NEW COMPUTATIONAL ASPECTS OF DISCREPANCY THEORY

BY ALEKSANDAR NIKOLOV

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Computer Science

Written under the direction of

S. Muthukrishnan

and approved by

_____

_____

_____

_____

New Brunswick, New Jersey

October, 2014

ABSTRACT OF THE DISSERTATION

New Computational Aspects of Discrepancy Theory

by Aleksandar Nikolov

Dissertation Director: S. Muthukrishnan

The main focus of this thesis work is *computational aspects of discrepancy theory*. Discrepancy theory studies how well discrete objects can approximate continuous ones. This question is ubiquitous in mathematics and computer science, and discrepancy theory has found numerous applications. In this thesis work, we (1) initiate the study of the polynomial time approximability of central discrepancy measures: we prove the first hardness of approximation results and design the first polynomial time approximation algorithms for combinatorial and hereditary discrepancy. We also (2) make progress on longstanding open problems in discrepancy theory, using insights from computer science: we give nearly tight hereditary discrepancy lower bounds for axis-aligned boxes in higher dimensions, and for homogeneous arithmetic progressions. Finally, we have (3) found new applications of discrepancy theory to (3a) fundamental questions in private data analysis and to (3b) communication complexity. In particular, we use discrepancy theory to design nearly optimal efficient algorithms for counting queries, in all parameter regimes considered in the literature. We also show that discrepancy lower bounds imply communication lower bounds for approximation problems in the one-way model. Directions for further research and connections to expander graphs, compressed sensing, and the design of approximation algorithms are outlined.

Acknowledgements

First, I would like to thank my advisor S. Muthukrishnan (Muthu) for his support and guidance. Taking his graduate course in Algorithms in my first semester is one of the reasons why I work in theory, and I can only hope to project the same infectious enthusiasm for algorithm design to my students. Working with him has taught me a lot about how to choose and approach problems. He has tirelessly worked to advance my career, and I am deeply grateful for that.

I would like to also thank my internship mentors at Microsoft's Silicon Valley research lab, Kunal Talwar and Cynthia Dwork. Their creativity and work ethic is an inspiration. They continued to be my mentors long after my internships ended, and have given me much invaluable advice.

I also thank my committee members, Swastik Kopparty, Mike Saks, and Salil Vadhan, for their guidance.

Many thanks to my other co-authors: Alantha Newman, Moses Charikar, Darakhshan Mir, Rebecca Wright, Ofer Neiman, Nadia Fawaz, Nina Taft, Jean Bolot, Li Zhang, Alex Andoni, Krzysztof Onak, Grigory Yaroslavtsev, Jiří Matoušek. I am very thankful in particular to Alantha, who taught me a lot about giving talks, technical writing, and doing research in a very early stage of my PhD.

I thank the Simons Foundation for generously funding the last two years of my PhD.

Thanks to my friends for their support through the last six years. Most of all, thanks to Alisha, who encouraged me and believed in me every step of the way. She would listen to my every complaint and celebrate every milestone with me. Her emotional support made this possible.

Thanks most of all to my parents Todor and Rositsa, whose sacrifice and support are the reason for my achievements. I love you and this work is dedicated to you.

Dedication

To Mom and Dad.

Table of Contents

Chapter 1

Introduction

Many questions in combinatorics and computer science can be phrased as questions about how well a "simple" probability measure can approximate a "complex" measure. Discrepancy theory provides useful tools to address such questions. A number of techniques in computer science, often developed independently from discrepancy theory, naturally relate to discrepancy problems: ϵ-nets, expander graphs, randomized and iterative rounding are just a few examples. Understanding these techniques within the framework of discrepancy theory provides context and a fresh viewpoint, which can lead to further progress. On the other hand, discrepancy theory itself raises interesting and under-explored computational questions.

In this thesis we initiate the study of the computational complexity of approximating combinatorial discrepancy measures. We show the first hardness of approximation results and design the first nontrivial polynomial time approximation algorithms. The geometric techniques we develop for our approximation algorithms allow us to resolve a number of important questions in discrepancy theory. They also have further applications in computer science: they allow us to characterize the necessary and sufficient noise required to answer statistical database queries while preserving individual privacy, in all parameter regimes considered in the literature.

We finish the thesis with directions for further research. We sketch connections between discrepancy theory and *communication complexity*, *expander graph constructions*, *compressed sensing*, and the design of *approximation algorithms*. We hope that investigating these connections further will prove fruitful.

We start this introductory chapter with a brief historical background on discrepancy theory, and an overview of some applications of discrepancy to computer science. Then

we introduce basic measures of discrepancy and provide a more detailed overview of the major results of the thesis.

1.1 Historical Background

Discrepancy theory has its origins in number theory and the theory of uniformity of distribution. Central objects of study in the latter area are uniform sequences, i.e. sequences of bounded real numbers that "hit" every interval of the same length equally often in the limit. An early and fundamental result is Weyl's criterion of uniformity [148], which can be used to show, for example, that the sequence $(i\alpha \bmod 1)_{i=1}^{\infty}$, which is fundamental in Diophantine approximation, is uniform in $[0, 1)$ for any irrational α.

The 1930s saw the emergence of a line of work inquiring into the necessary *irregularity* of discrete distributions. The uniformity of a sequence shows that the sequence in a sense converges to a uniform distribution. But what if we are interested in more precise information about the speed of convergence or in characterizing the most uniform sequence? The modern formulation of discrepancy grew out of such considerations in the work of van der Corput [143, 144], van Aardenne-Ehrenfest [141, 142], and Roth [126]. The latter paper gave an influential geometric reformulation of the question of quantifying the discrepancy of a sequence. It turns out that this question is equivalent to the problem of determining the smallest absolute deviation of a discrete counting measure supported on n points from the Lebesgue measure in the plane, where deviation is measured with respect to axis aligned rectangles. This reformulation naturally suggests investigating discrepancy with respect to other shapes and essentially started the field of geometric discrepancy theory.

Combinatorial discrepancy also has its origins in number theory, particularly in the study of irregularities with respect to long arithmetic progressions. Recall that the van der Waerden theorem implies that for any k there exists an n such that any bi-chromatic coloring of the integers $\{1, \ldots, n\}$ contains a monochromatic arithmetic progression of length at least k. This fundamental result in Ramsey theory is an example of extreme

discrepancy: it shows that for any two-coloring of a large enough set, there exist arithmetic progressions that are extremely imbalanced. But the value of n with respect to k in van der Waerden's theorem is enormous, and a natural question is how well one can simultaneously balance *long* arithmetic progressions on $\{1, \ldots, n\}$. A beautiful result of Roth from 1964 shows that no matter how we color the positive integers between 1 and n red and blue, some arithmetic progression will have $\Omega(n^{1/4})$ integers of one color in excess of the other [127]. This is a classical problem in combinatorial discrepancy theory, which is generally concerned with simultaneously balancing a collection of sets with respect to a bichromatic coloring of their union.

The modern definition of combinatorial discrepancy and its connection to classical discrepancy (i.e. uniformity of distribution) are due to Beck [19]. Informally, Beck observed that combinatorial discrepancy is equivalent to the discrepancy question of approximating a given counting measure by a counting measure with half the support of the given one. This intuition can be formalized into transference theorems between the two notions of discrepancy.

In a striking result, Beck showed that Roth's lower bound on the discrepancy of arithmetic progressions is nearly tight [20]. Beck's paper introduced the partial coloring method, which still remains one of the most powerful tools for proving discrepancy upper bounds. In another celebrated result (known as the Six Standard Deviations Suffice theorem), Spencer showed a tight upper bound of $O(\sqrt{n})$ on the discrepancy of $O(n)$ subsets of a universe of size n, by refining the partial coloring method of Beck [135]. Both Beck's result on arithmetic progressions and Spencer's result show that a careful coloring strategy can achieve much better discrepancy bounds than a simple random coloring.

Geometric vector balancing questions are closely related to combinatorial discrepancy theory. Typically, a vector balancing problem asks to give signs to a sequence of vectors from a normed vector space, so that the signed sum of the vectors is as small as possible in a prescribed norm. Equivalently, the problem is to partition a set of vectors into two sets, so that the sums over each are as close as possible. Such questions were considered by Dvoretzky, Barany and Grinberg [16], Giannopoulos [67]

Banaszczyk [10, 9], among others. Vector balancing problems are naturally related to notions of matrix discrepancy [93]: intuitively, the discrepancy of a matrix measures the minimum possible imbalance between 2-partitions of its columns. Matrix discrepancy generalizes total unimodularity, and like it, it is related to problems of approximating real valued vectors with integral vectors [93].

1.2 Connections with Computer Science

The applications of discrepancy theory to computer science are numerous and many of them are beautifully surveyed in the monograph of Chazelle [41]. Here we give just a few examples, with no claims to being exhaustive.

The theory of uniformity of distribution and related discrepancy theory questions are intimately connected to quasi-Monte Carlo methods in numerical analysis. A survey of this topic can be found in the monograph by Niederreiter [114], who covers applications to *numerical integration* (via the Koksma-Hlawka inequality and similar results), *optimization*, and *pseudorandom number generation*. Glasserman [69] gives applications of quasi-Monte Carlo methods and low discrepancy constructions in *financial engineering*. Shirley [134] gives applications to *computer graphics*.

Discrepancy theory has many important applications in *derandomization*. Particularly successful has been the use of ϵ-approximations and ϵ-nets in derandomizing algorithms in *computational geometry*. There has also been considerable interplay between techniques for deterministically constructing low discrepancy colorings and low discrepancy sets, and range searching problems. These connections are surveyed by Matoušek [100] and Chazelle [41].

Alon and Mansour [5] gave a *fast deterministic interpolation* algorithm for multivariate polynomials which relies on a construction of a low discrepancy set with respect to exponential sums. The results of Alon and Mansour are related to the foundational work of Naor and Naor on *ϵ-biased spaces* [111], i.e. low-discrepancy sets with respect to Fourier characters. ϵ-biased spaces have numerous applications, since they allow the

construction of very small sample spaces, on which the uniform distribution approximates a k-wise independent distribution. For example, ϵ-biased spaces have been used in work on work on property testing [71], color coding in parametrized complexity [7], and the construction of min-wise independent permutations [32], themselves useful in information retrieval and streaming algorithms.

Expander graphs are another classical derandomization tool with deep links to discrepancy theory. Expanders can be characterized as graphs which closely approximate a random graph with respect to cuts. This characterization is captured in the *Expander Mixing Lemma*, and its converse [25]. For general information on expander graphs see the book of Chung [44]. Linial, and Wigderson [79], and Vadhan [140] survey applications to derandomization and metric embeddings.

Some deterministic algorithms can be interpreted as derandomizations via discrepancy techniques, even though this connection may not be apparent at first. Chazelle [41] discusses such a view of the famous linear time deterministic median algorithm [27]. This interpretation of median finding inspires Chazelle's deterministic near-linear time minimum spanning tree algorithm [43] (the discrepancy theory view of the algorithm can also be found in [41]).

The above examples show that discrepancy is a useful tool for the design of efficient (deterministic) *algorithms*. It also turns out that both lower bounds on discrepancy and constructions of low-discrepancy sets are useful in understanding the *limitations of models of computation*. Chazelle [42] used lower bounds from combinatorial discrepancy theory to prove lower bounds on the size of *linear circuits* with bounded coefficients for *geometric range searching* problems. More recently, Larsen [89] used discrepancy lower bounds to give update time vs. query time trade-offs for *dynamic range searching* with bounded coefficients in the group model. Wei and Yi [147] gave lower bounds on the *space complexity* of approximate range counting data structures. The use of discrepancy for giving lower bounds on randomized communication complexity is classical (see the book [88]). Interestingly, in communication complexity, high complexity is certified by low discrepancy, in contrast to all other examples so far.

A recent line of work in combinatorial optimization connects discrepancy theory

and the design of approximation algorithms for hard combinatorial problems. These applications have a different flavor from the applications to numerical analysis, derandomization, and complexity theory. Recall that the Ghouila-Houri characterization of totally unimodular (TU) matrices [66] shows that a matrix is TU if and only if any subset of its columns can be almost perfectly balanced. This is a low-discrepancy property. TU matrices have considerable significance in combinatorial optimization, since any integer linear program whose constraints can be encoded by a TU matrix can be solved exactly using generic linear programming algorithms. Thus, extremely low discrepancy matrices allow for rounding linear programming solutions without sacrificing the quality of the solution. This is another example of how low-discrepancy objects provide a bridge between the continuous and the discrete. It turns out that a similar property holds in more generality: a solution of a linear program can be rounded to an integer solution, without increasing the cost or violating the constraints by much more than an appropriate discrepancy value associated with the constraint matrix [93]. This connection was recently exploited by Rothvoß [128] to give an improved approximation algorithm for the bin-packing problem. His work circumvents an earlier negative result by Newman, Neiman, and the author [113], also proved via discrepancy.

Despite the numerous applications of discrepancy theory to computer science, until recently relatively little was known about many central computational questions about discrepancy itself. Many non-trivial discrepancy upper bounds were first proved using non-constructive methods: the proofs only showed existence, but did not suggest an efficient algorithm to find a low-discrepancy set or coloring. Some exceptions are the work of Beck and Fiala [18], which is an early example of iterative rounding, and the work of Bohus [28] on the discrepancy of permutations (which uses similar methods). Moreover, simple randomized constructions are efficient, and also can usually be derandomized using standard techniques.

The situation was changed by a breakthrough result of Bansal, who gave a constructive version of Spencer's Six Standard Deviations Suffice theorem, based on semidefinite programming and randomized rounding via discretized Brownian motion [11]. Bansal's work still relied on the original partial coloring lemma used by Spencer. Lovett and

Meka gave a new constructive proof of Spencer' partial coloring lemma, with improved parameters [95]. An exceptionally simple algorithm which makes a vector balancing result of Giannopolous constructive, and essentially subsumes all of the above algorithmic results, was given recently by Rothvoß [129]. These new algorithms make a number of nontrivial discrepancy constructions efficient, and therefore available as algorithmic tools. Rothvoß's work on the bin-packing problem, for example, relies on Lovett and Meka's algorithm. The recent progress suggests that we can expect more algorithmic applications of deep results in discrepancy theory.

Even less was known until recently about the inherent computational complexity of computing measures of discrepancy itself. The only result we are aware of prior to the work presented in this thesis is Lovász's proof that 2-coloring a hypergraph is NP-hard [92]. As noted by Beck and Sós [21], 2-coloring a hypergraph is a special case of computing combinatorial discrepancy.

1.3 Notions of Discrepancy

In this section we introduce the basic notions of discrepancy that will be studied in the rest of the thesis.

1.3.1 Lebesgue Measure Discrepancy

Among the most well-studied discrepancy problems is the classical question of uniformity of distribution: how "uniform" can a set of n points in $[0,1]^d$ be, i.e. how well can the counting measure on a finite set of n points approximate the Lebesgue measure on $[-1,1]^d$ with respect to a family of subsets of $[0,1]^d$. Here the "complicated measure" is the continuous Lebesgue measure, and the simple measure is the counting measure on a discrete set. Formally, let $\mathcal{S}$ be a family of measurable subsets of $[0,1]^d$. Then the *Lebesgue measure discrepancy* of a point set $P \subseteq [-1,1]^d$ of size $|P| = n$ with respect to $\mathcal{S}$ is defined as

$$D(P,\mathcal{S}) \triangleq \sup_{S\in\mathcal{S}} |n\nu(S) - |S \cap P||,$$

where ν is the Lebesgue measure. More generally, we can define the discrepancy of P with respect to $\mathcal{S}$ in terms of any measure:

$$D(P, \mathcal{S}, \mu) \triangleq \sup_{S \in \mathcal{S}} |n\mu(S) - |S \cap P||.$$

We can now define the *discrepancy* of $\mathcal{S}$ with respect to the measure μ as the optimal discrepancy achievable by any n-point set P:

$$D(n, \mathcal{S}, \mu) \triangleq \inf_{P} D(P, \mathcal{S}, \mu),$$

where the infimum is taken over all subsets P of $[0,1]^d$ of size $|P| = n$. When μ is the Lebesgue measure ν, we simply use the notation $D(n, P)$.

1.3.2 Combinatorial Discrepancy

Another well-studied special case of the general question of approximating general measures by simple measure is *combinatorial discrepancy*. Combinatorial discrepancy studies how well a counting measure on a set of size n can be approximated by a counting measure on a set of size at most $n/2$. Formally, let $(\mathcal{S}, U)$ be a family of subsets of a set U of size $|U| = n$. We shall call $(\mathcal{S}, U)$ a *set system*. For $X \subseteq U$, define

$$\operatorname{disc}(X, \mathcal{S}) \triangleq \max_{S \in \mathcal{S}} ||S| - 2|X \cap S||.$$

Equivalently (and this is the more common definition), for a function $\chi \colon U \to \{-1, 1\}$, we can write

$$\operatorname{disc}(\chi, \mathcal{S}) \triangleq \max_{S \in \mathcal{S}} |\chi(S)|,$$

where $\chi(S) \triangleq \sum_{e \in S} \chi(e)$. Analogously to Lebesgue measure discrepancy, we can define the discrepancy of $\mathcal{S}$ as

$$\operatorname{disc}(\mathcal{S}) \triangleq \min_{\chi} \operatorname{disc}(\chi, \mathcal{S}),$$

where the minimum is taken over functions $\chi \colon U \to \{-1, 1\}$.

