definition of the approximately isotropic position, we have: $\mathbb{E}_{z\sim K}\|z\|^2 \leqslant d \cdot c^2 L_K^2 \mathrm{Vol}(K)^{2/d}$. By Jensen's inequality,

$$\frac{d+1}{\varepsilon}\mathop{\mathbb{E}}_{z\sim K}\|z\| \leqslant \frac{d+1}{\varepsilon}\sqrt{\mathop{\mathbb{E}}_{z\sim K}\|z\|^2} \leqslant O\left(\frac{cL_K d\sqrt{d}\mathrm{Vol}(K)^{1/d}}{\varepsilon}\right) = O\left(\frac{cL_K d\mathrm{vr}(K)}{\varepsilon}\right).$$

□

We can see that the previous upper bound is tight up to a factor of cL_K. Estimating L_K for general convex bodies is a well-known open problem in convex geometry. The best known upper bound for a general convex body $K \subseteq \mathbb{R}^d$ is $L_K \leqslant O(d^{1/4})$ due to Klartag [Kla06], improving over the estimate $L_K \leqslant O(d^{1/4}\log d)$ of Bourgain from '91. The conjecture is that $L_K = O(1)$.

Conjecture 6.3 (Hyperplane Conjecture)**.** *There exists $C > 0$ such that for every d and every convex set $K \subseteq \mathbb{R}^d$, $L_K < C$.*

Assuming this conjecture we get matching bounds for approximately isotropic convex bodies.

Theorem 6.4. *Let $\varepsilon > 0$. Assuming the hyperplane conjecture, for every $F \in [-1,1]^{d\times N}$ such that $K = FB_1^N$ is c-approximately isotropic, the K-norm mechanism $KM(F,d,\varepsilon)$ is ε-differentially private with error at most*

$$O\left(\frac{cd}{\varepsilon}\right)\cdot \min\left\{\sqrt{d}, \sqrt{\log\left(\frac{N}{d}\right)}\right\}. \tag{17}$$

7 Non-isotropic bodies

While the mechanism of the previous sections is near-optimal for near-isotropic queries, it can be far from optimal if K is far from isotropic. For example, suppose the matrix F has random entries from $\{+1,-1\}$ in the first row, and (say) from $\{\frac{1}{d^2}, -\frac{1}{d^2}\}$ in the remaining rows. While the Laplacian mechanism will add $O(\frac{1}{\varepsilon})$ noise to the first co-ordinate of Fx, the K-norm mechanism will add noise $O(d/\varepsilon)$ to the first co-ordinate. Moreover, the volume lower bound VolLB is at most $O(\varepsilon^{-1}\sqrt{d})$. Rotating F by a random rotation gives, w.h.p., a query for which the Laplacian mechanism adds ℓ_2 error $O(d/\varepsilon)$. For such a body, the Laplacian and the K-norm mechanisms, as well as the VolLB are far from optimal.

In this section, we will design a recursive mechanism that can handle such non-isotropic convex bodies. To this end, we will need to introduce a few more notions from convex geometry.

Suppose $K \subseteq \mathbb{R}^d$ is a centered convex body, i.e. $\int_K x\,\mathrm{d}x = 0$. The *covariance matrix of* K, denoted M_K is the $d \times d$ matrix with entry ij equal to $M_{ij} = \frac{1}{\mathrm{Vol}(K)}\int_K x_i x_j\,\mathrm{d}x$. That is, M_K is the covariance matrix of the uniform distribution over K.

```
NiKM($F,d,\varepsilon$):
  1. Let $K = FB_1^N$. Let $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_d$ denote the eigenvalues of the covariance matrix $M_K$. Pick a corresponding orthonormal eigenbasis $u_1,\ldots,u_d$.
  2. Let $d' = \lfloor d/2 \rfloor$ and let $U = \mathrm{span}\{u_1,\ldots,u_{d'}\}$ and $V = \mathrm{span}\{u_{d'+1},\ldots,v_d\}$.
  3. Sample $a \sim \mathrm{KM}(F,d,\varepsilon)$.
  4. If $d = 1$, output $P_V a$. Otherwise, output $\mathrm{NiKM}(P_U F, d', \varepsilon) + P_V a$.
```

Figure 3: Mechanism for non-isotropic bodies

7.1 A recursive mechanism

Having defined the covariance matrix, we can describe a recursive mechanism for the case when K is not in isotropic position. The idea of the mechanism is to act differently on different eigenspaces of the covariance matrix. Specifically, the mechanism will use a lower-dimensional version of $\mathrm{KM}(F,d',\varepsilon)$ on subspaces corresponding to few large eigenvalues.

Our mechanism, called $\mathrm{NiKM}(F,d,\varepsilon)$, is given a linear mapping $F\colon \mathbb{R}^N \to \mathbb{R}^d$, and parameters $d \in \mathbb{N}, \varepsilon > 0$. The mechanism proceeds recursively by partitioning the convex body K into two parts defined by the middle eigenvalue of M_K. On one part it will act according to the K-norm mechanism. On the other part, it will descend recursively. The mechanism is described in Figure 3

Remark 7.1. *The image of $P_U F$ above is a d'-dimensional subspace of $\mathbb{R}^d$. We assume that in the recursive call NiKM($P_U F, d', \varepsilon$), the K-norm mechanism is applied to a basis of this subspace. However, formally the output is a d-dimensional vector.*

To analyze our mechanism, first observe that the recursive calls terminate after at most $\log d$ steps. For each recursive step $m \in \{0,\ldots,\log d\}$, let a_m denote the distribution over the output of the K_m-norm mechanism in step 3. Here, K_m denotes the d_m-dimensional body given in step m.