1.3.3 Hereditary Discrepancy and the Transference Lemma

It is often beneficial to consider the combinatorial discrepancy of restrictions of $\mathcal{S}$, as $\operatorname{disc}(\mathcal{S})$ turns out to be too sensitive to changes in $\mathcal{S}$. In fact, any set system

of any discrepancy can be turned into a set system of discrepancy 0 by adding new elements to the universe. The restriction $\mathcal{S}|_W$ of a family $\mathcal{S}$ to $W \subseteq U$ is the family $\mathcal{S}|_W \triangleq \{S \cap W : S \in \mathcal{S}\}$. Then we define

$$\text{herdisc}(s, \mathcal{S}) \triangleq \max_{W \subseteq U : |W| \leq s} \text{disc}(\mathcal{S}|_W),$$

and we define the *hereditary discrepancy* of $\mathcal{S}$ as

$$\text{herdisc}(\mathcal{S}) \triangleq \text{herdisc}(|U|, \mathcal{S}).$$

Note that the definition $\text{herdisc}(s, \mathcal{S})$ makes sense even when U is an infinite set and $\mathcal{S}$ is an infinite family of subsets of U, and so does $\text{herdisc}(\mathcal{S})$ with the maximum function substituted by the supremum function.

The two notions of discrepancy introduced above – Lebesgue measure discrepancy and combinatorial discrepancy – are related by the transference lemma of Beck.

Lemma 1.1 ([19])**.** *Assume that* $D(n, \mathcal{S}) = o(n)$ *and that* $\text{herdisc}(n, \mathcal{S})$ *satisfies* $\text{herdisc}(2n, \mathcal{S}) \leq (2 - \epsilon)\,\text{disc}(n, \mathcal{S})$ *for some fixed constant* ϵ*. Then* $D(n, \mathcal{S}) = O(\text{herdisc}(n, \mathcal{S}))$.

1.3.4 Discrepancy of Matrices and Rounding Algorithms

Discrepancy and hereditary discrepancy can be extended in a natural way to matrices. Let A be the the *incidence matrix* of a set system $(\mathcal{S}, U)$, i.e. a matrix $A \in \mathbb{R}^{\mathcal{S} \times U}$ so that $a_{S,e} = 1$ if $e \in S$ and $a_{S,e} = 0$ otherwise. The discrepancy and hereditary discrepancy of $\mathcal{S}$ are equal respectively to the discrepancy and hereditary discrepancy of A, defined as

$$\text{disc}(A) \triangleq \min_{x \in \{-1,1\}^U} \|Ax\|_\infty;$$

$$\text{herdisc}(s, A) \triangleq \max_{W \subseteq U : |W| \leq s} \text{disc}(A_W);$$

$$\text{herdisc}(A) \triangleq \text{herdisc}(|U|, A),$$

where A_W is the submatrix of A consisting of the columns indexed by the set W. The above definitions are valid for any matrix A, and serve as the definitions of matrix discrepancy and hereditary discrepancy.

An important motivation for the study of the hereditary discrepancy of matrices is that it is related to *rounding algorithms*, themselves useful in approximation algorithms. This is yet another example of how discrepancy related to questions of approximating continuous quantities by discrete ones.

Theorem 1.1 ([93]). *For any matrix $A \in \mathbb{R}^{m\times n}$, and any vector $c \in [-1,1]^n$, there exists a vector $x \in \{-1,1\}^n$ such that*

$$\|Ax - Ac\|_\infty \leq 2\,\mathrm{herdisc}(A).$$

This theorem related the hereditary discrepancy of A to the *linear discrepancy*, defined as the worst-case error due to rounding:

$$\mathrm{lindisc}(A) \triangleq \max_{c\in[-1,1]^n} \min_{x\in\{-1,1\}^n} \|Ax - Ac\|_\infty.$$

In these terms, the theorem shows that $\mathrm{lindisc}(A) \leq 2\,\mathrm{herdisc}(A)$.

1.3.5 L_p-Discrepancy

A relaxed, average notion of Lebesgue measure discrepancy has also been extensively studied in the literature, and its importance is comparable to the worst-case discrepancy. Given a collection of measurable subsets $\mathcal{S}$ of $[0,1]^d$, and a measure μ on $\mathcal{S}$, the L_p discrepancy of an n-points set $P \subseteq U$ is

$$D_{p,\mu}(P,\mathcal{S}) \triangleq \left(\int_{\mathcal{S}} \Big|n\nu(S) - |S\cap P|\Big|^p d\mu(S)\right)^{1/p},$$

where ν is the d-dimensional Lebesgue measure.

A similar kind of average discrepancy can also be considered in the combinatorial setting. Namely, for a set system $(\mathcal{S},U)$, we define

$$\mathrm{disc}_p(\mathcal{S}) \triangleq \min_{\chi:U\to\{-1,1\}} \left(\frac{1}{|\mathcal{S}|}\sum_{S\in\mathcal{S}} |\chi(S)|^p\right)^{1/p}.$$

More generally, for a non-negative weight function $w\,\mathcal{S}\to[0,\infty)$, not identically 0, we similarly define

$$\mathrm{disc}_{p,w}(\mathcal{S}) \triangleq \min_{\chi:U\to\{-1,1\}} \left(\frac{1}{w(\mathcal{S})}\sum_{S\in\mathcal{S}} w(S)|\chi(S)|^p\right)^{1/p}.$$

We can define $\text{herdisc}_{p,w}(s, \mathcal{S})$ as the maximum of $\text{disc}_{p,w}(\mathcal{S}|_W)$ over all restrictions of $\mathcal{S}$ to sets W of size at most s, and $\text{herdisc}_{p,w}(\mathcal{S})$ as $\text{herdisc}_{p,w}(|U|, \mathcal{S})$. When w is equal to a constant function, i.e. the weights are uniform over the sets, we use the notation $\text{disc}_p(\mathcal{S})$, $\text{herdisc}_p(s, \mathcal{S})$ and $\text{herdisc}_p(\mathcal{S})$.

These notions extend to matrices in the natural way.

1.4 Main Results of the Thesis

In this thesis we initiate the study of the computational complexity of central measures of combinatorial discrepancy. Prior to our work it was known that computing combinatorial discrepancy *exactly* is NP-hard. We prove a much stronger statement, which essentially implies that no non-trivial approximation to combinatorial discrepancy is possible, unless P = NP. In particular, from the work of Spencer [135], it is known that the discrepancy of a set system of m subsets of a universe of size n is at most $O(\sqrt{n \log(2m/n)})$. In Chapter 2 we show that it is NP-hard to distinguish between set systems that achieve this maximum attainable discrepancy bound up to constants, and set systems with the minimum possible discrepancy of 0. Our results come in two regimes: $m = O(n)$, in which it is NP-hard to distinguish between discrepancy 0 and discrepancy $\Omega(\sqrt{n})$, and $m = n^c$ for a constant $c > 1$, in which it is NP-hard to distinguish between discrepancy 0 and discrepancy $\Omega(\sqrt{n \log n})$. Our main technique is a method of composing set systems achieving asymptotically maximum discrepancy with a set system of constant size sets which either has discrepancy 0 or a constant fraction of the sets have nonzero discrepancy, and it is NP-hard to distinguish between the two cases. The technique is general enough that it allows us to prove optimal hardness results for more restricted set systems: set systems with polynomially bounded primal shatter function, which are a subset of the systems with constant VC-dimension.

We then proceed to study the hardness of computing hereditary discrepancy. Hereditary discrepancy looks superficially harder than discrepancy, since it is the maximum discrepancy over an exponentially large collection of set systems (i.e. all $2^n - 1$ non-trivial restrictions of the original set system). Moreover, we do not know whether the

problem of deciding herdisc($\mathcal{S}$) $\leq t$ is in NP; it does naturally belong to Π_2^{P} in the second level of the polynomial hierarchy. Nevertheless, the richer structure of hereditary discrepancy can make it more tractable. A classical example of this is that set systems of the lowest possible hereditary discrepancy 1 are exactly the totally unimodular set systems (by [66]) and can be recognized in polynomial time, as shown by Seymour [133]. This already contrasts with the situation with discrepancy. In Chapter 2 we show that Seymour's result is the best possible, in the sense that it is NP-hard to distinguish between hereditary discrepancy at most 2 and at least 3; this also implies that it is NP-hard to approximate hereditary discrepancy by a factor smaller than 3/2 (this was later improved to a factor 2 hardness by Austrin, Guruswami and Håstad [8]).

Then in Chapter 4, we give the first polynomial time approximation algorithm for hereditary discrepancy. Our algorithm approximates the hereditary discrepancy of any $m \times n$ matrix (and therefore any set system of m subsets of a size n universe as well) within a factor of $O(\log^{3/2} m)$. Our result shows that the robustness of hereditary discrepancy does in fact make it more tractable.

The key to our approximation algorithm is a characterization of herdisc(A) by a geometric quantity associated with the matrix A: the side length of the smallest cube which contains an ellipsoid containing all the columns of A, seen as vectors in $\mathbb{R}^m$. We call this quantity the ellipsoid-infinity norm of A and denote it $\|A\|_{E\infty}$. We show that it is equal to the optimal value of a convex minimization problem, and therefore can be computed in polynomial time using standard techniques. Because hereditary discrepancy is not an NP optimization problem, coming up with a simple fractional relaxation of it seems impossible. Indeed showing that $\|A\|_{E\infty}$ gives upper and lower bounds on herdisc(A) is challenging in both directions. We use a known bound by Banaszczyk [9] for a vector balancing problem to show that herdisc(A) $= O(\sqrt{\log m}) \cdot \|A\|_{E\infty}$. We use convex programming duality and the Restricted Invertibility Principle of Bourgain and Tzafriri [30] to show that $\|A\|_{E\infty} = O(\log m)\,\text{herdisc}(A)$. Moreover, we can find in deterministic polynomial time a submatrix of A on which hereditary discrepancy is approximately maximized. An algorithm of Bansal [11] and a vector balancing upper bound we prove in Chapter 3 can be used to find a coloring of discrepancy at most

$O(\log m) \cdot \|A\|_{E\infty}$ for any submatrix of A. Here we do not quite match the best upper bound we can prove, because Banaszczyk's bound has so far resisted attempts to find a constructive proof.

In the remainder of the thesis we show a number of applications of the ellipsoid-infinity norm to fundamental questions in discrepancy theory and private data analysis. In Chapter 6 we give new tight upper and lower bounds on the discrepancy of natural set systems. The most prominent examples from geometry are set systems induced by axis-aligned boxes in constant dimension, and set systems of subcubes of the Boolean cube (i.e. axis-aligned boxes in high dimension). We also essentially determine the hereditary discrepancy of homogeneous arithmetic progressions; doing the same for the *discrepancy* remains a challenging open problem, known as the Erdős discrepancy problem. These results use a number of favorable properties of the ellipsoid infinity norm proved in Chapter 5: e.g. that it satisfies the triangle inequality and is multiplicative with respect to tensor products. We use these properties, and the fact that the ellipsoid-infinity norm approximates hereditary discrepancy, to deduce the new upper and lower bounds from bounds for simpler set systems, which are easy to compute.

Then, in Chapter 7 we apply the ellipsoid-infinity norm and discrepancy theory to problems in private data analysis. We study the popular model of *differential privacy*, which is applicable to computing aggregate queries on databases that contain the personal information of many individuals. Differential privacy requires that the algorithm that computes the query answers behave almost identically after adding or removing individuals to the database. Differential privacy provides strong semantic guarantees: it implies, for instance, that the cost (i.e. any notion of privacy risk measured by a non-negative utility function) is not increased significantly by participating in the database. These strong guarantees come at the price of sacrificing exact query answers. Indeed, this is inevitable: Dinur and Nissim [48] have shown that answering any large enough set of a simple class of queries, called counting queries, with too much accuracy allows an adversary to recover almost the entire database, clearly violating any sensible notion of privacy. We start our investigation of differential privacy by giving a discrepancy theory viewpoint of this kind of *reconstruction attack*. We show that any algorithm that

answers a set of counting queries with error less than (an appropriate variant of) the hereditary discrepancy of an associated matrix allows a reconstruction attack. In particular, such an algorithm violates differential privacy, showing that the necessary error for answering counting queries under differential privacy is bounded from below by (a variant of) the hereditary discrepancy. On the other hand, we use the ellipsoid infinity norm to show that there exists a simple efficient differentially private algorithm that answers any set of counting queries with error not much more than the hereditary discrepancy of the associated matrix. This exhibits an interesting threshold phenomenon: on one hand, error less than the discrepancy allows a dramatic breach of privacy in the form of a reconstruction attack; on the other hand, a simple efficient algorithm has error only slightly more than the discrepancy. Our results extend the results of Hardt and Talwar [78], who considered a stronger notion of privacy than we do; more importantly, our algorithms are more efficient and hopefully will be practical. Since we characterize the necessary and sufficient error to achieve differential privacy for any set of counting queries in terms of discrepancy, we can use the results from Chapter 6 to give nearly tight upper and lower error bounds for some natural sets of queries, such as range queries and marginals. We also use the discrepancy-based reconstruction attacks to prove tight lower bounds on error in a stronger model of privacy, pan-privacy, that applies to streaming algorithms and requires that the memory state of the algorithm itself be private. Even for simple queries, the state of the algorithm can be used to answer *many* counting queries, allowing us to use a reconstruction attack.

Unfortunately, the nearly optimal algorithm from Chapter 7 may not be usable when the database size is much smaller than the number of queries. In this case, the error of our algorithm may exceed the database size, i.e. the algorithm provides only trivial error guarantees. Nevertheless, it has long been known that it is possible to reduce the error under the assumption that the database is small, and in fact non-trivial error is achievable unless the number of queries asked is exponentially large in the size of the database. Indeed, the reconstruction attacks that prove the optimality of our algorithm in Chapter 7 use databases of size at least the number of queries, and are invalid if the database is smaller. In this context, it would be desirable to have an efficient algorithm

whose error is optimal for any pair of query set and database size bound. We achieve such a guarantee, but with respect to a weaker average measure of error. Our main algorithmic tool is sparse regression: we post-process the output of an algorithm very similar to the one in Chapter 7 by performing a regression step to enforce the constraint that the query answers must be consistent with the small database size bound. Via a geometric argument we show that this post-processing step indeed reduces the error. We then show that the reduced error is in fact nearly optimal, by giving a variant of the discrepancy-based reconstruction attacks that only uses small databases. To prove optimality we also need to extend our analysis of the ellipsoid infinity norm from Chapter 4.

In Chapter 9 we show an interesting connection between our work on differential privacy and communication complexity. Since we have shown that approximate answers to a set of queries allow a reconstruction attack via discrepancy, we can use Fano's inequality to give a lower bound on the mutual information between a random database and any random variable that allows giving approximate query answers. As an application we give a new proof of Woodruff's one-way distributional communication complexity lower bound for approximating hamming distance under the uniform distribution [151].

We conclude the thesis with directions for future work. We outline some connections between problems in computer science and discrepancy theory that seem promising: in particular we outline a discrepancy theory view of expander graphs, compressed sensing, and rounding algorithms for hard combinatorial optimization problems.

1.5 Basic Notation

We use the notation $[n]$ for the set $\{1, \ldots, n\}$.

We denote matrices by capital letters, for example $A \in \mathbb{R}^{m \times n}$ is an m by n real matrix; matrix entries are denoted as lower case letters, for example for the matrix A, a_{ij} is understood to be the entry in the i-th row and j-th column. We use the standard notation $\|x\|_p$ for the ℓ_p^n norm of a vector $x \in \mathbb{R}^n$, i.e. $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$. Moreover, the ℓ_∞^n norm is, as usual, defined as $\|x\|_\infty = \max_{i=1}^n |x_i|$.

The tensor product (or *Kronecker product*) $A \otimes B$ of two matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$ is a matrix whose rows are indexed by $[m] \times [p]$, columns are indexed by $[n] \times [q]$, and the entry corresponding to pairs $(i,k),(j,l)$ is defined as the product $A_{(i,k),(j,l)} = a_{ij} b_{kl}$. This is a matrix representation of the tensor product of the linear operators represented by A and B, with the basis of the image and domain of $A \otimes B$ chosen in the natural way using the corresponding bases of the image and domain of A and B. In block representation, $A \otimes B$ is

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{pmatrix}.$$

For a matrix A, we use $\sigma_i(A)$ to denote the i-th largest singular value of A. The notation $X \succeq 0$ (respectively $X \succ 0$) means hat X is positive semidefinite (resp. positive definite). The notation $X \succeq Y$ means that $X - Y \succeq 0$, i.e.X dominates Y in the positive semidefinite (PSD) sense.

We will use the term *set system* for a pair $(\mathcal{S}, U)$, where $\mathcal{S}$ is a family of subsets of the *universe* U. When there is no ambiguity, we will often denote the set system simply by $\mathcal{S}$. The degree $\Delta_{\mathcal{S}}(e)$ of an element $e \in U$ is the maximum number of sets in $\mathcal{S}$ to which e belongs, i.e. $\Delta_{\mathcal{S}}(e) = \|\{S \in \mathcal{S} : e \in S\}|$. The maximum degree of the set system $\mathcal{S}$ is $\Delta_{\mathcal{S}} = \max_{e \in U} \Delta_{\mathcal{S}}(e)$.

For a set system $(\mathcal{S}, U)$, where $\mathcal{S} = \{S_1, \ldots, S_m\}$ and $U = \{e_1, \ldots, e_n\}$, the *incidence matrix* A of the set system is defined by

$$A_{ij} = \begin{cases} 1, & e_j \in S_i \\ 0, & \text{otherwise} \end{cases}.$$

In other words, the i-th row of A is the indicator vector of S_i.

Chapter 2

Computational Hardness

2.1 Overview

One of the corner-stone results of combinatorial discrepancy theory is Spencer's Six Standard Deviations Suffice theorem. Using the probabilistic method, it is easy to see that for any set system $\mathcal{S}$ of m subsets of a universe of size n, a random coloring χ achieves $\text{disc}(\chi, \mathcal{S}) = O(\sqrt{n \log m})$ with very high probability. Qualitatively, Spencer's theorem shows that when $m = O(n)$ we can do much better, and the $\sqrt{\log m}$ term is unnecessary. We state this result next.

Theorem 2.1 ([135])**.** *For any set system $(\mathcal{S}, U)$ with $|\mathcal{S}| = m$ and $|U| = n$,* $\text{disc}(\mathcal{S}) = O(\sqrt{n \log \frac{2m}{n}})$.

Moreover, Spencer showed that when $m = n$, the constant in the asymptotic notation is at most 6, which gives the name of the paper. Besides Spencer's proof, several others are known: an independent geometric proof by Gluskin [70], a simplification of Gluskin's proof by Giannopoulos [67], and a simplification of Spencer's proof using an entropy argument, due to Boppana (see [6, Chap. 13]). However, all of the above mentioned proofs crucially use a pigeonhole argument with an exponential number of pigeons and holes, and therefore they do not easily yield efficient algorithms. Until recently, it was not known whether any efficient algorithm can find a coloring matching Spencer's discrepancy bound, given a set system is input (and in fact Spencer conjectured otherwise). By contrast, the sub-optimal random coloring argument yields a trivial randomized algorithm, and can be derandomized using standard techniques (see e.g. [41] for details). Nevertheless, in a breakthrough paper, Bansal [11] showed that there is an efficient algorithm to find a coloring matching Spencer's bound.

Theorem 2.2 ([11, 13])**.** *There exists a deterministic polynomial time algorithm that, on input a set system* $(\mathcal{S}, U)$ *with* $|\mathcal{S}| = m$*,* $|U| = n$*, and* $m = O(n)$*, outputs a coloring* $\chi : U \to \{-1, 1\}$ *such that* $\operatorname{disc}(\chi, \mathcal{S}) = O(\sqrt{n})$.

Bansal used the key technical lemma from Spencer's proof as a black box. Following his work, Lovett and Meka [95] and Rothvoß [129] gave completely constructive arguments of Spencer's theorem.

Bansal's algorithm gives a coloring with discrepancy that matches (within constants) the worst case discrepancy for all set systems with $m = O(n)$ sets. This left open the question: Can we achieve discrepancy bounds tailored to the optimal discrepancy of the input instance instead of the worst case discrepancy over all instances? In particular, can we get better guarantees for discrepancy if the optimal discrepancy for the input instance is small? Given that the existence of an efficient algorithm for achieving worst case discrepancy was open until recently, it is not surprising that very little is known about these questions.

In this chapter, we show strong hardness results that rule out any better discrepancy guarantees for efficient algorithms. We show that from the perspective of computational efficiency, Bansal's results are tight for general set systems. Specifically, it is NP-hard to distinguish between set systems of discrepancy 0, and set systems of discrepancy $\Omega(\sqrt{n \log \frac{m}{n}})$. This means that even if the optimal solution has discrepancy zero, we cannot hope to efficiently find a coloring with discrepancy $o(\sqrt{n})$. The proof goes via composing a family of high discrepancy set systems with a family for which it is NP-hard to distinguish instances with discrepancy zero from instances in which a constant fraction of the sets have discrepancy $\Omega(1)$. The composition amplifies this zero versus $\Omega(1)$ gap.

The methods are general enough that we also obtain a similar theorem for set systems with bounded shatter function. For such set systems, we show that the upper bounds due to Matoušek [99] are tight. The proof for this latter result involves using high discrepancy set systems that have bounded shatter function in the composition, and proving that the resulting set system also has bounded shatter function. Thus, our methods suggest a general framework where we can obtain computational lower

bounds for computing the discrepancy on other restricted set systems. In particular, our composition consists of two main steps that need to be tailored to a specified type of restricted set system: (i) plug in a lower bound (i.e. high discrepancy) instance for a set system with certain specified properties, and (ii) show that the final set system maintains these specified properties. If these two steps can be carried out, the discrepancy of the lower bound instance will be translated to computational hardness of distinguishing between discrepancy zero and discrepancy equal to that of the lower bound instance.

We finish the chapter with a constant hardness of approximation result for hereditary discrepancy, which complements the approximation algorithm presented in Chapter 4. Subsequent to publication, this result has been improved by Austrin, Guruswami and Håstad [8].

2.2 Preliminaries

In the proof of our hardness result, we show that using a strong discrepancy lower bound, we can amplify a small hardness of approximation gap. The strong discrepancy lower bound is a lower bound on a slight relaxation of discrepancy in which we allow "coloring" with slightly larger numbers than ± 1. The problem that gives a small hardness of approximation gap for discrepancy is the Max-2 − 2-Set-Splitting problem, closely related to NAE-Sat. We introduce both ingredients in this section.

The relaxation of discrepancy we use is called *b-bounded discrepancy*, and is defined for a set system $(\mathcal{S}, U)$ over a universe U and a positive integer b as follows:

$$\operatorname{disc}_{\phi}^{[b]}(\mathcal{S}) \triangleq \min_{\chi} \max_{S \in \mathcal{S}} \sum_{e \in S} \chi(e),$$

where χ ranges over $\chi : U \to \{-b, \ldots, b\}$ such that $|\{e : \chi(e) \neq 0\}| \geq \phi |U|$. We observe that, similarly to discrepancy, $\operatorname{disc}^{[b]}(\mathcal{S})$ is equal to the quantity

$$\operatorname{disc}_{\phi}^{[b]}(A) \triangleq \min_{x} \|Ax\|_{\infty},$$

for $A \in \mathbb{R}^{m \times n}$ the incidence matrix of $\mathcal{S}$ and x ranging over vectors in $\{-b, \ldots, b\}^n$ such that $|\{i : x_i \neq 0\}| \geq \phi n$.

Next we define the MAX-2 − 2-SET-SPLITTING problem.

MAX-2 − 2-SET-SPLITTING

Input: A set system $(\mathcal{S}, U)$, where each set $S \in \mathcal{S}$ contains exactly 4 elements.

Output: A coloring $\chi : U \to \{-1, 1\}$ such that $|\{S \in \mathcal{S} : \sum_{e \in S} \chi(e) = 0\}|$ is maximized.

The following hardness result is due to Guruswami. The observation that we may assume that any universe element in a hard instance belongs to a bounded number of sets was made in [36].

Theorem 2.3 ([75]). *There exists a positive integer B such that for any $\phi < \frac{1}{12}$ the following holds. Given an instance $(\mathcal{S}, U)$ of* MAX-2−2-SET-SPLITTING *with maximum degree $\Delta_{\mathcal{S}} \leq B$, it is* NP*-hard to distinguish between the following two cases:*

Completeness $\mathrm{disc}(\mathcal{S}) = 0$*;*

Soundness $\forall \chi : U \to \{-1, 1\} \quad |\{S \in \mathcal{S} : \sum_{e \in S} \chi(e) \neq 0\}| \geq \phi m$.

An immediate consequence of Theorem 2.3 is that it is NP-hard to decide whether the discrepancy of a set system is 0. However, we prove a stronger result in the next section.

For some of our hardness results we need to use the well-known notion of 4-wise independent sample spaces. We give a definition and a basic existence and construction result next.

Definition 2.1. *A set $S \subseteq \{-1, 1\}^n$, $|S| = m$, is a k-wise independent sample space on $\{-1, 1\}^n$ if for each set $T \subseteq [n]$ of size at most k, and each vector $b \in \{-1, 1\}^T$,*

$$|\{x \in S : x_i = b_i \ \forall i \in T\}| = m2^{-k}.$$

For any constant k, k-wise independent sample spaces can be constructed in deterministic polynomial time. The following result is due to Alon, Babai, and Itai [3], and an exposition is also given in [6].

Lemma 2.1 ([3]). *For any natural number k and any $n = 2^t - 1$, there exists k-wise independent sample space S on $\{-1, 1\}^n$ of size $2(n+1)^{\lfloor k/2 \rfloor}$. S can be constructed in deterministic polynomial time in n^k.*

2.3 Hardness for Arbitrary Set Systems

To carry out our NP-hardness reduction, we need to be able to construct instances of set systems with high discrepancy reasonably efficiently. The next two lemmas gives such constructions for the regimes $m = O(n)$ and $m = \omega(n)$.

Lemma 2.2. *There exists a deterministic algorithm that, for any $n = 2^k$ for a positive integer k, outputs in time polynomial in n a set system $(\mathcal{S}, U)$ such that $|\mathcal{S}| = |U| = n$ and $\operatorname{disc}_\phi^{[b]}(\mathcal{S}) \geq \frac{1}{3}\sqrt{\phi n}$ for any positive integer b and any positive ϕ.*

Proof. Let H be the $n \times n$ Hadamard matrix, i.e. a matrix $H \in \{-1, 1\}^{n \times n}$ such that $H^\intercal H = nI$. Such matrices are known to exist for each $n = 2^k$ and can be constructed in time $O(n \log n)$. We let $U = [n]$ and let $\mathcal{S}$ be the system whose incidence matrix is $A = \frac{1}{2}(H + J)$, where J is the $n \times n$ all-ones matrix. Let us fix an arbitrary $x \in \{-b, \ldots, -1, 1, \ldots, b\}^n$ such that $|\{i : x_i \neq 0\}| \geq \phi n$. We have

$$\|Ax\|_\infty \geq \frac{1}{2}\|Hx\|_\infty - \frac{1}{2}|\sum_{i=1}^{n} x_i| \geq \frac{1}{2}\|Hx\|_\infty - \frac{1}{2}\|Ax\|_\infty$$

The first inequality follows from the triangle inequality, and the second from the fact that the all-ones vector is one of the rows of A. It follows that $\|Ax\|_\infty \geq \frac{1}{3}\|Hx\|_\infty$. Since H is a Hadamard matrix, and using the relationships between the ℓ_2 and ℓ_∞ norms, we have

$$\|Hx\|_\infty \geq \frac{1}{\sqrt{n}}\|Hx\|_2 = \frac{1}{\sqrt{n}}\sqrt{x^\intercal H^\intercal H x} = \|x\|_2 \geq \sqrt{\phi n}.$$

The last inequality follows because for at least ϕ fraction of the coordinates i, $|x_i| \geq 1$. Because x was arbitrary, we have $\operatorname{disc}_\phi^{[b]}(\mathcal{S}) = \operatorname{disc}_\phi^{[b]}(A) \geq \frac{1}{3}\sqrt{n}$, as desired. □

The following lemma is well-known.

Lemma 2.3. *Let $s \in \{-1, 1\}^n$ be picked uniformly from a 4-wise independent sample space. Then, for any vector $x \in \mathbb{R}^n$,*

$$\Pr[|\langle s, x\rangle| \geq \alpha\|x\|_2] > \frac{1}{3}(1 - \alpha^2)^2.$$

Proof. Let $z = |\langle s, x\rangle|^2$. We need to show that $z \geq \alpha^2\|x\|_2^2$ with probability at least $\frac{1}{3}(1-\alpha^2)^2$. Because s is sampled from a 4-wise independent sample space, $\mathbb{E}z = \|x\|_2^2$. Also, we can upper bound the second moment of z as follows:

$$\mathbb{E}z^2 = \sum_{i=1}^{n} x_i^4 + 6\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} x_i^2 x_j^2 = 3\|x\|_2^4 - 2\|x\|_4^4 < 3\|x\|2^4.$$

Then, by the Paley-Zygmund inequality,

$$\Pr[z \geq \alpha^2\|x\|_2^2] = \Pr[z \geq \alpha^2\mathbb{E}z] \geq (1-\alpha^2)^2\frac{(\mathbb{E}z)^2}{\mathbb{E}z^2} \geq \frac{1}{3}(1-\alpha^2)^2.$$

This completes the proof. □

The next lemma (i.e. the bounded discrepancy lower bound for $m = \omega(n)$) was (essentially) proved for $b = 1$ by Rabani and Shpilka, and it is easy to see that their proof can be adapted to any b. We describe their construction and the (modified) analysis here for completeness.

Lemma 2.4 ([121])**.** *There exists a deterministic algorithm that, for any positive integer k, outputs in time polynomial in $n = k(2^k - 1)$ a set system $(\mathcal{S}, U)$ such that $|U| = n$, $|\mathcal{S}| = O(n^3/\log^3 n)$, and $\operatorname{disc}_{\phi}^{[b]}(\mathcal{S}) = \Omega(\phi^{3/2}\sqrt{n\log n})$ for any positive integer b.*

Proof. Let $S \in \{-1,1\}^{m_0\times n_0}$ be a matrix whose rows form a 4-wise independent sample space, and one of the rows is the all-ones row. By Lemma we can take $m_0 \triangleq 2^{2k+1}$ and $n_0 \triangleq 2^k - 1$. Let Σ be the matrix whose rows form the set $\{-1,1\}^k$, and define the matrix $M \in \{-1,1\}^{2^k m_0 \times k n_0}$ as the tensor product $M \triangleq \Sigma \otimes S$. Moreover, for $\sigma \in \{-1,1\}^k$, define the $m_0 \times kn_0$ matrix $M^\sigma = \sigma^\intercal \otimes S$. Clearly, the rows of M are the union of the rows of M^σ over all $\sigma \in \{-1,1\}^k$.

Then, the set system $(\mathcal{S}, U)$ is the one with incidence matrix $A = \frac{1}{2}(M + J)$, where J is the $n \times n$ all-ones matrix. By construction, $|U| = k(2^k - 1)$ and $|\mathcal{S}| = 2^{3k+1}$, as required.

It remains to verify that $\operatorname{disc}_{\phi}^{[b]}(\mathcal{S}) = \operatorname{disc}_{\phi}^{[b]}(A) = \Omega(\phi^{3/2}\sqrt{n\log n})$. To this end, let us fix an arbitrary $x \in \{-b, \ldots, b\}^n$, such that $|\{i : x_i \neq 0\}| \geq \phi n$. As in the proof of Lemma 2.2, because A contains the all-ones row, it suffices to prove $\|Mx\|_\infty = \Omega(\phi^{3/2}\sqrt{n\log n})$. Let us write $x = e_1 \otimes x^1 + \ldots + e_k \otimes x^k$, where e_t is the t-th standard

basis vector in $\mathbb{R}^k$ and each x^t is a vector in $\mathbb{R}^{n_0}$. Pick s to be a uniformly random row from S, and define $X(t)$ to be the indicator random variable for the event $\{|\langle s, x^t\rangle| \geq \frac{1}{3}\|x^t\|_2\}$. By Lemma 2.3, $\mathbb{E}X(t) > \frac{2^6}{3^5}$, and, by linearity of expectation,

$$\mathbb{E}[\sum_{t=1}^{k} |\langle s, x^t\rangle|] \geq \sum_{t=1}^{k} \frac{1}{3}\|x^t\|_2 \mathbb{E}X(t) > \frac{2^6}{3^6}\sum_{t=1}^{k} \|x^t\|_2$$

Therefore, by averaging, there exists a row $s^i = S_{i*}$ of S so that $\sum_{t=1}^{k} |\langle s^i, x^t\rangle| > \frac{2^6}{3^6}\sum_{t=1}^{k} \|x^t\|_2$. Define $\sigma \in \{-1, 1\}^k$ by $\sigma_t = \text{sign}(\langle s^i, x^t\rangle)$. By construction of M, we have

$$\|Mx\|_\infty \geq \|M^\sigma x\|_\infty \geq \left|\sum_{t=1}^{k} \sigma_t \langle s^i, x^t\rangle\right| = \sum_{t=1}^{k} |\langle s^i, x^t\rangle| > \frac{2^6}{3^6}\sum_{t=1}^{k} \|x^t\|_2. \tag{2.1}$$

To complete the proof we need to lower bound $\sum_{t=1}^{k} \|x^t\|_2$. Let T be the set of block indexes t such that $|\{(t-1)n_0 + 1 \leq i \leq tn_0 : x_i \neq 0\}| \geq \phi n_0/3$. Observe that $|T| > 2\phi k/3$; indeed, otherwise the number of zero coordinates of x would be at least

$$(1 - \frac{2}{3}\phi)(1 - \frac{1}{3}\phi)n_0 k > (1 - \phi)n_0 k = (1 - \phi)n,$$

contradicting our choice of x. For any $t \in T$, $\|x^t\|_2 \geq \sqrt{\phi n_0/3}$, and

$$\sum_{t=1}^{k} \|x^t\|_2 \geq \sum_{t \in T} \|x^t\|_2 \geq \frac{2\sqrt{3}}{9}\phi^{3/2} k\sqrt{n_0} = \frac{2\sqrt{3}}{9}\phi^{3/2}\sqrt{nk}.$$

Since the choice of x was arbitrary, together with (2.1) this completes the proof. □

The following question, which asks for an improvement of the above construction, appears to be open.