Lemma 7.2. *The mechanism NiKM(F,d,ε) satisfies ($\varepsilon \log d$)-differential privacy.*

Proof. We claim that for every step $m \in \{0,\ldots,\log d\}$, the distribution over a_m is ε-differentially private. Notice that this claim implies the lemma, since the joint distribution of $a_0,a_1,\ldots,a_m$ is $\varepsilon\log(d)$-differentially private. In particular, this is true for the final output of the mechanism as it is a function of $a_0,\ldots,a_m$.

To see why the claim is true, observe that each K_m is the d_m-dimensional image of the ℓ_1-ball under a linear mapping. Hence, the K_m-norm mechanism guarantees ε-differential privacy by Theorem 4.3. □

The error analysis of our mechanism requires more work. In particular, we need to understand how the volume of $P_U K$ compares to the norm of $P_V a$. As a first step we will analyze the volume of $P_U K$.

7.2 Volume in eigenspaces of the covariance matrix

Our goal in this section is to express the volume of K in eigenspaces of the covariance matrix in terms of the eigenvalues of the covariance matrix. This will be needed in the analysis of our mechanism for non-isotropic bodies.

We start with a formula for the volume of central sections of isotropic bodies. This result can be found in [MP89].

Proposition 7.3. *Let $K \subseteq \mathbb{R}^d$ be an isotropic body of unit volume. Let E denote a k-dimensional subspace for $1 \leqslant k \leqslant d$. Then,*

$$\mathrm{Vol}_k(E \cap K)^{1/(d-k)} = \Theta\left(\frac{L_{B_K}}{L_K}\right).$$

Here, B_K is an explicitly defined isotropic convex body.

From here on, for an isotropic body K, let $\alpha_K = \Omega(L_{B_K}/L_K)$ be a lower bound on $\mathrm{Vol}_k(E \cap K)^{1/(d-k)}$ implied by the above proposition. For a non-isotropic K, let α_K be α_{TK} when T is the map the brings K into isotropic position. Notice that if the Hyperplane Conjecture is true, then $\alpha_K = \Omega(1)$. Moreover, α_K is $\Omega(d^{\frac{1}{4}})$ due to the results of [Kla06].

Corollary 7.4. *Let $K \subseteq \mathbb{R}^d$ be an isotropic body with* $\mathrm{Vol}(K) = 1$. *Let E denote a k-dimensional subspace for $1 \leqslant k \leqslant d$ and let P denote an orthogonal projection operator onto the subspace E. Then,*

$$\mathrm{Vol}_k(PK)^{1/(d-k)} \geqslant \alpha_K.$$

Proof. Observe that the PK contains $E \cap K$ since P is the identity on E. □

We cannot immediately use these results since they only apply to isotropic bodies and we are specifically dealing with non-isotropic bodies. The trick is to apply the previous results after transforming K into an isotropic body while keeping track how much this transformation changed the volume.

As a first step, the following lemma relates the volume of projections of an arbitrary convex body K to the volume of projections of TK for some linear mapping T.

Lemma 7.5. *Let $K \subseteq \mathbb{R}^d$ be a symmetric convex body. Let T be a linear map which has eigenvectors $u_1, \ldots, u_d$ with eigenvalues $\lambda_1, \ldots, \lambda_d$. Let $1 \leqslant k \leqslant d$ and suppose $E = \mathrm{span}\{u_1, u_2, \ldots, u_k\}$, Denote by P be the projection operator onto the subspace E. Then,*

$$\mathrm{Vol}_k(PK) \geqslant \mathrm{Vol}_k(PTK) \prod_{i=1}^{k} \lambda_i^{-1}.$$

Proof. For simplicity, we assume that the eigenvectors of T are the standard basis vectors $e_1, \ldots, e_d$; this is easily achieved by applying a rotation to K. Now, it is easy to verify that $P = PT^{-1}T = SPT$ where $S = \mathrm{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_{\mathrm{k}}^{-1}, 0, \ldots, 0)$. Thus we can write

$$\mathrm{Vol}_k(PK) = \det(S_{|E}) \mathrm{Vol}_k(PTK) = \frac{1}{\prod_{i=1}^{k} \lambda_i} \mathrm{Vol}_k(PTK).$$

□

Before we can finish our discussion, we will need the fact that the isotropic constant of K can be expressed in terms of the determinant of M_K.

Fact 7.6 ([Gia03, MP89])**.** *Let $K \subseteq \mathbb{R}^d$ be a convex body of unit volume. Then,*

$$L_K^2 \mathrm{Vol}(K)^{\frac{2}{d}} = \det(M_K)^{1/d}. \tag{18}$$

Moreover, K is in isotropic position if and only if $M_K = L_K^2 \mathrm{Vol}(K)^{2/d} I$.

We conclude with the following Proposition 7.7.

Proposition 7.7. *Let $K \subseteq \mathbb{R}^d$ be a symmetric convex body. Let M_K have eigenvectors $u_1, \ldots, u_d$ with eigenvalues $\sigma_1, \ldots, \sigma_d$. Let $1 \leqslant k \leqslant \lceil \frac{d}{2} \rceil$ with and suppose $E = \mathrm{span}\{u_1, u_2, \ldots, u_k\}$. Denote by P be the projection operator onto the subspace E. Then,*

$$\mathrm{Vol}_k(PK)^{1/(d-k)} \geqslant \Omega(1) \cdot \alpha_K \left(\prod_{i=1}^{k} \sigma_i^{1/2} \right)^{1/(d-k)}, \tag{19}$$

where α_K is $\Omega(1/d^{\frac{1}{4}})$. Moreover, assuming the Hyperplane conjecture, $\alpha_K \geqslant \Omega(1)$.