Question 1. *Is there a deterministic polynomial time algorithm which, for infinitely many n and any constant c, constructs a set system $(\mathcal{S}, [n])$ such that $|\mathcal{S}| = O(n^{1+c})$ and* $\text{disc}(\mathcal{S}) = \Omega(n \log n)$. *Note that such set systems exist, by a randomized construction, and Lemma 2.4 can be modified to give a construction of set systems of size $|\mathcal{S}| = O(n^{2+c})$ for any c and* $\text{disc}(\mathcal{S}) = \Omega(n \log n)$.

As a warm-up, we prove an easier hardness result for discrepancy of *matrices* with bounded integer entries. An additional trick allows to make the hard matrix binary (and therefore an incidence matrix of a set system) by blowing up the number of rows slightly.

Theorem 2.4. *There exists a constant B, such that for matrices $A \in \{0,\ldots,B\}^{O(n)\times n}$ it is* NP*-hard to distinguish between the cases* ***(1)*** $\mathrm{disc}(A) = 0$ *and* ***(2)*** $\mathrm{disc}(A) = \Omega(\sqrt{n})$*. Moreover, for matrices $A \in \{0,\ldots,B\}^{O(n^3)\times n}$, it is* NP*-hard to distinguish between the cases* ***(1)*** $\mathrm{disc}(A) = 0$ *and* ***(2)*** $\mathrm{disc}(A) = \Omega(\sqrt{n\log n})$*.*

Proof. We prove the theorem by reduction from Max-2 − 2-Set-Splitting. We give the proof of the first statement, and the second follows analogously. Let $(\mathcal{S}_0, U)$, $|U| = n$, be an instance of Max-2 − 2-Set-Splitting with maximum degree $\Delta_{\mathcal{S}_0} \le B$; we have $m_0 \triangleq |\mathcal{S}_0| \le Bn$. Moreover, we can assume that m_0 is a power of 2, by adding at most Bn new elements, each appearing in a unique singleton set. Let A_0 be the incidence matrix of the resulting set system. Let $(\mathcal{S}_1, U_1)$ be the set system output by the algorithm in Lemma 2.2 for $k = \log_2|\mathcal{S}_0|$, and let A_1 be its incidence matrix. The reduction outputs the matrix $A = A_1A_0$. It is clear that this is a polynomial time reduction, and that A has $m_0 = O(n)$ rows. Since each column of A_0 has at most B ones, the entries of A are non-negative integers bounded by B. It remains to analyze $\mathrm{disc}(A)$.

Completeness When $\mathrm{disc}(\mathcal{S}_0) = 0$, there exists an $x \in \{-1,1\}^n$ such that $A_0x = 0$, and therefore $Ax = A_1(A_0x) = 0$, and $\mathrm{disc}(A) = 0$.

Soundness If for all $\chi : U \to \{-1,1\}$, $|\{S \in \mathcal{S}_0 : \sum_{e\in S}\chi(e) \neq 0\}| \ge \phi m_0$, then for any $x \in \{-1,1\}^n$, $A_0x \in \{-4,-2,0,2,4\}^{m_0}$, and $|\{i : (A_0x)_i \neq 0\}| \ge \phi m_0$. Then, by the definition of 4-bounded discrepancy,

$$\mathrm{disc}(A) = \min_{x\in\{-1,1\}^n} \|A_1(A_0x)\|_\infty \ge \mathrm{disc}^{[4]}_\phi(A_1) = \Omega(\sqrt{n}).$$

Since, by Theorem 2.3 it is NP-hard to distinguish between the Completeness and Soundness cases for Max-2 − 2-Set-Splitting instance with maximum degree B, the first part of the theorem follows. The second part is proved analogously, by using the set system from Lemma 2.4 as $(\mathcal{S}_1, U_1)$. □

To adapt the reduction above to output a set system (or equivalently, a binary matrix), we need a simple technical lemma, stated and proved next.

Lemma 2.5. *Let $(\mathcal{S}, U)$ be a set system with maximum degree $\Delta_{\mathcal{S}}$. Then $\mathcal{S}$ can be partitioned in polynomial time into at most $\Delta_{\mathcal{S}} + 1$ parts, each of which is a set system with maximum degree 1, i.e. no two sets share an element.*

Proof. Construct a graph $G = (V, E)$, where each vertex in V is associated with one set in $\mathcal{S}$ and there is an edge between two vertices in V if the associated pair of sets have a non-empty intersection. Each vertex in V has degree at most $\Delta_{\mathcal{S}}$, and therefore G can be colored with at most $\Delta_{\mathcal{S}} + 1$ colors in polynomial time using the standard greedy algorithm. The color classes partition V into $\Delta_{\mathcal{S}} + 1$ independent sets, where each independent set is associated with a collection of pairwise disjoint sets from $\mathcal{S}$. Therefore, we can partition $\mathcal{S}$ so that each part is the union of the vertices associated to a color class of vertices in V. □

We are now ready to prove our main result.

Theorem 2.5. *Given a set system $(\mathcal{S}, U)$ with $|\mathcal{S}| = m$, $|U| = n$ and $m = O(n)$, it is* NP*-hard to distinguish between the cases **(1)** $\operatorname{disc}(\mathcal{S}) = 0$ and **(2)** $\operatorname{disc}(\mathcal{S}) = \Omega(\sqrt{n})$. Moreover, given a set system $(\mathcal{S}, U)$ with $|\mathcal{S}| = m$, $|U| = n$ and $m = O(n^3)$, it is* NP*-hard to distinguish between the cases **(1)** $\operatorname{disc}(A) = 0$ and **(2)** $\operatorname{disc}(A) = \Omega(\sqrt{n \log n})$.*

Proof. Once again we prove the first statement, and the second statement (after "Moreover") follows analogously. Again, we use a reduction from Max-2−2-Set-Splitting. Let $(\mathcal{S}_0, U)$, $|U| = n$, and A_0 be as in the proof of Theorem 2.4. Let, furthermore, $(\mathcal{S}_1, U_1)$ again be the set system output by the algorithm in Lemma 2.2 for $k = \log_2 |\mathcal{S}_0|$, and let A_1 be its incidence matrix. By Lemma 2.5, $\mathcal{S}_0$ can be partitioned into $B + 1$ parts, each part consisting of disjoint sets; let us call the parts $\mathcal{S}_0^1, \ldots, \mathcal{S}_0^{B+1}$. Let us write $A_0 = A_0^1 + \ldots + A_0^{B+1}$, where A_0^t is a matrix whose non-zero entries form an incidence matrix for $\mathcal{S}_0^t$. In other words, A_0^t is the projection of A_0 onto the rows corresponding to the sets in $\mathcal{S}_0^t$. The reduction outputs the union of the set systems $\mathcal{S}^1, \ldots, \mathcal{S}^{B+1}$, where $\mathcal{S}^t$ is the set system with incidence matrix $A^t = A_1 A_0^t$.

It is clear that this is a polynomial time reduction, and that $\mathcal{S}$ has $(B+1)m_0 = O(n)$ sets. Since each column of A_0^t for each t has at most a single 1, the entries of each A^t

are binary, and therefore each A^t is indeed an incidence matrix. It remains to analyze $\mathrm{disc}(\mathcal{S})$.

Completeness When $\mathrm{disc}(\mathcal{S}_0) = 0$, there exists an $x \in \{-1,1\}^n$ such that $A_0^t x = 0$ for all t, and therefore $A^t x = A_1(A_0^t x) = 0$ for all t. Since $\mathrm{disc}(\mathcal{S}) = \max_{t=1}^{B+1} \mathrm{disc}(\mathcal{S}^t) = \max_{t=1}^{B+1} \mathrm{disc}(A^t)$, we have $\mathrm{disc}(\mathcal{S}) = 0$.

Soundness If for all $\chi : U \to \{-1,1\}$, $|\{S \in \mathcal{S}_0 : \sum_{e\in S} \chi(e) \neq 0\}| \geq \phi m_0$, then for any $x \in \{-1,1\}^n$, $A_0 x \in \{-4,-2,0,2,4\}^{m_0}$, and $|\{i : (A_0 x)_i \neq 0\}| \geq \phi m_0$. But $A_0 x = \sum_{t=1}^{B+1} A_0^t x$, and each $A_0^t x$ is a projections of $A_0 x$ onto a coordinate subspace. Therefore, by averaging, there exists a t such that $|\{i : (A_0^t x)_i \neq 0\}| \geq \frac{1}{B+1}\phi m_0$. Then, by the definition of 4-bounded discrepancy,

$$\mathrm{disc}(A^t) = \min_{x\in\{-1,1\}^n} \|A_1(A_0 x)\|_\infty \geq \mathrm{disc}^{[4]}_{\phi/(B+1)}(A_1) = \Omega(\sqrt{n}).$$

As noted in the Completeness case, the discrepancy of $\mathcal{S}$ is at least as large as $\mathrm{disc}(\mathcal{S}^t) = \mathrm{disc}(A^t)$, and therefore, $\mathrm{disc}(\mathcal{S}) = \Omega(\sqrt{n})$ in this case.

Since, by Theorem 2.3 it is NP-hard to distinguish between the Completeness and Soundness cases for Max-$2-2$-Set-Splitting instance with maximum degree B, the first part of the theorem follows. The second part is again proved analogously, by using the set system from Lemma 2.4 as $(\mathcal{S}_1, U_1)$. □

2.4 Hardness for Set Systems with Bounded Shatter Function

For some special classes of set systems there exist bounds that improve on the guarantees of Spencer's theorem. For example, Matoušek [99] showed improved discrepancy bounds for set systems whose shatter function is polynomially bounded. Such set systems arise frequently in computational geometry and computational learning theory. Moreover, Matoušek's bounds can be made constructive using the work of Lovett and Meka [95]. In this section, we show tight inapproximability results for the discrepancy of set systems with polynomially bounded shatter function. They are proved using the same approach that was used for proving Theorem 2.5.

Let $(U, \mathcal{S})$ be a set system on $n = |U|$ elements and $m = |\mathcal{S}|$ sets. Given $W \subseteq U$, recall that the restriction of $\mathcal{S}$ to W is $\mathcal{S}|_W = \{S \cap W : S \in \mathcal{S}\}$.

Definition 2.2. *The* primal shatter function $\pi_{\mathcal{S}}(s)$ *of* $\mathcal{S}$ *evaluated at* s *is equal to the maximum number of distinct sets in any restriction* $\mathcal{S}|_W$ *to a set* W *of size* $|W| = s$.

Matoušek [105] proved that for set systems $(U, \mathcal{S})$ such that $\pi_{\mathcal{S}}(s) = O(s^d)$, herdisc$(\mathcal{S}) = O(n^{1/2-1/2d})$. The proof relies on the entropy lemma; since Lovett and Meka [95] gave a constructive version of the lemma, Matoušek's bound can be proved constructively as well. We show that this is essentially best possible.

Theorem 2.6. *Given a set system* $(U, \mathcal{S})$*, with* $|U| = n$ *and* $\pi_{\mathcal{S}}(s) = O(s^d)$*, it is NP-hard to distinguish between the cases* ***(1)*** herdisc$(\mathcal{S}) = 0$*, and* ***(2)*** herdisc$(\mathcal{S}) = \Omega(n^{1/2-1/2d})$.

2.4.1 Generalizing Alexander's Bound

One of the main ingredients in the proof of Theorem 2.5 is a family of high discrepancy set systems: in the $m = O(n)$ regime this was the Hadamard set systems, which are a tight example for Spencer's theorem. Analogously, in the proof of Theorem 2.6 we use a family of high discrepancy set systems with polynomially bounded shatter function. The family consists of systems of sets defined by halfspaces. The discrepancy lower bound for such set systems was proved by Alexander [2]. We present the result as it appears in Chazelle [41]. We need to extend the original result to b-bounded discrepancy, which we do via the proof technique introduced in [40].

We first need to introduce a new definition. For a set P of points in $\mathbb{R}^d$, let $\Delta(P) = \max_{x,y \in U} ||x - y||_2$, and, similarly, $\delta(P) = \min_{x,y \in P} ||x - y||_2^2$. I.e. $\Delta(P)$ is the diameter of P and $\delta(P)$ is the distance between the closest pair of points.

Definition 2.3. *A set* P *of* n *points in* $\mathbb{R}^d$ *is* c-spread *if* $\Delta(P)/\delta(P) \leq cn^{1/d}$.

Observe that the set of vertices of a regular grid inside a d-dimensional cube is 1-spread.

The following simple fact will be useful in the proof of Theorem 2.6.

Lemma 2.6. *Let P be a c-spread set of n points in $\mathbb{R}^d$. If $W \subseteq P$ and $|W| \geq \phi n$, then W is $(c/\phi^{1/d})$-spread.*

Proof. Since $\Delta(W) \leq \Delta(P)$ and $\delta(W) \geq d_{min}(P)$, $\Delta(Y)/\delta(Y) \leq \Delta(P)/\delta(P) \leq cn^{1/d}$. By $|W| \geq \phi n$ we have $cn^{1/d} \leq \frac{c}{\phi^{1/d}}|W|^{1/d}$, and this completes the proof. □

We can now state the generalized version of Alexander's lower bound.

Lemma 2.7. *Let P be a $O(1)$-spread set of n points in $\mathbb{R}^d$, and let $\mathcal{S}$ be the set system induced by closed halfspaces on P. Then, $\mathrm{disc}_{\phi}^{[b]}(\mathcal{S}) = \Omega(n^{1/2-1/2d})$ for all constant b and ϕ.*

Proof. First we give a lower bound on $\mathrm{disc}_{1}^{[b]}(\mathcal{S})$ for every constant b, and then we deduce the lower bound for every constant ϕ via Lemma 2.6. The bound on $\mathrm{disc}_{1}^{[b]}(\mathcal{S})$ follows from a small modification of the argument in [40]. We sketch the modification, following Section 3.3. of Chazelle [41].

First, we introduce notation that closely follows Chazelle's. Let P be a well-spread point set in $\mathbb{R}^d$, and let $v = (v_1, 0, \ldots, 0)$ be a vector in $\mathbb{R}^d$, where v_1 is a small real number to be specified later. We consider a union of P with $t = \lceil d/2 \rceil + 1$ copies of itself, each translated by a multiple of v:

$$P_v = \bigcup_{j=0}^{t} (P + jv).$$

Fix an assignment $\chi : P \to \{\pm 1, \ldots, \pm b\}$. The coloring is extended to P_v as follows:

$$\chi(p + jv) = (-1)^j \binom{t}{j} \chi(p).$$

For a hyperplane h, let h^+ denote the closed halfspace above h, i.e. the halfspace bounded by h that does not contain the origin. Let $D(h)$ denote the discrepancy of $h^+ \cap P$, and let $D_v(h)$ denote the discrepancy of $h^+ \cap P_v$ with respect to the extended coloring. Consider a cube that encloses P, and pick a random hyperplane through the cube according to the measure on hyperplanes invariant under rigid motion. By averaging, $\mathbb{E}[D(h)^2] \geq \max_h D(h)^2$, where the expectation is taken over picking a random hyperplane as described above. Chazelle [41] shows that

$$\mathbb{E}[D(h)^2] = \Omega(\mathbb{E}[D_v(h)^2]).$$

The next step in the proof is to bound $\mathbb{E}[D_v(h)^2]$ from below. Define a weight function $G(p,q)$ as

$$G(p,q) \triangleq \begin{cases} \sum_{j=-t}^{t} (-1)^j \binom{2t}{t+j} |p-q+jv| & \text{if } p \neq q, \\ -\binom{2t-2}{t-1}\|v\| & \text{if } p = q. \end{cases}$$

Chazelle further proves the following facts:

$$\mathbb{E}[D_v(h)^2] = -\sum_{p,q \in X} \chi(p)\chi(q)G(p,q); \tag{2.2}$$

$$\sum_{x \neq y} |G(x,y)| = O(\|v\|^{2t} n^{1+(2t-1)/d}). \tag{2.3}$$

All the statements so far are independent of the range of the assignment function χ. Next we show how to modify the proof in order to accommodate the larger domain of assignments.