Proof. Consider the linear mapping $T = M_K^{-1/2}$. this is well defined since M_K is a positive symmetric matrix. It is easy to see that after applying T, we have $M_{TK} = I$. Hence, by Fact 7.6, TK is in isotropic position and has volume $\mathrm{Vol}(TK)^{1/d} = 1/L_{TK} = 1/L_K$, since $\det(M_{TK}) = 1$. Scaling TK by $\lambda = L_K^{1/d}$ hence results in $\mathrm{Vol}(\lambda TK) = 1$. Noting that λT has eigenvalues $\lambda \sigma_1^{-\frac{1}{2}}, \lambda \sigma_2^{-\frac{1}{2}}, \ldots, \lambda \sigma_d^{-\frac{1}{2}}$, we can apply Lemma 7.5 and get

$$\mathrm{Vol}_k(PK) \geqslant \mathrm{Vol}_k(P\lambda TK) \prod_{i=1}^{k} \frac{\sqrt{\sigma_i}}{\lambda}$$

Since λTK is in isotropic position and has unit volume, Corollary 7.4 implies that

$$\mathrm{Vol}_k(P\lambda TK)^{1/(d-k)} \geqslant \alpha_K. \tag{20}$$

Thus the required inequality holds with an additional $\lambda^{-\frac{k}{d-k}}$ term. By assumption on k, $\frac{k}{d-k}$ is at most 2. Moreover, $\lambda = L_K^{1/d} \leqslant d^{1/d} \leqslant 2$, so that this additional term is a constant. As discussed above, α_K is $\Omega(d^{-\frac{1}{4}})$ by [Kla06], and $\Omega(1)$ assuming the Hyperplane Conjecture 6.3. Hence the claim. □

7.3 Arguing near optimality of our mechanism

Our next lemma shows that the expected squared Euclidean error added by our algorithm in each step is bounded by the square of the optimum. We will first need the following fact.

Fact 7.8. *Let $K \subseteq \mathbb{R}^d$ be a centered convex body. Let $\sigma_1 \geqslant \sigma_2 \geqslant \ldots \geqslant \sigma_d$ denote the eigenvalues of M_K with a corresponding orthonormal eigenbasis $u_1, \ldots, u_d$. Then, for all $1 \leqslant i \leqslant d$,*

$$\sigma_i = \max_{\theta} \mathop{\mathbb{E}}_{x \in K} \langle \theta, x \rangle^2 \tag{21}$$

where the maximum runs over all $\theta \in \mathbb{S}^{d-1}$ such that θ is orthogonal to $u_1, u_2, \ldots, u_{i-1}$.

Lemma 7.9. *Let a denote the random variable returned by the K-norm mechanism in step (3) in the above description of NiKM(F, d, ε). Then,*

$$\mathrm{VolLB}(F,\varepsilon)^2 \geqslant \Omega(\alpha_K^2)\,\mathbb{E}\|P_V a\|_2^2.$$

Proof. For simplicity, we will assume that d is even and hence $d - d' = d'$. The analysis of the K-norm mechanism (Theorem 4.3 with $p = 2$) shows that the random variable a returned by the K-norm mechanism in step (3) satisfies

$$\begin{aligned}
\mathbb{E}\|P_V a\|_2^2 = \frac{\Gamma(d+3)}{\varepsilon^2\Gamma(d+1)} = \frac{(d+2)(d+1)}{\varepsilon^2} \mathop{\mathbb{E}}_{z\in K}\|P_V z\|_2^2 & \\
= O\left(\frac{d^2}{\varepsilon^2}\right)\sum_{i=d'+1}^{d} \mathop{\mathbb{E}}_{z\in K}\langle z, u_i\rangle^2 & \\
= O\left(\frac{d^2}{\varepsilon^2}\right)\sum_{i=d'+1}^{d} \sigma_i & \qquad \text{(by Fact 7.8)} \\
\leqslant O\left(\frac{d^3}{\varepsilon^2}\right)\cdot \sigma_{d'+1}. & \qquad (22)
\end{aligned}$$

On the other hand, by the definition of VolLB,

$$\begin{aligned}
\mathrm{VolLB}(F,\varepsilon)^2 &\geqslant \Omega\left(\frac{d^3}{\varepsilon^2}\right)\cdot \mathrm{Vol}_{d'}(P_U K)^{2/d'} \\
&\geqslant \Omega\left(\frac{d^3}{\varepsilon^2}\right)\alpha_K^2\left(\prod_{i=1}^{d'}\sigma_i\right)^{1/d'} && \text{(by Proposition 7.7)} \\
&\geqslant \Omega\left(\frac{d^3}{\varepsilon^2}\right)\alpha_K^2\sigma_{d'}.
\end{aligned}$$

Since $\sigma_{d'} \geqslant \sigma_{d'+1}$, it follows that

$$\mathrm{VolLB}(F,\varepsilon)^2 \geqslant \Omega(\alpha_K^2)\,\mathbb{E}\|P_V a\|^2. \qquad (23)$$

The case of odd d is similar except that we define K' to be the projection onto the first $d'+1$ eigenvectors. □

Lemma 7.10. *Assume the Hyperplane Conjecture. Then, the ℓ_2-error of the mechanism NiKM(F, d, ε) satisfies*

$$\mathrm{err}(NiKM, F) \leqslant O\left(\sqrt{\log(d)}\cdot \mathrm{VolLB}(F,\varepsilon)\right).$$