We separate the cross terms in the expression (7.2) for $\mathbb{E}[D_v(h)^2]$, and show that even if the points in P are assigned colors from $\{\pm 1, \pm 2, \ldots, \pm b\}$, the cross terms are dominated by the remaining terms. Note that for any $p, q \in X$, $|\chi(p)\chi(q)| \leq b^2$, and $\chi(p)^2 \geq 1$. Then,

$$\begin{aligned} \mathbb{E}[D_v(H)^2] &= -\sum_{p} \chi(p)^2 G(p,p) - \sum_{p \neq q} \chi(p)\chi(q)G(p,q) \\ &\geq -\sum_{p} G(p,p) - b^2 \sum_{p \neq q} |G(p,q)|. \end{aligned}$$

By the definition of $G(p,q)$, and the bound (2.3), we have

$$\mathbb{E}[D_v(h)^2] = \Omega(n\|v\| - b^2\|v\|^{2t} n^{1+(2t-1)/d}).$$

Setting $\|v\| = cn^{-1/d}$ gives $\mathbb{E}[D_v(h)^2] = \Omega((c - b^2c^{2t})n^{1-1/d})$. Choosing c small enough so that $c > b^2c^{2t}$ completes the proof of $\mathrm{disc}_1^{[b]}(\mathcal{S}) = \Omega(n^{1/2-1/2d})$.

It remains to deduce the lower bound for constant $\phi < 1$. Observe that $\mathrm{disc}_\phi^{[b]}(\mathcal{S})$ is equal to the minimum of $\mathrm{disc}_1^{[b]}(\mathcal{S}|_W)$ over all $W \subseteq P$ of size $|W| \geq \phi n$. Since every such W is $O(1)$-spread by Lemma 2.6, and $\mathcal{S}|_W$ is equal to the set system induced by closed halfspaces on W, we have $\mathrm{disc}_1^{[b]}(\mathcal{S}|_W) = \Omega(n^{1/2-1/2d})$. This completes the proof. □

It is a well known fact that a set system $(P, \mathcal{S})$ of halfspaces in $\mathbb{R}^d$ has $\pi_{\mathcal{S}}(s) = O(s^d)$ (see e.g. [105]). Thus, such set systems are a tight example for Matoušek's upper bound.

2.4.2 The Reduction

Our proof of the hardness of approximating discrepancy on set systems with polynomially bounded shatter function follows the structure of the proof of Theorem 2.5. The two key steps in the proof of Theorem 2.6 are using systems of halfspaces instead of Hadamard set systems, and showing that the shatter function of the final construction is bounded by $O(s^d)$. The following lemma is helpful in achieving this second goal.

Lemma 2.8. *Let $\mathcal{S}_0$ be a set system of pairwise disjoint sets, with incidence matrix $A_0 \in \mathbb{R}^{m_0 \times n}$. Furthermore, let $\mathcal{S}_1$ be a set system such that $\pi_{\mathcal{S}_1}(s) = O(s^d)$, and let $A_1 \in \mathbb{R}^{m \times m_0}$ be its incidence matrix. Then the set system $\mathcal{S}$ with incidence matrix $A = A_1 A_0$ has shatter function $\pi_{\mathcal{S}}(s) \leq \pi_{\mathcal{S}_1}(s) = O(s^d)$.*

Proof. Assume without loss of generality that the ground set of $\mathcal{S}_0$ is $[n]$, and fix W to be an arbitrary subset of $[n]$ of size s. Let, furthermore, X be the subset of $[m_0]$ indexing the non-zero rows of A_0. Clearly, $A_W = (A_1 A_0)_W = (A_1)_X (A_0)_{X,W}$, where $(A_0)_{X,W}$ is the restriction of A_0 to rows indexed by X and columns indexed by W. Moreover, $|X| \leq s$ because each column of A_0 has at most a single nonzero entry. Then, $(A_1)_X$ has at most $\pi_{\mathcal{S}_1}(s)$ distinct rows, and, therefore, $A_W = (A_1)_X (A_0)_{X,W}$ has at most as many distinct rows as well. □

Proof of Theorem 2.6. Let $(\mathcal{S}_0, U)$, $|U| = n$, be an instance of Max-2−2-Set-Splitting with maximum degree at most B and $m_0 \triangleq |\mathcal{S}_0| \leq Bn$ sets. Furthermore, let $(\mathcal{S}_1, P)$ be a set system induced by all closed halfspaces on a $O(1)$-spread point set P of size $|P| = m_0$. Let A_0 and A_1 be the incidence matrices respectively of $\mathcal{S}_0$ and $\mathcal{S}_1$. Using Lemma 2.5, we partition $\mathcal{S}_0$ into $B + 1$ set system $\mathcal{S}_0^1, \ldots, \mathcal{S}_0^{B+1}$, each consisting of pairwise disjoint sets. As in the proof of Theorem 2.5, we write $A_0 = A_0^1 + \ldots + A_0^{B+1}$, where A_0^t is the projection of A_0 onto the rows corresponding to sets in $\mathcal{S}_0^t$. Then the reduction outputs the union $\mathcal{S}$ of the set systems $\mathcal{S}^1, \ldots, \mathcal{S}^{B+1}$, where $\mathcal{S}^t$ is the set system with incidence matrix $A^t = A_1 A_0^t$.

The analysis of completeness and soundness is analogous to the analysis in Theorem 2.5, but substituting Lemma 2.7 for Lemma 2.2. It remains to prove that $\mathcal{S}$ has shatter function bounded as $\pi_{\mathcal{S}}(s) = O(s^d)$. From the definition of $\pi_{\mathcal{S}}(s)$ and the union bound, it is immediate that

$$\pi_{\mathcal{S}}(s) \leq \sum_{t=1}^{B+1} \pi_{\mathcal{S}^t}(s),$$

so it suffices to show that for any t, $\pi_{\mathcal{S}^t}(s) = O(s^d)$. This last bound follows from Lemma 2.8, and this completes the proof. □

2.5 Hardness of Approximating Hereditary Discrepancy

While no non-trivial approximation to discrepancy is possible (unless $\mathsf{P} = \mathsf{NP}$), we will see in Chapter 4 that hereditary discrepancy admits a polylogarithmic approximation. Here we show a complementary negative result: approximating herdisc better than a factor of 3/2 is NP-hard.

Theorem 2.7. *Given a set system $(\mathcal{S}, U)$, it is* NP*-hard to distinguish between the two cases* ***(1)*** $\text{herdisc}(\mathcal{S}) \leq 2$ *and* ***(2)*** $\text{herdisc}(\mathcal{S}) \geq \text{disc}(\mathcal{S}) \geq 3$.

Theorem 2.7 implies that it is NP-hard to decide if a set system has hereditary discrepancy 2. By contrast, there exists a polynomial time algorithm that recognizes matrices (and therefore set systems) with hereditary discrepancy 1. Matrices with hereditary discrepancy 1 are exactly the totally unimodular matrices [66], and an efficient algorithm for their recognition was given by Seymour [133].

The proof of Theorem 2.7 is a straight-forward reduction from the 2-colorability problem for 3-uniform set systems. Recall that a set system is r-uniform if all sets in it have size r. We also have the following definition.

Definition 2.4. *A set system $(\mathcal{S}, U)$, is* 2-colorable *if and only if there exists a set $T \subseteq U$ such that for all $S \in \mathcal{S}$, $S \cap T \neq \emptyset$ and $S \cap T \neq S$. The set T is called a* transversal *of $\mathcal{S}$.*

The hardness of deciding whether a 3-uniform set system is 2-colorable follows from Schaefer characterization of the hardness of binary constraint satisfaction problems.

Lemma 2.9 ([130])**.** *There exists a family of 3-uniform set systems such that deciding whether a set system in the family is 2-colorable is* NP*-complete.*

Given the lemma, the hardness reduction for hereditary discrepancy is straightforward.

Proof of Theorem 2.7. The proof is a reduction from the problem of deciding if a 3-uniform set system is 2-colorable. We show that a 2-colorable 3-uniform set system has hereditary discrepancy at most 2, while a set system that does not have a transversal has discrepancy at least 3. Then the theorem follows from Lemma 2.9

Completeness If a 3-uniform set system $(\mathcal{S}, U)$ is 2-colorable, this is witnessed by a transversal $T \subseteq U$. We define a coloring by $\chi(e) = 1$ if $e \in T$, and $\chi(e) = -1$ otherwise. This coloring witnesses $\operatorname{disc}(\mathcal{S}|_W) \leq 2$ for any $W \subseteq U$. Indeed, because T is a transversal, any $S \in \mathcal{S}$ has at most two elements given the same color by χ. This clearly holds for any subset of S as well, and, therefore, $|\chi(S \cap W)| \leq 2$.

Soundness If a 3-uniform set system $(\mathcal{S}, U)$ is not 2-colorable, then for any $\chi : U \to \{-1, 1\}$, $\operatorname{disc}(\mathcal{S}, \chi) \geq 3$. Indeed, if there exists a coloring such that $\operatorname{disc}(\mathcal{S}, \chi) \leq 2$, then $T \triangleq \{e \in U : \chi(e) = 1\}$ forms a transversal.

□

We remark that subsequent to publishing this result, Austrin, Guruswami and Håstad [8] have shown an improved hardness of $2 - \varepsilon$ for any $\varepsilon > 0$. A factor of two is *a natural barrier* for techniques used here and in that work, where in the low discrepancy case, the same coloring works for all restrictions of the set system. Giving either a constant factor approximation to hereditary discrepancy or a super-constant hardness result remains an open problem:

Question 2. *Can hereditary discrepancy be approximated within some fixed constant in polynomial time?*

Bibliographic Remarks

A preliminary version of the hardness results for approximating the discrepancy of general set systems with $O(n)$ sets, and the discrepancy of set systems with bounded shatter function exponent was published in [37]. The hardness result for general set systems of $n^{\omega(1)}$ sets appears for the first time in this thesis: I thank Swastik Kopparty for pointing me to the work of Shpilka and Rabani on explicit constructions of covering codes. The proofs in this chapter are simplified compared to those in [37], and present a more linear-algebraic view of the reduction. The reduction from the problem of 2-coloring hypergraphs to approximating hereditary discrepancy first appeared in the full version of [118].

Chapter 3

Vector Discrepancy and the Komlós Problem

3.1 Overview

Vector discrepancy is a convex relaxation of discrepancy, and an important tool in constructive discrepancy minimization. Unlike discrepancy, vector discrepancy is efficiently computable, since it is a convex minimization problem. We shall also see in subsequent chapters that vector discrepancy is key in designing efficient approximation algorithms for hereditary discrepancy, and has an interesting relationship with differential privacy.

In this chapter we lay out the background for these results. We define vector discrepancy and review an important result of Bansal that relates hereditary vector discrepancy and hereditary discrepancy. Bansal's result reduces approximating hereditary discrepancy to approximating hereditary vector discrepancy. However, to have any hope to approximate the latter, we need an upper bound. The main new result of this chapter is a solution to a vector discrepancy analogue of the Komlós problem, which will be used in Chapter 4 to give near tight upper bounds on hereditary vector discrepancy. Our upper bound on vector discrepancy uses strong duality for semi-definite programming.

3.2 Definition and Relationship with Hereditary Discrepancy

Let $(\mathcal{S}, U)$ be a set system. Vector discrepancy is defined analogously to discrepancy, but we "color" U with unit n-dimensional vectors rather than ± 1:

$$\operatorname{vecdisc}(\mathcal{S}) \triangleq \min_{\chi:U\to\mathbb{S}^{n-1}} \max_{S\in\mathcal{S}} \left\| \sum_{e\in S} \chi(e) \right\|_2,$$

where $\mathbb{S}^{n-1}$ is the unit sphere in $\mathbb{R}^n$. Hereditary vector discrepancy is also defined analogously, i.e. $\operatorname{hvdisc}(\mathcal{S}) = \max_{W\subseteq U} \operatorname{vecdisc}(\mathcal{S}|_W)$.

Like discrepancy, vector discrepancy generalizes to matrices $A \in \mathbb{R}^{m\times n}$:

$$\operatorname{vecdisc}(A) \triangleq \min_{u_1,\ldots,u_n\in\mathbb{S}^{n-1}} \max_{i=1}^{m} \Big\|\sum_{j=1}^{n} A_{ij}u_j\Big\|_2.$$

Notice that when A is the incidence matrix of $\mathcal{S}$, the definitions agree, i.e. $\operatorname{vecdisc}(\mathcal{S}) = \operatorname{vecdisc}(A)$. The corresponding notion of hereditary vector discrepancy of matrices is $\operatorname{hvdisc}(A) = \max_{J\subseteq[n]} \operatorname{hvdisc}(A_J)$. For the rest of the chapter we shall focus on vector discrepancy for matrices, since this is the more general notion.

Vector discrepancy can be equivalently defined as the optimal value of a convex program. In particular, $\operatorname{vecdisc}(A)^2$ can be written as the optimal solution to the *semidefinite program*

$$\text{Minimize } D \quad \text{s.t.} \tag{3.1}$$

$$(AXA^\intercal)_{ii} \le D \qquad \forall 1 \le i \le m \tag{3.2}$$

$$x_{jj} = 1 \qquad \forall 1 \le j \le n \tag{3.3}$$

$$X \succeq 0. \tag{3.4}$$

To see the equivalence, write the vectors $u_1, \ldots, u_n$ forming a vector coloring as the columns of the matrix U and set $X = U^\intercal U \succeq 0$. Also, by the Cholesky decomposition of positive semidefinite matrices, any $X \succeq 0$ can be written as $X = U^\intercal U$ where the columns of U are unit vectors and therefore give a vector coloring. Since semidefinite programs can be optimized in polynomial time, $\operatorname{vecdisc}(A)$ can be approximated to within an additive ϵ in time polynomial in $m, n, \log \epsilon^{-1}$ [72] (see also the book [65]).

Vector discrepancy is a relaxation of discrepancy, i.e. $\operatorname{vecdisc}(A) \le \operatorname{disc}(A)$ for all matrices A: a coloring $x \in \mathbb{R}^n$ achieving $\operatorname{disc}(A)$ induces a vector coloring $\{u_i = x_i v\}_{i=1}^n$ achieving the same value for vector discrepancy, where v is an arbitrary unit vector. An immediate corollary is that $\operatorname{hvdisc}(A) \le \operatorname{herdisc}(A)$ for all A. A partial converse is implied by the following result of Bansal.

Theorem 3.1 ([11])**.** *For any matrix $A \in \mathbb{R}^{m\times n}$, $\operatorname{disc}(A) = O(\log m)\operatorname{hvdisc}(A)$. Moreover, there exists a polynomial time randomized algorithm that computes a coloring $x \in \{-1, 1\}^n$ such that, with high probability, $\|Ax\|_\infty = O(\log m)\operatorname{hvdisc}(A)$.*

The converse is the following corollary of Theorem 3.1.

Corollary 3.2. *There exists a fixed constant C, such that for any matrix A,*

$$\operatorname{hvdisc}(A) \leq \operatorname{herdisc}(A) \leq (C \log m) \cdot \operatorname{hvdisc}(A).$$

Hereditary discrepancy is a maximum over an exponential number of NP-hard problems. Hereditary *vector* discrepancy is a maximum over an exponential number of convex optimization problems, which is intuitively more tractable. Nevertheless, it is not clear how to give non-trivial upper bounds on hereditary vector discrepancy. In this chapter we develop a tool that will allow us to give such upper bounds.

An interesting question is whether the upper bound in Corollary 3.2 can be improved to $O(\sqrt{\log m})$. This would be tight, since the power set $(2^U, U)$ has discrepancy $\lceil n/2 \rceil = \Omega(\log m)$ and vector discrepancy $\sqrt{n} = O(\sqrt{\log m})$, as witnessed by taking $X = I$ in (3.1)–(3.4).

3.3 Relationship with L_2-discrepancy

In a sense vector discrepancy is a relaxation of average discrepancy. This fact is captured by the following proposition.