Proof. We have to sum up the error over all recursive calls of the mechanism. To this end, let $P_{V_m}a_m$ denote the output of the K-norm mechanism a_m in step m projected to the corresponding subspace V_m. Also, let $a \in \mathbb{R}^d$ denote the final output of our mechanism. We then have,

$$\begin{aligned}
\mathbb{E}\|a\|_2 &\leqslant \sqrt{\mathbb{E}\|a\|_2^2} && \text{(Jensen's inequality)}\\
&= \sqrt{\sum_{m=1}^{\log d} \mathbb{E}\|P_{V_m} a_m\|_2^2}\\
&\leqslant \sqrt{\sum_{m=1}^{\log d} O(\alpha_{K_m}^{-2}) \cdot \mathrm{VolLB}(F,\varepsilon)^2} && \text{(by Lemma 7.9)}\\
&\leqslant O(\sqrt{\log d})\Big(\max_m \alpha_{K_m}^{-1}\Big)\mathrm{VolLB}(F,\varepsilon).
\end{aligned}$$

Here we have used the fact that $\mathrm{VolLB}(F,\varepsilon) \geqslant \mathrm{VolLB}(P_U F,\varepsilon)$. Finally, the hyperplane conjecture implies $\max_m \alpha_{K_m}^{-1} = O(1)$. □

Corollary 7.11. *Let $\varepsilon > 0$. Suppose $F\colon \mathbb{R}^N \to \mathbb{R}^d$ is a linear map. Further, assume the hyperplane conjecture. Then, there is an ε-differentially private mechanism M with error*

$$\mathrm{err}(M,F) \leqslant O(\log(d)^{3/2} \cdot \mathrm{VolLB}(F,\varepsilon)).$$

Proof. The mechanism $\mathrm{NiKM}(F,d,\varepsilon/\log(d))$ satisfies ε-differential privacy, by Lemma 7.2. The error is at most $\log(d)\sqrt{\log d} \cdot \mathrm{VolLB}(F,\varepsilon)$ as a direct consequence of Lemma 7.10. □

Thus our lower bound VolLB and the mechanism NiKM are both within $O(\log^{3/2} d)$ of the optimum.

8 More efficient implementation using geometric random walks

We will first describe how to implement our basic mechanism $\mathrm{KM}(F,d,\varepsilon)$. As we saw, this mechanism is optimal when FB_1^N is in roughly isotropic position. In Section 8.1, we extend our discussion to $\mathrm{NiKM}(F,d,\varepsilon)$ thus getting an efficient nearly optimal mechanism even when FB_1^N is not in isotropic position.

Recall that we first sample $r \sim \mathrm{Gamma}(d+1,\varepsilon^{-1})$ and then sample a point a uniformly at random from rK. The first step poses no difficulty. Indeed, when $U_1,\dots,U_d$ are independently distributed uniformly over the interval $(0,1]$, then a standard fact tells us that

$$\frac{1}{\varepsilon}\sum_{i=1}^{d+1} -\ln(U_i) \sim \mathrm{Gamma}(d+1,\varepsilon^{-1}).$$

Sampling uniformly from K on the other hand may be hard. However, there are ways of sampling nearly uniform points from K using various types of rapidly mixing random walks. In this section, we will use the *Grid Walk* for simplicity even though there are more efficient walks that will work for us. We refer the reader to the survey of Vempala [Vem05] or the original paper of Dyer, Frieze and Kannan [DFK91] for a description of the Grid walk and background information. Informally, the Grid walk samples nearly uniformly from a grid inside K, i.e., $\mathcal{L} \cap K$ where we take $\mathcal{L} = \frac{1}{d^2}\mathbb{Z}^d$. The Grid Walk poses two requirements on K:

1. Membership in K can be decided efficiently.

2. K is bounded, in the sense that $B_2^d \subseteq K \subseteq dB_2^d$.

Both conditions are naturally satisfied in our case where $K = FB_1^N$ for some $F \in [-1,1]^{d\times N}$. Indeed, $K \subseteq B_\infty^d \subseteq \sqrt{d}B_2^d$ and we assume throughout this section that $B_2^d \subseteq K$. This is without loss of generality, since we may replace K by $K' = K + B_2^d$. This will only increase the noise level by 1 in Euclidean distance. Notice that K' is convex.

The exact notion of membership oracle that we need is given in the next definition.

Definition 8.1. A *β-weak* separation oracle for K is a blackbox that says 'YES' when given $u \in \mathbb{R}^d$ with $(u + \beta B_2^d) \subseteq K$ and outputs 'NO' when $u \notin K + \beta B_2^d$.

In order to implement a weak membership oracle for K, we need to be able to decide for a given $a \in \mathbb{R}^d$, whether there exists an $x \in B_1^N$ such that $Fx = a$. These constraints can be encoded using a linear program. In the case of K' this can be done using standard convex programming techniques [GLS94].

Lemma 8.2. *Let $\beta > 0$. We can implement a β-weak separation oracle for K and also $K + B_2^d$ in time polynomial in $N, d, 1/\beta$.*

The mixing time of the Grid walk is usually quantified in terms of the total variation (or L_1) distance between the random walk and its stationary distribution. The stationary distribution of the grid Walk is the uniform distribution over $\mathcal{L} \cap K$. Standard arguments show that an L_1-bound gives us (ε,δ)-differential privacy where δ can be made exponentially small in polynomial time. In order to get $(\varepsilon, 0)$-differential privacy we instead need a multiplicative guarantee on the density of the random walk at each point in K.

It is not difficult to show that the Grid Walk actually satisfies mixing bounds in a pointwise multiplicative sense. We also need to take care of the fact that the stationary distribution is a priori not uniform over K.

Theorem 8.3. *There is a mechanism M' with expected runtime polynomial in N, d and ε^{-1} such that*

1. *M' is ε-differentially private,*

2. $\mathrm{err}(M', F) = O(\mathrm{err}(KM(F, d, \varepsilon), F))$.

To prove the theorem we need the next lemma that essentially directly follows from [DFK91].

Lemma 8.4. *There is a randomized algorithm* $\mathrm{Sample}(K,\beta)$ *running in time* $\mathrm{poly}(N, d, \beta^{-1})$ *whose output distribution is pointwise within a $(1 \pm \beta)$ factor from the uniform distribution over a body $\widehat{K}$ such that $K \subseteq \widehat{K} \subseteq (1+\beta)K$.*

Proof sketch. In order to implement $\mathrm{Sample}(K,\beta)$ we consider the t-step grid walk over K using a $\beta/2$-weak separation oracle and a fine enough grid $\mathcal{L} = \frac{\beta}{2}\mathbb{Z}^d \cap dB_2^d$ where $\beta = \mathrm{poly}(1/d)$. By Lemma 8.2, we can implement the separation oracle in time $\mathrm{poly}(N, d, \beta^{-1})$.

It is known [DFK91] that the t-step Grid Walk for $t = \mathrm{poly}(d, {}^1\!/\!_\beta, \log({}^1\!/\!_\Delta))$ gets within statistical distance at most Δ of the uniform distribution over $\mathcal{L} \cap K'$ where K' is a body satisfying $(1 - \beta/2)K \subseteq K' \subseteq (1 + \beta/2)K$, Setting Δ to be much smaller than the number of atoms in the Grid Walk, i.e., $\Delta = \text{poly}(\varepsilon\beta/|\mathcal{L}|)$ we end up with a distribution that is point-wise within at $(1 \pm \beta)$-factor of the uniform distirbution over $\mathcal{L} \cap K'$. Note that $\log(1/\Delta) = \text{poly}(d, 1/\varepsilon, 1/\beta)$.

Let Z be a sample from the grid walk described above, and let $\widehat{Z}$ be a random point from an ℓ_∞-ball of radius $\beta/4$ centered at Z. Then $\widehat{Z}$ is a nearly uniform sample from a body $\widehat{K}$ which has the property that $(1-\beta)K \subseteq \widehat{K} \subseteq (1+\beta)K$. □

Proof of Theorem 8.3. Our algorithm first samples $r \sim \text{Gamma}(d+1, 10\varepsilon^{-1})$, and then outputs $Fx + rz$ where z is the output of $\text{Sample}(K, \beta)$ for $\beta = \min\{\varepsilon/10d, 1/10r\}$. Let g denote the resulting density function of our mechanism. Let $a \in \mathbb{R}^d$. We will compare $g(a)$ to the density $f(a)$ of $\text{KM}(F, d, \varepsilon/10)$.

We can repeat the calculation for the density at a point a in equation (9). Indeed for a point a with $\|a - Fx\|_K = R$, the density at a conditioned on a sample r from the Gamma distribution, is $(1 \pm \beta)/\text{Vol}(r\widehat{K})$ whenever $a \in r\widehat{K}$, and zero otherwise. By our choice of β, $\text{Vol}(\widehat{K}) = (1 \pm \varepsilon/5)\text{Vol}(K)$. Moreover, since $K \subseteq \widehat{K} \subseteq (1+\beta)K$ we have $a \in r\widehat{K}$ for $r \geqslant R$. Thus the density at a is

$$g(a) \geqslant \frac{1-\varepsilon/5}{\varepsilon^{-d}\Gamma(d+1)} \int_R^\infty \frac{e^{-\varepsilon t} t^d}{\text{Vol}(tK)} \mathrm{d}t \geqslant \frac{\exp(-\varepsilon/2) e^{-\varepsilon R}}{\Gamma(d+1)\text{Vol}(\varepsilon^{-1}K)} = \exp(-\varepsilon/2) f(a).$$

Similarly, using the fact that $a \notin r\widehat{K}$ for $r < R/(1+\beta)$. we have

$$g(a) \leqslant \frac{\exp(\varepsilon/5)}{\varepsilon^{-d}\Gamma(d+1)} \int_{R/(1+\beta)}^\infty \frac{e^{-\varepsilon t} t^d}{\text{Vol}(tK)} \mathrm{d}t \leqslant \frac{\exp(\varepsilon/5) e^{-\varepsilon R/(1+\beta)}}{\Gamma(d+1)\text{Vol}(\varepsilon^{-1}K)}$$

Since we chose $\beta \leqslant 1/10r$ it follows that $e^{-\varepsilon R/(1+\beta)} \leqslant e^{-\varepsilon R + \varepsilon/5}$.

We conclude that $g(a)$ is pointwise within a $\exp(\pm\varepsilon/2)$ factor of the ideal density that gives $\varepsilon/10$-differential privacy by our choice of parameters. Hence, our mechanism satisfies ε-differential privacy. Further since $K \subseteq \widehat{K} \subseteq 2K$, it follows that our mechanism satisfies the stated error bound. Finally, the bound on the moments of the Gamma distribution from Fact 2.10 implies that the expected running time of this algorithm is polynomial in N, d, ε^{-1}. □

8.1 An efficient implementation of NiKM

Theorem 8.3 extends to our mechanism for the non-isotropic case.

Theorem 8.5. *There is a mechanism M' with runtime polynomial in N, d and ε^{-1} such that*

1. *M' is ε-differentially private,*

2. $\text{err}(M', F) = O(\text{err}(\mathit{NiKM}(F, d, \varepsilon), F))$.