Proposition 3.1. *For any matrix $A \in \mathbb{R}^{m \times n}$,*

$$\operatorname{vecdisc}(A) \leq \max_w \operatorname{disc}_{2,w}(A).$$

Proof. Let us define $\max_w \operatorname{disc}_{2,w}(A)$ as the value of a zero-sum game. The strategy set of the Max player is $[m]$, and the strategy set of the Min player is $\{-1, 1\}^n$. The pay-off for a pair of strategies (i, x) is $\left(\sum_{j=1}^n A_{ij} x_i\right)^2$. It is easy to verify that $\max_w \operatorname{disc}_{2,w}(A)$ is the value of this game. Then, by von Neumann's min-max theorem,

$$\max_w \operatorname{disc}_{2,w}(A) = \min_{\Pi} \max_i \left(\mathbb{E}_{x \sim \Pi} \big(\sum_{j=1}^n A_{ij} x_j \big)^2 \right)^{1/2},$$

where Π ranges over all probability distributions on $\{-1, 1\}^n$ and $\mathbb{E}_{x \sim \Pi}$ is the expectation operator when x is sampled from Π. The right hand side is an upper bound on $\operatorname{vecdisc}(A)$, because for each distribution Π we can define an assignment of unit vectors

χ with the same discrepancy. It is more convenient to define the positive semidefinite matrix X in the program (3.1)–(3.4). Let Π be a distribution on $\{-1,1\}^n$, and define $X_{i,j} = \mathbb{E}_{x\sim\Pi}[x_i x_j]$, i.e. the correlation matrix of Π. As a correlation matrix, X is a PSD matrix. Moreover, $X_{jj} = \mathbb{E}[x_j^2] = 1$, and

$$AXA^\intercal = A\mathbb{E}[xx^\intercal]A^\intercal = \mathbb{E}[(Ax)(Ax)^\intercal],$$

so $(AXA^\intercal)_{ii}$ is equal to $\mathbb{E}\big(\sum_{j=1}^n A_{ij}x_j\big)^2$. Therefore, X for an optimal mixed strategy Π for the Min player is a feasible solution to (3.1)–(3.4) with objective value equal to the value of the zero sum game. □

Proposition 3.1 and Corollary 3.2 together imply that

$$\max_w \operatorname{herdisc}_{2,w}(A) \le \operatorname{herdisc}(A) \le O(\log m)\cdot \max_w \operatorname{herdisc}_{2,w}(A).$$

There are several natural improvements to these bounds that remain open, and would have interesting consequences. One such improvement is replacing hereditary discrepancy and hereditary L_2-discrepancy with discrepancy and L_2-discrepancy. Another is replacing the bond $O(\log m)$ with $O(\sqrt{\log m}$. We are not aware of any counterexamples to either of these strengthenings. On the other hand, the power set example we mentioned in relation to Corollary 3.2 shows that the factor $O(\sqrt{\log m})$ would be tight. Indeed $(2^U, U)$ has discrepancy $\lceil n/2\rceil$ and L_2 discrepancy $\sqrt{n}$ for any weights $w : 2^U \to \mathbb{R}$. To show the latter claim, pick a coloring $\chi : U \to \{-1,1\}$ uniformly at random, and observe that

$$\mathbb{E}\left[\frac{1}{w(2^U)}\sum_{S\in 2^U} w(S)\chi(S)^2\right] = \frac{1}{w(2^U)}\sum_{S\in 2^U} w(S)\mathbb{E}[\chi(S)^2] \le \sqrt{n},$$

where the expectation is taken over the choice of χ. By averaging, there exists some χ which achieves L_2 discrepancy bounded above by the expectation.

3.4 Duality for Vector Discrepancy

Let us first state the strong duality theorem for general semidefinite programs (SDPs). Consider an SDP in the form:

$$\text{Minimize } \operatorname{tr}(F^\intercal X) \quad \text{s.t.} \tag{3.5}$$

$$\operatorname{tr}(A_i^\intercal X) = b_i \qquad \forall 1 \le i \le k, \tag{3.6}$$

$$\operatorname{tr}(C_j^\intercal X) \ge d_j \qquad \forall 1 \le j \le \ell, \tag{3.7}$$

$$X \succeq 0. \tag{3.8}$$

Above $A_1, \ldots, A_k$ and $C_1, \ldots, C_\ell$ are matrices with the appropriate dimension so that the matrix product is well-defined.

The dual SDP is:

$$\text{Maximize } b^\intercal y + d^\intercal z \quad \text{s.t.} \tag{3.9}$$

$$\sum_{i=1}^{k} y_i A_i + \sum_{j=1}^{\ell} z_j C_j \preceq F, \tag{3.10}$$

$$z_j \ge 0 \qquad \forall 1 \le j \le \ell. \tag{3.11}$$

The strong duality theorem of semidefinite programming identifies sufficient conditions under which the optimal values of (3.5)–(3.8) is equal to the optimal value of (3.9)–(3.11).

Theorem 3.3. *If the optimal value of* (3.5)–(3.8) *is finite, and there is a feasible* $X \succ 0$ *such that* $\operatorname{tr}(C_j^\intercal X) > d_j$ *for all* $j \in [\ell]$ *, then the optimal value of* (3.5)–(3.8) *is equal to the optimal value of* (3.9)–(3.11)*.*

For a proof see [65, Chap. 4] or [31, Sec. 5.9.1]. This is a special case of the more general duality theory for cone programming.

Theorem 3.3 easily leads to the following dual characterization of vector discrepancy. The dual given below was independently derived by Matoušek in his work on the determinant lower bound [106].

Proposition 3.2. *For any matrix $A \in \mathbb{R}^{m\times n}$:*

$$\text{vecdisc}(A)^2 = \max \text{tr}(Q) \quad s.t. \tag{3.12}$$

$$Q \preceq A^\intercal P A, \tag{3.13}$$

$$P \text{ diagonal }, P \succeq 0, \text{tr}(P) \leq 1, \tag{3.14}$$

$$Q \text{ diagonal }. \tag{3.15}$$

Proof. We put (3.1)–(3.4) in the form

$$\begin{aligned}
&\text{Minimize } \text{tr}(E_{n+1,n+1}\tilde{X}) \quad \text{s.t.} \\
&\qquad \text{tr}(E_{j,j}\tilde{X}) = 1 && \forall 1 \leq j \leq n, \\
&\qquad \text{tr}(E_{n+1,n+1}\tilde{X}) - \text{tr}(A^\intercal E_{i,i} A X) \geq 0 && \forall i \leq i \leq m, \\
&\qquad X \succeq 0.
\end{aligned}$$

Above we use $E_{i,j}$ to denote a standard basis matrix, i.e. the matrix with (i,j) entry 1, and all other entries 0. The new variable $\tilde{X}$ should be thought of as the direct sum of X from (3.1)–(3.4) and the variable t. Let r be the maximum ℓ_2 norm of any row of A. One can verify that that for an arbitrary $\epsilon > 0$,

$$\tilde{X} \triangleq \begin{pmatrix} I & 0 \\ 0 & r^2 + \epsilon \end{pmatrix}$$

is a strictly feasible positive definite solution to the above program. Moreover, the optimal value of the program is bounded in the interval $[0, r^2]$. The proposition then follows directly from Theorem 3.3 if we take P to be the diagonal matrix whose entries are the dual variables corresponding to inequality constraints, and Q to be the diagonal matrix whose entries are the dual variables corresponding to equality constraints. □

In the sequel we shall analyze a strengthening of vecdisc(A), defined by the following SDP:

$$\mu(A)^2 \triangleq \max\ t \quad \text{s.t.} \tag{3.16}$$

$$AXA^\intercal \preceq tI, \tag{3.17}$$

$$x_{jj} = 1 \qquad \forall 1 \leq j \leq n. \tag{3.18}$$

It is not hard to see that $\mathrm{vecdisc}(A) \leq \mu(A)$, since for any matrices X and X, and any $i \in [m]$, $AXA^\intercal \preceq tI$ implies $(AXA^\intercal)_{ii} = e_i^\intercal AXA^\intercal e_i \leq t$, where e_i is the i-th standard basis vector, i.e. a vector with 1 in position i and 0s everywhere else. On the other hand, $\mu(A)$ is more robust with respect to linear transformations of A, which will be a crucial for giving a near tight upper bound on $\mathrm{herdisc}(A)$.

A derivation very similar to the proof of Proposition 3.2 gives a dual characterization of $\mu(A)$.

Lemma 3.1. *For all $A \in \mathbb{R}^{m\times n}$:*

$$\mu(A)^2 = \max \mathrm{tr}(Q) \quad s.t. \tag{3.19}$$

$$Q \preceq A^\intercal P A, \tag{3.20}$$

$$P \succeq 0, \mathrm{tr}(P) \leq 1, \tag{3.21}$$

$$Q \text{ } \textit{diagonal.} \tag{3.22}$$

Proof. We can write (3.16)–(3.18) in the form

$$\begin{aligned} &\text{Minimize } t \quad \text{s.t.} \\ &\qquad tI - AXA^\intercal = Y, \\ &\qquad X, Y, t \succeq 0, \\ &\qquad x_{jj} = 1 \qquad\qquad\qquad\qquad \forall 1 \leq j \leq m. \end{aligned}$$

The constraint $tI - AXA^\intercal = Y$ can be thought of as n^2 equality constraints. This program can then be put in standard form analogously to the proof of Proposition 3.2. Setting $X = I$, $t = \|A\|_2^2 + \epsilon$, and $Y = tI - AA^\intercal$, where $\epsilon > 0$ is arbitrary and $\|A\|_2$ is the $\ell_2 \to \ell_2$ operator norm, gives a positive definite feasible solution. The optimal value is bounded in the interval $[0, \|A\|_2^2]$. The lemma then follows from Theorem 3.3 after taking Q to be the diagonal matrix whose entries are the dual variables corresponding to the constraints $x_{ii} = 1$, and P to be the $n \times n$ matrix whose entries correspond to the constraints $(tI - AXA^\intercal)_{i,j} = y_{i,j}$. □

3.5 The Komlós Problem

Recall that we denote the maximum degree of a set system $\mathcal{S}$ by $\Delta_{\mathcal{S}}$. By a classical result of Beck and Fiala [18], $\operatorname{disc}(\mathcal{S}) \leq 2\Delta_{\mathcal{S}} - 1$. Furthermore, Beck and Fiala conjectured that $\operatorname{disc}(\mathcal{S}) = O(\sqrt{\Delta_{\mathcal{S}}})$. Proving this conjecture remains an elusive open problem in discrepancy theory.

Let $\|A\|_{1\to 2}$ be the $\ell_1 \to \ell_2$ norm of the matrix A, equal to the maximum ℓ_2 norm of its columns. If A is the incidence matrix of $\mathcal{S}$, then $\|A\|_{1\to 2} = \sqrt{\Delta_{\mathcal{S}}}$. The Komlós conjecture is a strengthening of the Beck-Fiala conjecture, stating that there exists an absolute constant C such that $\operatorname{disc}(A) \leq C \cdot \|A\|_{1\to 2}$ for any matrix A. While this conjecture also remains open, a partial result due to Banaszczyk [9] shows that $\operatorname{disc}(A) \leq O(\sqrt{\log m}) \cdot \|A\|_{1\to 2}$. The Komlós conjecture belongs to a class of *vector balancing problems*, considered in generality by Bárány and Grinberg [16]. These problems ask to determine, given two norms $\|\cdot\|_K$ and $\|\cdot\|_L$, the supremum of

$$\min_{x\in\{-1,1\}^n} \Big\|\sum_{i=1}^{n} x_i v_i\Big\|_L,$$

over all n and all sequences $v_1, \ldots, v_n$ sich that $\|v_i\|_K \leq 1$ for all i. Beck and Fiala's proof in fact bounds this supremum by 2 when the sequence of vectors have ℓ_1 norm at most 1, and the goal is to bound the ℓ_∞ norm of their signed combination.

The Beck-Fiala theorem and Banaszczyk's theorem, and vector balancing upper bounds in general, give upper bounds on hereditary discrepancy without any additional effort. The reason is that the assumptions of such results hold for any restricted set system, or, respectively, submatrix. I.e. $\Delta_{\mathcal{S}'} \leq \Delta_{\mathcal{S}}$ for any $\mathcal{S}' = \mathcal{S}|_W$, and $\|A_J\|_{1\to 2} \leq \|A\|_{1\to 2}$. This makes vector balancing results useful for giving general upper bounds on hereditary discrepancy. In the Chapter 4 we will see an application of this fact to the design of an approximation algorithm for hereditary discrepancy.

Next we state the main new result of this chapter, which resolves a vector discrepancy version of the Komlós problem. Recall the strengthening of vector discrepancy $\mu(A)$, defined in (3.16)–(3.18).

Theorem 3.4. *For any $m \times n$ real matrix A,* $\operatorname{vecdisc}(A) \leq \mu(A) \leq \|A\|_{1\to 2}$.

This upper bound is the best possible, and is tight, for example, for the identity matrix I. Theorem 3.4 implies that if the efficient upper bound in Theorem 3.1 can be strengthened to $O(\sqrt{\log m})$, we would have a new and algorithmic proof of Banaszczyk's result. Banaszczyk's proof itself does not seem to yield an efficient algorithm to find a coloring matching his discrepancy upper bound.

A weaker bound of $\text{vecdisc}(A) = O(\sqrt{\log m}) \cdot \|A\|_{1\to 2}$ can be derived in a variety of ways: directly from Banaszczyk's upper bound; from the existence of constant discrepancy partial colorings for the Komlós conjecture; from Matoušek's recent upper bound [106] on vector discrepancy in terms of the determinant lower bound of Lovász, Spencer, and Vesztergombi [93]. An alternative proof giving $\text{vecdisc}(A) \leq C\|A\|_{1\to 2}$ for a constant $C > 1$ was communicated to us by Oded Regev, Raghu Meka, and Shachar Lovett. Such a bound also follows from Proposition 3.1 and a result Matoušek [101]. However, none of these proofs yield the tight constant of 1.

To prove Theorem 3.4, we need a well-known lemma, also known as the Cauchy Interlace Theorem. It follows easily from the variational characterization of eigenvalues, see e.g. [24, Sec. 3.1].

Lemma 3.2. *Let M be a symmetric real matrix with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n$. Let also $U \in \mathbb{R}^{n\times k}$ be a matrix with mutually orthogonal unit columns, and let $\mu_1 \geq \ldots \geq \mu_k$ be the top k eigenvalues of $U^\intercal MU$. Then for any $1 \leq i \leq k$, $\lambda_{n-k+i} \leq \mu_i \leq \lambda_i$.*

The following is an immediate consequence of Lemma 3.2.

Lemma 3.3. *Let $M \in \mathbb{R}^{n\times n} : M \succeq 0$ be a symmetric real matrix with eigenvalues $\lambda_1 \geq \ldots \geq \lambda_n \geq 0$. Let also $U \in \mathbb{R}^{n\times k}$ be a matrix with mutually orthogonal unit columns. Then $\det(U^\intercal MU) \leq \lambda_1 \ldots \lambda_k$.*

The final lemma we need states that if a sequence x of positive reals multiplicatively majorizes a sequence y, then the sum of the terms in y dominates the sum of the terms in x. We give two proofs: one based on Schur convexity and a self-contained elementary one.

Lemma 3.4. *Let $x_1 \geq \ldots \geq x_n > 0$ and $y_1 \geq \ldots \geq y_n > 0$ such that*

$$\forall k \leq n : x_1 \ldots x_k \geq y_1 \ldots y_k \tag{3.23}$$

Then,

$$x_1 + \ldots + x_n \geq y_1 + \ldots + y_n. \tag{3.24}$$

Proof. First we prove the lemma using the general tool of Schur convexity. Define the new sequence $a_1 \geq \ldots \geq a_n$ by $a_i = \ln x_i$, and, similarly, $b_1 \geq \ldots \geq b_n$ by $b_i = \ln y_i$. By assumption and the monotonicity of the logarithm, $a = (a_i)_{i=1}^n$ weakly majorizes $b = (b_i)_{i=1}^n$ from below, i.e.

$$\forall 1 \leq k \leq n : \sum_{i=1}^{k} a_i \geq \sum_{i=1}^{k} b_i.$$

This is written $b \prec_w a$. Consider the function $f(a) = \sum_{i=1}^n e^{a_i}$. Since it is symmetric with respect to permuting the coordinates, convex, and monotonically increasing in each coordinate, $b \prec_w a$ implies $f(b) \leq f(a)$ (see [98, Sec. 3.C]), which proves the lemma.

We now give an alternative self-contained proof using a powering trick. We will show that for all positive integers L, $(x_1 + \ldots + x_n)^L \geq \frac{1}{n!}(y_1 + \ldots + y_n)^L$. Taking L-th roots, we get that $x_1 + \ldots + x_n \geq \frac{1}{(n!)^{1/L}}(y_1 + \ldots + y_n)$. Letting $L \to \infty$ and taking limits yields the desired result.

By the multinomial theorem,

$$(x_1 + \ldots + x_n)^L = \sum_{i_1 + \ldots + i_n = L} \frac{L!}{i_1! \ldots i_n!} x_1^{i_1} \ldots x_n^{i_n}. \tag{3.25}$$

The inequalities (3.23) imply that whenever $i_1 \geq \ldots \geq i_n$, $x_1^{i_1} \ldots x_n^{i_n} \geq y_1^{i_1} \ldots y_n^{i_n}$. Therefore,

$$(x_1 + \ldots + x_n)^L \geq \sum_{\substack{i_1 \geq \ldots \geq i_n \\ i_1 + \ldots + i_n = L}} \frac{L!}{i_1! \ldots i_n!} y_1^{i_1} \ldots y_n^{i_n}. \tag{3.26}$$

Given a sequence $i_1, \ldots, i_n$, let σ be a permutation on n elements such that $i_{\sigma(1)} \geq \ldots \geq i_{\sigma(n)}$. Since $y_1 \geq \ldots \geq y_n$, we have that $y_1^{i_{\sigma(1)}} \ldots y_n^{i_{\sigma(n)}} \geq y_1^{i_1} \ldots y_n^{i_n}$. Furthermore, there are at most $n!$ distinct permutations of $i_1, \ldots, i_n$ (the bound is achieved exactly when all $i_1, \ldots, i_n$ are distinct). These observations and the multinomial theorem imply that

$$(y_1 + \ldots + y_n)^L \leq \sum_{\substack{i_1 \geq \ldots \geq i_n \\ i_1 + \ldots + i_n = L}} \frac{n!L!}{i_1! \ldots i_n!} y_1^{i_1} \ldots y_n^{i_n}. \tag{3.27}$$

Inequalities (3.26) and (3.27) together imply $(x_1 + \ldots + x_n)^L \geq \frac{1}{n!}(y_1 + \ldots + y_n)^L$ as desired. □

Proof of Theorem 3.4. We will show that for any positive semidefinite matrix $P \succeq 0$, and any diagonal matrix Q such that $Q \preceq A^\intercal PA$, we have $\mathrm{tr}(Q) \leq \|A\|_{1\to 2}^2 \mathrm{tr}(P)$. Together with Lemma 3.1, this implies the theorem.

We can assume, by homogeneity, that $\|A\|_{1\to 2} = 1$; under this assumption, we need to show that for P and Q feasible for (3.19)–(3.22), $\mathrm{tr}(Q) \leq \mathrm{tr}(P)$. Let us define $q_i = q_{ii}$, and also define $p_i = \lambda_i(P) \geq 0$ to be the i-th largest eigenvalue value of P. Let, without loss of generality, $q_1 \geq \ldots \geq q_{n'} > 0 \geq q_{n'+1} \geq \ldots \geq q_n$. Observe that $\sum_{i=1}^{n'} q_i \geq \sum_{i=1}^{n} q_i = \mathrm{tr}(Q)$. Therefore, it suffices to show

$$\mathrm{tr}(Q) \leq \sum_{i=1}^{n'} q_i \leq \sum_{i=1}^{n'} p_i \leq \sum_{i=1}^{m} p_i = \mathrm{tr}(P). \tag{3.28}$$

Denote by A_k the matrix consisting of the first k columns of A, and by Q_k the diagonal matrix with $q_1, \ldots, q_k$ on the diagonal. We first show that

$$\forall k \leq n' : \det(A_k^\intercal P A_k) \leq p_1 \ldots p_k. \tag{3.29}$$

Let $u_1, \ldots u_k$ be an orthonormal basis for the range of A_k and let U_k be the matrix $(u_1, \ldots u_k)$. Then $A_k = U_k U_k^\intercal A_k$, since $U_k U_k^\intercal$ acts as an orthogonal projection on the range of A_k, and therefore it leaves the columns of A_k invariant. Each column of the square matrix $U_k^\intercal A_k$ has norm at most $\|A\|_{1\to 2} = 1$, and, by Hadamard's inequality,

$$\det(A_k^\intercal U_k) = \det(U_k^\intercal A_k) \leq 1.$$

Therefore,

$$\forall k \leq n' : \det(A_k^\intercal P A_k) = \det(U_k^\intercal A_k)^2 \det(U_k^\intercal P U_k) \leq \det(U_k^\intercal P U_k).$$

By Lemma 3.3, we have that $\det(U_k^\intercal P U_k) \leq p_1 \ldots p_k$, which proves (3.29).

By constraint (3.20), we have that for all k and for all $u \in \mathbb{R}^k$, $u^\intercal A_k^\intercal P A_k u \geq u^\intercal Q_k u$. Then, we have that

$$\forall k \leq n' : \det(A_k^\intercal P A_k) \geq \det(Q_k) = q_1 \ldots q_k \tag{3.30}$$

Combining (3.29)) and (3.30), we have that

$$\forall k \leq n' : p_1 \dots p_k \geq q_1 \dots q_k \tag{3.31}$$

By Lemma 3.4, (3.31) implies (3.28), and completes the proof of the theorem. □

Bibliographic Remarks

A preliminary version of a weaker form of the result of the current chapter appears in [115].

Chapter 4

Approximating Hereditary Discrepancy

4.1 Overview

In Chapter 2 we showed that no non-trivial efficient approximation for discrepancy is possible, unless $\mathsf{P} = \mathsf{NP}$. In this chapter we will see that the robustness of hereditary discrepancy has a computational consequence: there exists a deterministic polynomial time algorithm that approximates the hereditary discrepancy of any given matrix within a polylogarithmic factor. The technical tools we develop to approximate hereditary discrepancy have further applications. In Chapter 7 we will use them to characterize the error of differentially private algorithms for linear queries; in Chapter 6 we give applications to discrepancy theory, most prominently a new and near-tight lower bound for the combinatorial discrepancy of axis-aligned boxes.

Recall that by Corollary 3.2, a factor $\alpha(m, n)$ approximation to hereditary vector discrepancy implies a factor $\alpha(m, n) \log m$ approximation to hereditary discrepancy. In this chapter we do give such an approximation result with $\alpha(m, n) = O(\log m)$.

Theorem 4.1. *There exists a polynomial time algorithm that approximates* $\operatorname{hvdisc}(A)$ *within a factor of* $O(\log m)$ *for any* $m \times n$ *matrix* A*. Moreover, the algorithm finds a submatrix* A_J *of* A*, such that* $\operatorname{hvdisc}(A) = O(\log m) \operatorname{vecdisc}(A_J)$.

Theorem 4.1 follows from a geometric characterization of hereditary vector discrepancy. We show that, up to a factor of $O(\log m)$, $\operatorname{hvdisc}(A)$ is equal to the smallest value of $\|E\|_\infty$ over all 0-centered ellipsoids that contain the columns of A. Here, $\|E\|_\infty$ is just the maximum ℓ_∞^m norm of all points in E, or, equivalently, the maximum width of E in the directions of the standard basis vectors $e_1, \ldots, e_m$. For a given matrix A, we denote the minimum achievable value of $\|E\|_\infty$ over ellipsoids E containing the columns

of A by $\|A\|_{E\infty}$ and call it the *ellipsoid infinity norm* of A. *A priori*, it is not clear how to relate this quantity in either direction to the hvdisc(A), as it is not a fractional "relaxation" in the traditional sense. It is in fact non-trivial to prove either of the two inequalities relating $\|A\|_{E\infty}$ to hvdisc(A).

Proving that the ellipsoid infinity norm is an upper bound on hereditary vector discrepancy relies on the upper bound for the Komlós problem proved in Theorem 3.4. We apply a linear transformation T to the containing ellipsoid E achieving $\|A\|_{E\infty}$, so that TE is the unit ball. Then Theorem 3.4 applies; because of the transformation, we need to make sure that in the transformed space the vector discrepancy is low in a set of directions different from the standard basis. This is where it is crucial that Theorem 3.4 gives an upper bound on $\mu(A)$, rather than merely on the vector discrepancy. While, on the face of things, this argument only gives an upper bounds on the vector discrepancy of A, it in fact also works for *any submatrix* as well, because if E contains all columns of A, it also contains all the columns of any submatrix of A. This simple observation is crucial to the success of our arguments.

To show that $\|A\|_{E\infty}$ also gives a lower bound on hvdisc(A), we analyze the *convex dual* of the optimization problem defining $\|A\|_{E\infty}$. We can transform dual certificates for this problem to dual certificates for vector discrepancy of some submatrix of A. The dual of the problem of minimizing $\|E\|_\infty$ over ellipsoids E containing the columns of A is a problem of maximizing the *nuclear norm* (i.e. the sum of singular values) over re-weightings of the columns and rows of A. To get dual certificates for vector discrepancy for some submatrix, we need to be able to extract a submatrix with a large least singular value from a matrix of large nuclear norm. We accomplish this using the *restricted invertibility principle* of Bourgain and Tzafriri [30]: a powerful theorem from functional analysis which states, roughly, that any matrix with many approximately equal singular values contains a large well-conditioned submatrix. Using a constructive proof of the theorem by Spielman and Srivastava [137], we can also find the well-conditioned submatrix in deterministic polynomial time; this gives us a submatrix of A on which hereditary vector discrepancy is approximately maximized.

Theorem 4.1 immediately implies a $O(\log^2 m)$ approximation of herdisc via Corollary 3.2. However, we can improve this bound to an $O(\log^{3/2} m)$ approximation:

Theorem 4.2. *There exists a polynomial time algorithm that approximates* herdisc(A) *within a factor of* $O(\log^{3/2} m)$ *for any* $m \times n$ *matrix* A*. Moreover, the algorithm finds a submatrix* A_J *of* A*, such that* $\text{herdisc}(A) \leq O(\log^{3/2} m)\,\text{vecdisc}(A_J)$.

To prove Theorem 4.2, we retain the same lower bound on hereditary discrepancy, but prove a new upper bound. Rather than bounding vector discrepancy from above in terms of $\|A\|_{E\infty}$, and then bounding discrepancy in terms of vector discrepancy, we give a direct upper bound on discrepancy in terms of $\|A\|_{E\infty}$. For this purpose, we use a general form of Banaszczyk's upper bound for the Komlós problem that we mentioned in Chapter 3. Banaszczyk's general result [9] shows that for any convex body K of large Gaussian volume, and a matrix A with columns of at most unit Euclidean norm, there exists a $x \in \{-1, 1\}^n$ such that $Ax \in CK$ for a constant C. We use this theorem analogously to the way we used Theorem 3.4: we find a linear transformation T that maps the ellipsoid E achieving $\|A\|_{E\infty}$ to the unit ball, and we specify a body K such that if some ± 1 combination of the columns of TA is in K, then the corresponding combination of the columns of A is in the infinity ball scaled by $O(\sqrt{\log m})$.

Lovász, Spencer and Vesztergombi [93] defined the following quantity, commonly called the *determinant lower bound*:

$$\text{detlb}(A) = \max_k \max_B |\det(B)|^{1/k},$$

where for each k, B ranges over $k \times k$ submatrices of A. They proved that $\text{detlb}(A)$ gives a lower bound on hereditary discrepancy.

Theorem 4.3 ([93])**.** *For any matrix* A*,* $\text{detlb}(A) \leq 2\,\text{herdisc}(A)$.

Matoušek [106] showed that this lower bound is tight up to $O(\log^{3/2} m)$. These results do not immediately yield an approximation algorithm for hereditary discrepancy, as the determinant lower bound is a maximum over exponentially many quantities and not known to be efficiently computable. We show that $\|A\|_{E\infty}$ approximates $\text{detlb}(A)$ up to a factor of $O(\log m)$ via a determinant-based version of the restricted invertibility

principle. We give an elementary self-contained proof of this version of the principle. This provides also an elementary proof (without using the restricted invertibility principle) of the lower bound in Theorem 4.2.

In Chapter 5 we give some useful properties of the ellipsoid infinity norm, and we prove that there exist examples for which our upper and lower on bounds herdisc(A) in terms of $\|A\|_{E\infty}$ are tight (such examples were discovered independently by Matoušek. The properties of $\|\cdot\|_{E\infty}$ will be later used in Chapter 6 to prove upper and lower bounds on the discrepancy of natural set systems.

4.2 Preliminaries

In this section we review the tools that will be needed in the proof of our approximation result.

4.2.1 Restricted Invertibility

For a matrix M, we denote by $\|M\|_2 = \|M\|_{2\to 2} = \sigma_{\max}(M)$ the spectral norm of M and $\|M\|_{HS} = \sqrt{\sum_i \sigma_i^2(M)} = \sqrt{\sum_{i,j} a_{i,j}^2}$ the Hilbert-Schmidt (or Frobenius) norm of M. Recall that the $\ell_1 \to \ell_2$ norm $\|M\|_{1\to 2}$ is equal to the maximum Euclidean length of the columns of the matrix $M = (a_i)_{i=1}^n$, i.e. $\|M\|_{1\to 2} = \max_{x:\|x\|_1=1} \|Mx\|_2 = \max_{i\in[n]} \|M_i\|_2$.

A matrix M trivially contains an invertible submatrix of k columns as long as $k \le \mathrm{rank}(M)$. An important result of Bourgain and Tzafriri [30] (later strengthened by Vershynin [146], and Spielman and Srivastava [137]) shows that when k is strictly less than the robust rank $\|M\|_{HS}^2/\|M\|_2^2$ of M, we can find k columns of M that form a *well-invertible* submatrix. This result is usually called the *restricted invertibility principle*. Next we state a tight algorithmic version of it, due to Spielman and Srivastava.

Theorem 4.4 ([137])**.** *Let $\epsilon \in (0,1)$, and let M be an m by n real matrix. For any integer k such that $k \le \epsilon^2 \frac{\|M\|_{HS}^2}{\|M\|_2^2}$ there exists a subset $J \subseteq [n]$ of size $|J| = k$ such that $\sigma_{\min}(M_J)^2 \ge (1-\epsilon)^2 \frac{\|M\|_{HS}^2}{n}$. Moreover, J can be computed in deterministic polynomial time.*

We will need the following *weighted* version of Theorem 4.4, which can be proved by a slight modification of the argument of Spielman and Srivastava.

Theorem 4.5. *Let $\epsilon \in (0,1)$, let M be an $m \times n$ real matrix, and let W be an $n \times n$ diagonal matrix such that $W \succeq 0$ and $\mathrm{tr}(W) = 1$. For any integer k such that $k \leq \epsilon^2 \frac{\|MW^{1/2}\|_{HS}^2}{\|MW^{1/2}\|_2^2}$ there exists a subset $J \subseteq [n]$ of size $|J| = k$ such that $\sigma_{\min}(M_J)^2 \geq (1-\epsilon)^2 \|MW^{1/2}\|_{HS}^2$. Moreover, J can be computed in deterministic polynomial time.*

For completeness, we also give a reduction from Theorem 4.5 to Theorem 4.4. The reduction is based on the following simple lemma.

Lemma 4.1. *Let $W \succeq 0$ be a diagonal matrix with rational entries, such that $\mathrm{tr}(W) = 1$. Then for any m by n matrix M, there exists a $m \times \ell$ matrix L such that $LL^\intercal = \ell MWM^\intercal$. Moreover, all columns of L are columns of M.*

Proof. Let ℓ be the least common denominator of all diagonal entries of W, i.e. $\ell W = D$ for an integral diagonal matrix D. Denote the j-th column of M by v_j, and let L be a matrix with d_{jj} copies of the j-th column of v_j. Clearly,

$$LL^\intercal = \sum_{j=1}^{n} d_{jj} v_j v_j^\intercal = MDM^\intercal = \ell AWA^\intercal.$$

Observe, finally, that the number of columns of L is equal to $\sum_{j=1}^n d_{jj} = \ell \sum_{j=1}^n w_{jj} = \ell$, since $\mathrm{tr}(W) = 1$ by assumption. □

Proof. Proof of Theorem 4.5 By introducing a tiny perturbation to W, we can make it rational while changing $\|MW^{1/2}\|_{HS}$ and $\|MW^{1/2}\|_2$ by an arbitrarily small amount. Therefore, we may assume that W is rational. Then, by Lemma 4.1, there exists a matrix L with ℓ columns all of which are columns of M, such that $LL^\intercal = \ell MWM^\intercal$. Let k be an integer such that

$$k \leq \epsilon^2 \frac{\|MW^{1/2}\|_{HS}^2}{\|MW^{1/2}\|_2^2} = \epsilon^2 \frac{\mathrm{tr}(MWM^\intercal)}{\lambda_{\max}(MWM^\intercal)} = \epsilon^2 \frac{\ell \mathrm{tr}(LL^\intercal)}{\ell \lambda_{\max}(LL^\intercal)} = \epsilon^2 \frac{\|L\|_{HS}^2}{\|L\|_2^2},$$

where $\lambda_{\max}(MWM^\intercal$ is the largest eigenvalue of $MWM^\intercal$. By Theorem 4.4, there exists a set J of size k, such that

$$\sigma_{\min}(L_J)^2 \geq (1-\epsilon)^2 \frac{\|L\|_{HS}^2}{\ell} = (1-\epsilon)^2 \frac{\mathrm{tr}(LL^\intercal)}{\ell} = (1-\epsilon)^2 \mathrm{tr}(MWM^\intercal) = (1-\epsilon)^2 \|MW^{1/2}\|_{HS}^2.$$

But since all columns of L are also columns of M, and no column in L_J can be repeated or otherwise $\sigma_{\min}(L_J) = 0$, there exists a set $K \subseteq [n]$ of size k such that $\sigma_{\min}(M_K)^2 \geq (1-\epsilon)^2 \|MW^{1/2}\|_{HS}^2$. □

We also use an elementary lemma which can be thought of as a version of the restricted invertibility principle for determinants. This result is much easier to prove. A similar argument was used in [106].