Proof. To implement $\text{NiKM}(F, d, \varepsilon)$ efficiently, we additionally need to compute the subspaces U and V to project onto (Step 2 of the algorithm). Note that these subspaces themselves depend only on the query F and not on the database x. Thus these can be published and the mechanism maintains its privacy for an arbitrary choice of subspaces U and V. The choice of U, V in Section 7 depended on the covariance matrix M, which we do not know how to compute exactly. We next describe a method to choose U and V that is efficient such that the resulting mechanism has essentially the same error. The sampling from K can then be replaced by approximate sampling as in the previous subsection, resulting in a polynomial-time differentially private mechanism with small error.

Without loss of generality, K has the property that $B_2^d \subseteq K \subseteq dB_2^d$. In this case, $x_i x_j \leqslant d^2$ so that with $O(d^2 \log d)$ (approximately uniform) samples from K, Chernoff bounds imply that the sample covariance matrix approximates the covariance matrix well in every entry. In other words, we can construct a matrix $\widetilde{M}$ such that with high probability each entry of $\widetilde{M}$ is within $neg(d)$ of the corresponding entry in M. Here and in the rest of the section, $neg(d)$ denotes a function bounded from above by d^{-C} for a large enough constant $C > 0$. The constant varies depending on the context. We also note that with high probability $\widetilde{M}$ is positive semidefinite. This uses the fact that $K \supseteq B_2^d$.

Let the eigenvalues of $\widetilde{M}$ be $\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_d$ with corresponding eigenvectors $\widetilde{u}_1, \ldots, \widetilde{u}_d$. Let $\widetilde{T} = \widetilde{M}^{-\frac{1}{2}}$, and let $\widetilde{P}$ be the projection operator onto the span of the first d' eigenvectors of $\widetilde{M}$. This defines our subspaces $\widetilde{U}$ and $\widetilde{V}$, and hence the mechanism. We next argue that Lemma 7.9 continues to hold.

First note that for any $i \geqslant d' + 1$

$$\mathop{\mathbb{E}}_{a \in K} \langle a, \widetilde{u}_i \rangle^2 = \left| \widetilde{u}_i^T M \widetilde{u}_i \right| = |\widetilde{u}_i^T \widetilde{M} \widetilde{u}_i| + |\widetilde{u}_i^T (M - \widetilde{M}) \widetilde{u}_i| = \widetilde{\sigma}_i + neg(d).$$

Thus, Equation 22 continues to hold with $\widetilde{\sigma}_{d'+1}$ replacing $\sigma_{d'+1}$.

To prove that Proposition 7.7 continues to hold (with $\widetilde{M}, \widetilde{T}, \widetilde{P}$ replacing M, T, P), we note that the only place in the proof that we used that M is in fact the covariance matrix of K is (20), when we require TK to be isotropic. We next argue that (20) holds for $\widetilde{T}K$ if $\widetilde{M}$ is a good enough approximation to M. This would imply Proposition 7.7 and hence the result.

First recall that Wedin's theorem [Wed72] states that for non-singular matrices R, $\widetilde{R}$,

$$\|R^{-1} - \widetilde{R}^{-1}\|_2 \leqslant \frac{1+\sqrt{5}}{2} \|R - \widetilde{R}\|_2 \cdot \max\left\{ \|R^{-1}\|_2^2, \|\widetilde{R}^{-1}\|_2^2 \right\}.$$

Using this for the matrices $M^{\frac{1}{2}}, \widetilde{M}^{\frac{1}{2}}$ and using standard perturbation bounds gives (see e.g. [KM08]):

$$\|\widetilde{T} - T\|_2 \leqslant O(1) \cdot \|T\|_2^2 \cdot \|\widetilde{M}^{\frac{1}{2}} - M^{\frac{1}{2}}\|_2. \tag{24}$$

Since $\|T\|_2$ is at most $poly(d)$ and the second term is $neg(d)$, we conclude that $\|\widetilde{T} - T\|_2$ is $neg(d)$. It follows that

$$TK \subseteq \widetilde{T}K + neg(d) B_2^d. \tag{25}$$

Moreover, since TK is in isotropic position, it contains a ball $\frac{1}{d} B_2^d$. **[Moritz's Note: needs an argument?]** It follows from the next lemma (applied to $A = d\widetilde{T}K$) that

$$\frac{1}{2d} B_2^d \subseteq \widetilde{T}K. \tag{26}$$

Lemma 8.6. *Let A be a convex body in $\mathbb{R}^d$ such that $B_2^d \subseteq A + rB_2^d$ for some $r < 1$. Then a dilation $(1-r)B_2^d$ is contained in A.*

Proof. Let $z \in \mathbb{R}^d$ be a unit vector. Suppose that $z' = (1-r)z \notin A$. Then by the Separating Hyperplane theorem (see, e.g., [BV04]), there is a hyperplane H separating z' from A. Thus there is a unit vector w and a scalar b such that $\langle z', w \rangle = b$ and $\langle u, w \rangle < b$ for all $u \in A$. Let $v = z' + rw$. Then by triangle inequality, $\|v\| \leqslant 1$. Moreover,

$$d(v, A) = \inf_{u \in A} \|u - v\| \geqslant \inf_{u \in A} \langle v - u, w \rangle \geqslant b + r - \sup_{u \in A} \langle u, w \rangle > r.$$

This however contradicts the assumption that that $v \in B_2^d \subseteq A + rB_2^d$. Since z was arbitrary, the lemma is proved. □

We can thus conclude

$$\begin{aligned} \left(1 - \tfrac{1}{d}\right) TK &\subseteq \left(1 - \tfrac{1}{d}\right) \widetilde{T}K + neg(d) B_2^d && \text{(using (25))} \\ &\subseteq \left(1 - \tfrac{1}{d}\right) \widetilde{T}K + neg(d) \widetilde{T}K && \text{(using (26))} \\ &\subseteq \widetilde{T}K, \end{aligned}$$

where the last containment follows from the fact that $\widetilde{T}K$ is convex and contains the origin. Thus

$$(1 - \frac{1}{d}) \widetilde{P}TK \subseteq \widetilde{P}\widetilde{T}K. \tag{27}$$

Since Corollary 3.4 still lower bounds the volume of $\widetilde{P}TK$, we conclude from (27) that

$$\mathrm{Vol}_k(\widetilde{P}\widetilde{T}K)^{1/k} \geqslant \frac{1}{e} \mathrm{Vol}_k(\widetilde{P}TK)^{1/k} \geqslant \frac{\alpha_K^{\frac{d-k}{k}}}{e},$$

where we have used the fact that $k \leqslant d$ so that $(1 - \frac{1}{d})^k \geqslant \frac{1}{e}$. For $k = d'$, $\frac{d-k}{k}$ is $\Theta(1)$ so that $\mathrm{Vol}_k(\widetilde{P}\widetilde{T}K)^{1/(d-k)} \geqslant \Omega(\alpha_K)$. Thus we have shown that up to constants, (20) holds for $\mathrm{Vol}_k(\widetilde{P}\widetilde{T}K)^{1/(d-k)}$ which completes the proof. □

9 Generalizations of our mechanism

Previously, we studied linear mappings $F\colon \mathbb{R}^N \to \mathbb{R}^d$ where $\mathbb{R}^N$ was endowed with the ℓ_1-metric. However, the K-norm mechanism is well-defined in a much more general context. The only property of K used here is its convexity. In general, let $\mathcal{D}$ be an arbitrary domain of databases with a distance function σ. Given a function $F : \mathcal{D} \to \mathbb{R}^d$, we could define $K_0 = \{(F(x) - F(x'))/dist(x, x') : x, x' \in \mathcal{D}\}$ and let K be the convex closure of K_0. Then the K-norm mechanism can be seen to be differentially private with respect to *dist*. Indeed note that that $|q(d,a) - q(d',a)| = |F(d) - a|_K - |F(d') - a|_K \leqslant |F(d) - F(d')|_K \leqslant dist(d, d')$, and thus privacy follows from the exponential mechanism.

Moreover, in cases when one does not have a good handle on K itself, one can use any convex body K' containing K.

Databases close in ℓ_2-norm For example, McSherry and Mironov [MM09] can transform their input data set so that neighboring databases map to points within Euclidean distance at most R for a suitable parameter R. Thus *dist* here is the ℓ_2 norm and for any linear query, K is an ellipsoid.

Local Sensitivity Nissim, Raskhodnikova and Smith [NRS07] define *smooth sensitivity* and show that one can design approximately differentially private mechanism that add noise proportional to the smooth sensitivity of the query. This can be significant improvement when the local sensitivity is much smaller than the global sensitivity. Notice that such queries are necessarily non-linear. We point out that one can define a local sensitivity analogue of the K-norm mechanism by considering the polytopes $K_x = \mathrm{conv}\left\{\frac{F(x')-F(x)}{\sigma(x,x')} : x' \in \mathcal{D}\right\}$ and adapting the techniques of [NRS07] accordingly.

References

[BCD+07] Boaz Barak, Kamalika Chaudhuri, Cynthia Dwork, Satyen Kale, Frank McSherry, and Kunal Talwar. Privacy, accuracy, and consistency too: a holistic solution to contingency table release. In *Proc. 26th Symposium on Principles of Database Systems (PODS)*, pages 273–282. ACM, 2007.

[BDMN05] Avrim Blum, Cynthia Dwork, Frank McSherry, and Kobbi Nissim. Practical privacy: the SuLQ framework. In *Proc. 24th Symposium on Principles of Database Systems (PODS)*, pages 128–138. ACM, 2005.

[BF88] I. Barany and Z. Furedi. Approximation of the sphere by polytopes having few vertices. *Proceedings of the American Mathematical Society*, 102(3):651–659, 1988.

[BKN10] Amos Beimel, Shiva Prasad Kasiviswanathan, and Kobbi Nissim. Bounds on the sample complexity for private learning and private data release. In *Proc. 7th TCC*, pages 437–454. Springer, 2010.

[BLR08] Avrim Blum, Katrina Ligett, and Aaron Roth. A learning theory approach to non-interactive database privacy. In *Proc. 40th Symposium on Theory of Computing (STOC)*, pages 609–618. ACM, 2008.

[BV04] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, March 2004.

[CM08] Kamalika Chaudhuri and Claire Monteleoni. Privacy-preserving logistic regression. In *Proc. 22nd Conference on Neural Information Processing Systems (NIPS)*, pages 289–296, 2008.

[De11] Anindya De. Lower bounds in differential privacy. *CoRR*, abs/1107.2183, 2011.

[DFK91] Martin E. Dyer, Alan M. Frieze, and Ravi Kannan. A random polynomial time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, 1991.

[DKM+06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *Proc. 25th EUROCRYPT*, pages 486–503. Springer, 2006.

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Proc. 3rd TCC*, pages 265–284. Springer, 2006.

[DMT07] Cynthia Dwork, Frank McSherry, and Kunal Talwar. The price of privacy and the limits of LP decoding. In *Proc. 39th Symposium on Theory of Computing (STOC)*, pages 85–94. ACM, 2007.

[DN03] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. In *Proc. 22nd PODS*, pages 202–210. ACM, 2003.

[DNR$^+$09] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N. Rothblum, and Salil P. Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proc. 41st Symposium on Theory of Computing (STOC)*, pages 381–390. ACM, 2009.

[DRV10] Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. Boosting and differential privacy. In *Proc. 51st Foundations of Computer Science (FOCS)*. IEEE, 2010.

[Dwo06] Cynthia Dwork. Differential privacy. In *Proc. 33rd ICALP*, pages 1–12. Springer, 2006.

[Dwo09] Cynthia Dwork. The differential privacy frontier (extended abstract). In *TCC*, pages 496–502. Springer, 2009.

[Dwo11] Cynthia Dwork. A firm foundation for private data analysis. *Communications of the ACM*, January 2011. Available from the author's web site.

[DY08] Cynthia Dwork and Sergey Yekhanin. New efficient attacks on statistical disclosure control mechanisms. In *Proc. 28th CRYPTO*, pages 469–480. Springer, 2008.

[FFKN09] Danny Feldman, Amos Fiat, Haim Kaplan, and Kobbi Nissim. Private coresets. In *Proc. 41st Symposium on Theory of Computing (STOC)*, pages 361–370. ACM, 2009.

[GH02] Apostolos Giannopoulos and Marianna Hartzoulaki. Random spaces generated by vertices of the cube. *Discrete and Computational Geometry*, V28(2):255–273, 2002.

[GHRU11] Anupam Gupta, Moritz Hardt, Aaron Roth, and Jon Ullman. Privately releasing conjunctions and the statistical query barrier. In *Proc. 43nd Symposium on Theory of Computing (STOC)*, pages 803–812. ACM, 2011.

[Gia03] Apostolos Giannopoulos. Notes on isotropic convex bodies. Preprint, 2003.

[GLM$^+$10] Anupam Gupta, Katrina Ligett, Frank McSherry, Aaron Roth, and Kunal Talwar. Differentially private approximation algorithms. In *Proceedings of the Twenty First Annual ACM-SIAM Symposium on Discrete Algorithms*, 2010. To appear.

[GLS94] Martin Grötschel, Laszlo Lovász, and Alexander Schrijver. *Geometric Algorithms and Combinatorial Optimization (Algorithms and Combinatorics)*. Springer, December 1994.

[GRS09] Arpita Ghosh, Tim Roughgarden, and Mukund Sundararajan. Universally utility-maximizing privacy mechanisms. In *Proc.* 41*st Symposium on Theory of Computing (STOC)*, pages 351–360. ACM, 2009.

[HR10] Moritz Hardt and Guy Rothblum. A multiplicative weights mechanism for privacy-preserving data analysis. In *Proc.* 51*st Foundations of Computer Science (FOCS)*, pages 61–70. IEEE, 2010.

[KK09] Bo'az Klartag and Gady Kozma. On the hyperplane conjecture for random convex sets. *Israel Journal of Mathematics*, 170(1):253–268, 2009.

[Kla06] Bo'az Klartag. On convex perturbations with a bounded isotropic constant. *Geometric and Functional Analysis (GAFA)*, 16(6):1274–1290, December 2006.

[KLS97] Ravi Kannan, László Lovász, and Miklós Simonovits. Random walks and an $O^*(n^5)$ volume algorithm for convex bodies. *Random Struct. Algorithms*, 11(1):1–50, 1997.

[KM08] David Kempe and Frank McSherry. A decentralized algorithm for spectral analysis. *Journal of Computer and System Sciences*, 74:70–83, 2008.

[KRSU10] Shiva Kasiviswanathan, Mark Rudelson, Adam Smith, and Jonathan Ullman. The price of privately releasing contingency tables and the spectra of random matrices with correlated rows. In *Proc.* 42*nd Symposium on Theory of Computing (STOC)*, pages 775–784. ACM, 2010.

[LPRN05] A. E. Litvak, A. Pajor, M. Rudelson, and Tomczak-Jaegermann N. Smallest singular value of random matrices and geometry of random polytopes. *Adv. Math.*, 195(2):491–523, 2005.

[MM09] Frank McSherry and Ilya Mironov. Differentially private recommender systems: building privacy into the net. In *Proc.* 15*th KDD*, pages 627–636. ACM, 2009.

[MP89] V.D. Milman and A. Pajor. Isotropic position and inertia ellipsoids and zonoids of the unit ball of a normed n-dimensional space. *Geometric Aspects of Functional Analysis*, 1376:64–104, 1989.

[MT07] Frank McSherry and Kunal Talwar. Mechanism design via differential privacy. In *Proc.* 48*th Foundations of Computer Science (FOCS)*, pages 94–103. IEEE, 2007.

[NRS07] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. Smooth sensitivity and sampling in private data analysis. In *Proc.* 39*th Symposium on Theory of Computing (STOC)*, pages 75–84. ACM, 2007.

[RR10] Aaron Roth and Tim Roughgarden. Interactive privacy via the median mechanism. In *Proc.* 42*nd Symposium on Theory of Computing (STOC)*, pages 765–774. ACM, 2010.

[RSH07] Vibhor Rastogi, Dan Suciu, and Sungho Hong. The boundary between privacy and utility in data publishing. In *VLDB '07: Proceedings of the 33rd international conference on Very large data bases*, pages 531–542. VLDB Endowment, 2007.

[Vem05] Santosh Vempala. Geometric random walks: a survey. *MSRI Volume on Combinatorial and Computational Geometry*, 52:577–616, 2005.

[Wed72] P.Å. Wedin. Perturbation bounds in connection with the singular value decomposition. *BIT*, 12:99–111, 1972.