Lemma 4.2. *Let M be an $k \times n$ matrix, and let W be an $n \times n$ diagonal matrix such that $W \succeq 0$ and $\operatorname{tr}(W) = 1$. Then there exists a k-element set $J \subseteq [n]$ such that*

$$|\det(M_J)|^{1/k} \geq \sqrt{k/e} \cdot |\det(MWM^\intercal)|^{1/2k}.$$

Proof. Applying the Binet–Cauchy formula to the matrix $MW^{1/2}$ and slightly simplifying, we have

$$\det(MWM^\intercal) = \sum_J \det(M_J)^2 \prod_{j \in J} w_{jj}.$$

Now $\sum_J \prod_{j \in J} w_{jj} \leq \frac{1}{k!}\left(\sum_{j=1}^n w_{jj}\right)^k = \frac{1}{k!}$, because each term of the left-hand side appears $k!$-times on the right-hand side (and the weights w_{jj} are nonnegative and sum to 1). Therefore

$$\begin{aligned} \det(MWM^\intercal) &\leq \left(\max_J \det(M_J)^2\right) \sum_J \prod_{j \in J} w_{jj} \\ &\leq \frac{1}{k!} \max_J \det(M_J)^2. \end{aligned}$$

So there exists a k-element J with

$$|\det(M_J)|^{1/k} \geq (k!)^{1/2k} |\det(MWM^\intercal)|^{1/2k} \geq \sqrt{k/e} \cdot |\det(MWM^\intercal)|^{1/2k},$$

where the last inequality follows from the estimate $k! \geq (k/e)^k$. □

4.2.2 Geometry

We review some basic notions from convex geometry.

A *convex body* is a convex compact subset of $\mathbb{R}^m$. For a convex body $K \subseteq \mathbb{R}^m$, the *polar body* K° is defined by $K^\circ = \{y : \langle y, x\rangle \leq 1 \ \forall x \in K\}$. A basic fact about polar

bodies is that for any two convex bodies K and L, $K \subseteq L \Leftrightarrow L^\circ \subseteq K^\circ$. A related fact is that for any convex bodies K and L, $(K \cap L)^\circ = \mathrm{conv}\{K^\circ \cup L^\circ\}$. Moreover, a symmetric convex body K and its polar body are dual to each other, in the sense that $(K^\circ)^\circ = K$.

A convex body K is *(centrally) symmetric* if $-K = K$. The *Minkowski norm* $\|x\|_K$ induced by a symmetric convex body K is defined as $\|x\|_K \triangleq \min\{r \in \mathbb{R} : x \in rK\}$. The Minkowski norm induced by the polar body K° of K is the *dual norm* of $\|x\|_K$ and also has the form $\|y\|_{K^\circ} = \max_{x \in K} \langle x, y \rangle$. It follows that we can also write $\|x\|_K$ as $\|x\|_K = \max_{y \in K^\circ} \langle x, y \rangle$. For a vector y of unit Euclidean length, $\|y\|_{K^\circ}$ is the *width* of K in the direction of y, i.e. half the Euclidean distance between the two supporting hyperplanes of K orthogonal to y. For symmetric body K, we denote by $\|K\| = \max_{x \in K} \|x\|$ the radius of K under the norm $\|\cdot\|$.

Of special interest are the ℓ_p^m norms, defined for any $p \geq 1$ and any $x \in \mathbb{R}^m$ by $\|x\|_p = (\sum_{i=1}^m |x|^p)^{1/p}$. The ℓ_∞^m norm is defined for as $\|x\|_\infty = \max_{i=1}^m |x_i|$. The norms ℓ_p^m and ℓ_q^m are dual if and only if $\frac{1}{p} + \frac{1}{q} = 1$, and ℓ_1^m is dual to ℓ_∞^m. We denote the unit ball of the ℓ_p^m norm by $B_p^m = \{x : \|x\|_p \leq 1\}$. As with the unit ball of any norm, B_p^m is convex and centrally symmetric for $p \in [1, \infty]$.

An *ellipsoid* in $\mathbb{R}^m$ is the image of the ball B_2^m under an affine map. All ellipsoids we consider are symmetric, and therefore, are equal to an image FB_2^m of the ball B_2^m under a linear map F. A full dimensional ellipsoid $E = FB_2^d$ can be equivalently defined as $E = \{x : x^\intercal (FF^\intercal)^{-1} x \leq 1\}$. The polar body of a symmetric ellipsoid $E = FB_2^d$ is the ellipsoid $E^\circ = \{x : x^\intercal FF^\intercal x \leq 1\}$. It follows that for $E = FB_2^m$ and for any x, $\|x\|_E = \sqrt{x^\intercal (FF^\intercal)^{-1} x}$ and for any y, $\|y\|_{E^\circ} = \sqrt{y^\intercal (FF^\intercal) y}$.

4.2.3 Convex Duality

Here we review the theory of Lagrange duals for convex optimization problems. Assume we are given the following optimization problem:

$$\text{Minimize } f_0(x) \tag{4.1}$$

s.t.

$$\forall 1 \leq i \leq m : f_i(x) \leq 0. \tag{4.2}$$

The Lagrange dual function associated with (4.1)–(4.2) is defined as $g(y) = \inf_x f_0(x) + \sum_{i=1}^{m} y_i f_i(x)$, where the infimum is over the intersection of the domains of $f_1, \ldots, \ldots f_m$, and $y \in \mathbb{R}^m$, $y \geq 0$. Since $g(y)$ is the infimum of affine functions, it is a concave upper-semicontinuous function.

For any x which is feasible for (4.1)–(4.2), and any $y \geq 0$, $g(y) \leq f_0(x)$. This fact is known as *weak duality*. The *Lagrange dual problem* is defined as

$$\text{Maximize } g(y) \text{ s.t. } y \geq 0. \tag{4.3}$$

Strong duality holds when the optimal value of (4.3) equals the optimal value of (4.1)–(4.2). Slater's condition is a commonly used sufficient condition for strong duality. We state it next.

Theorem 4.6 (Slater's Condition)**.** *Assume $f_0, \ldots, f_m$ in the problem* (4.1)–(4.2) *are convex functions over their respective domains, and for some $k \geq 0$, $f_1, \ldots, f_k$ are affine functions. Let there be a point x in the relative interior of the domains of $f_0, \ldots, f_m$, so that $f_i(x) \leq 0$ for $1 \leq i \leq k$ and $f_j(x) < 0$ for $k+1 \leq j \leq m$. Then the minimum of* (4.1)–(4.2) *equals the maximum of* (4.3)*, and the maximum of* (4.3) *is achieved if it is finite.*

For more information on convex programming and duality, we refer the reader to the book by Boyd and Vandenberghe [31].

4.3 Ellipsoid Upper Bounds on Discrepancy

In this section we show that ellipsoids of small infinity norm provide upper bounds on both hereditary vector discrepancy and hereditary discrepancy. Giving such an upper

bound is in general challenging because it must hold for all submatrices simultaneously. The proofs use Theorem 3.4, and Banaszczyk's general vector balancing result, stated next.

Theorem 4.7 ([9])**.** *There exists a universal constant C such that the following holds. Let A be an m by n real matrix such that $\|A\|_{1\to 2}$, and let K be a convex body in $\mathbb{R}^m$ such that $\Pr[g \in K] \geq 1/2$ where $g \in \mathbb{R}^m$ is a standard m-dimensional Gaussian random vector, and the probability is taken over the choice of g. Then there exists $x \in \{-1, 1\}^n$ such that $Ax \in CK$.*

We start our argument with the main technical lemmas.

Lemma 4.3. *Let $A = (a_j)_{j=1}^n \in \mathbb{R}^{m\times n}$, and let $F \in \mathbb{R}^{m\times m}$ be a rank m matrix such that $\forall j \in [n] : a_j \in E = FB_2^m$. Then there exists a matrix $X \succeq 0$ such that $\forall j \in [n] : X_{jj} = 1$ and $AXA^\intercal \preceq FF^\intercal$.*

Proof. Observe that, $a_j \in E \Leftrightarrow F^{-1}a_j \in B_2^m$. This implies $\|F^{-1}A\|_{1\to 2} \leq 1$, and, by Theorem 3.4, there exists an X with $X_{jj} = 1$ for all j such that $(F^{-1}A)X(F^{-1}A)^\intercal \preceq I$. Multiplying on the left by F and on the right by $F^\intercal$, we have $AXA^\intercal \preceq FF^\intercal$, and this completes the proof. □

Lemma 4.3 is our main tool for approximating hereditary vector discrepancy. By the relationship between vector discrepancy and discrepancy established by Bansal (Corollary 3.2), this is sufficient for a poly-logarithmic approximation to hereditary discrepancy. However, to get tight upper bounds on discrepancy (and improved approximation ratio), we give a direct argument using Banaszczyk's theorem.

Lemma 4.4. *Let $A = (a_j)_{j=1}^n \in \mathbb{R}^{m\times n}$, and let $F \in \mathbb{R}^{m\times m}$ be a rank m matrix such that $\forall j \in [n] : a_j \in E = FB_2^m$. Then, for any set of vectors $v_1, \ldots, v_k \in \mathbb{R}^m$, there exists $x \in \{\pm 1\}^n$ such that $\forall i \in [k] : |\langle Ax, v_i\rangle| \leq C\sqrt{(v_i^\intercal FF^\intercal v_i)\log k}$ for a universal constant C.*

Proof. Let $P \triangleq \{y : |\langle y, v_i\rangle| \leq \sqrt{v_i^\intercal FF^\intercal v_i}\ \forall i \in [k]\}$. We need to prove that there exists an $x \in \{-1, 1\}^n$ such that $Ax \in (C\sqrt{\log k})P$ for a suitable constant C. Set

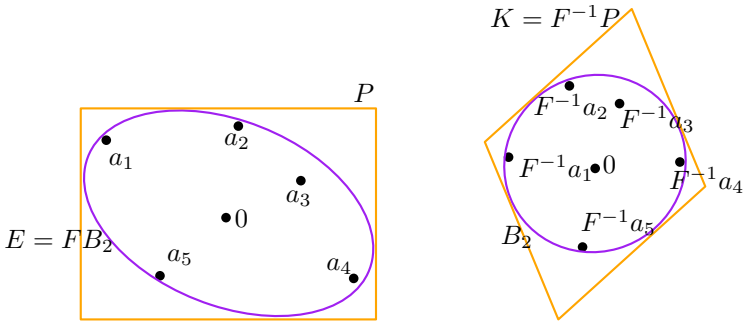


Figure 4.1: A linear transformation allows to apply Banaszczyk's theorem.

$K \triangleq F^{-1}P$. To show that there exists an x such that $Ax \in (C\sqrt{\log k})P$, we will show that there exists an $x \in \{-1,1\}^n$ such that $F^{-1}Ax \in (C\sqrt{\log k})K$. For this, we will use Theorem 4.7. As in the proof of Lemma 4.3, $\|F^{-1}A\|_{1\to 2} \le 1$. To use Theorem 4.7, we also need to argue that for a standard Gaussian g, $\Pr[g \in (C\sqrt{\log k})K] \ge \frac{1}{2}$.

For an intuitive explanation of the proof, see Figure 4.1. When the vectors v_i are unit length, the quantity $\sqrt{v_i^\intercal FF^\intercal v_i}$ is just half the width of E in the direction of v_i, and the bounding halfspaces of the polytope P are supporting hyperplanes of E. It follows that P contains E, which contains the columns of A. The map F^{-1} transforms E to a ball B_2^m, and P to the polytope K which contains B_2^m. The lower bound $\Pr[g \in (C\sqrt{\log k})K] \ge \frac{1}{2}$ follows from standard facts about Gaussians: either Sidak's lemma, or the Chernoff bound.

Let us first derive a representation of K as the intersection of slabs:

$$\begin{aligned} tK = tF^{-1}P &= \{tF^{-1}y : |\langle y, v_i\rangle| \le \sqrt{v_i^\intercal FF^\intercal v_i}\ \forall i \in [k]\} \\ &= \{z : |\langle Fz, v_i\rangle| \le t\sqrt{v_i^\intercal FF^\intercal v_i}\ \forall i \in [k]\} \\ &= \{z : |\langle z, F^\intercal v_i\rangle| \le t\sqrt{v_i^\intercal FF^\intercal v_i}\ \forall i \in [k]\}. \end{aligned}$$

Let g be a standard m-dimensional Gaussian vector. Then $\mathbb{E}_g|\langle g, F^\intercal v_i\rangle|^2 = v_iFF^\intercal v_i$; by standard concentration bounds, $\Pr[|\langle g, F^\intercal v_i\rangle|^2 > t^2(v_iFF^\intercal v_i)] < \exp(-t^2/2)$. Setting $t = \sqrt{2\ln 2k}$ and taking a union bound over all $i \in [k]$ gives us that $\Pr[g \notin \sqrt{2\ln 2k}\ K] < 1/2$. By Theorem 4.7, this implies that there exists an $x \in \{-1,1\}^n$ such that $F^{-1}Ax \in C\sqrt{2\ln 2k}\ K$, and, by multiplying on both sides by F, it follows that $Ax \in C\sqrt{2\ln 2k}\ P$.

□

The property that all columns of a matrix A are contained in E is *hereditary*: if it is satisfied for A, then it is satisfied for any submatrix of A. This elementary fact lends the power of Lemmas 4.3 and 4.4: the bound given by ellipsoids is *universal* in the sense that the discrepancy bound for any direction v_i holds for all submatrices A_J of A simultaneously. This fact makes it possible to upper bound hereditary discrepancy in arbitrary norms, and in the rest of the chapter we do this for ℓ_∞^m, which is the norm of interest for standard definitions of discrepancy. We consider ellipsoids E that contain the columns of A and minimize the quantity $\|E\|_\infty$: the largest ℓ_∞ norm of the points of E. Note that $\|E\|_\infty$, for an ellipsoid $E = FB_2^m$, can be written as

$$\|E\|_\infty = \max_{x\in E, y:\|y\|_1=1} \langle x, y\rangle = \max_{y:\|y\|_1=1} \|y\|_{E^\circ} = \max_{i\in[n]} \sqrt{e_i^\intercal F F^\intercal e_i}, \tag{4.4}$$

where the first identity follows since ℓ_1 is the dual norm to ℓ_∞, and the final identity follows from the formula for $\|\cdot\|_{E^\circ}$ and the fact that a convex function over the ℓ_1 ball is always maximized at a vertex, i.e. a standard basis vector e_i (e_i has 1 in the i-th coordinate and 0s everywhere else). The next definition and theorem give our main upper bound on hereditary (vector) discrepancy, which is in terms of $\|E\|_\infty$.

Definition 4.1. *For a matrix $A = (a_i)_{i=1}^n \in \mathbb{R}^{m\times n}$, the* ellipsoid-infinity norm *of A is defined as*

$$\|A\|_{E\infty} = \min\{\|E\|_\infty : a_i \in E\ \forall i \in n\}.$$

For a set system $\mathcal{S}$ with incidence matrix A, we define $\|\mathcal{S}\|_{E\infty} = \|A\|_{E\infty}$.

Theorem 4.8. *For any matrix $A \in \mathbb{R}^{m\times n}$,* $\operatorname{hvdisc}(A) \le \|A\|_{E\infty}$, *and* $\operatorname{herdisc}(A) = O(\sqrt{\log m}) \cdot \|A\|_{E\infty}$.

Proof. Let $\epsilon \ge 0$ be arbitrarily small and let F be a rank m matrix such that the ellipsoid $E = FB_2^m$ contains the columns of A and satisfies $\|E\|_\infty \le \|A\|_{E\infty} + \epsilon$. Let A_J be an arbitrary submatrix of A ($J \subseteq [n]$). Since all columns of A are contained in E, this holds for all columns of A_J as well, and by Lemma 4.3, we have that there exists $X \succeq 0$ with $X_{jj} = 1$ for all $j \in J$, and $A_J X A_J^\intercal \preceq FF^\intercal$. Therefore, for all $i \in [m]$, $e_i^\intercal A_J X A_J^\intercal e_i \le e_i F F^\intercal e_i \le \|E\|_\infty^2$, by (4.4). Since J was arbitrary and ϵ can be made

as small we as we like, this implies the bound on hvdisc(A). To bound herdisc(A), in Lemma 4.4 set $k = m$ and $v_i = e_i$ for $i \in [m]$ and e_i the i-th standard basis vector. $\square$

4.4 Lower Bounds on Discrepancy

In Section 4.3 we showed that the hereditary (vector) discrepancy of a matrix A can be *bounded from above* in terms of the $\|A\|_{E\infty}$. In this section we define $\|A\|_{E\infty}$ as a convex optimization problem, and show that it provides lower bounds for discrepancy as well. We use convex duality and the restricted invertibility theorem for this purpose. The lower bound we derive is new in discrepancy theory and we give further applications of it in Chapter 6.

4.4.1 The Ellipsoid Minimization Problem and Its Dual

Recall that for a block matrix

$$X = \begin{pmatrix} A & B \\ B^\intercal & C \end{pmatrix},$$

the *Schur complement* of an invertible block C in X is $A - B^\intercal C^{-1} B$. When $C \succ 0$, $X \succeq 0$ if and only if $A - B^\intercal C^{-1} B \succeq 0$.

To formulate the problem of minimizing $\|E\|_\infty = \max_{x \in E} \|x\|_\infty$ as a convex optimization problem we need the following well-known lemma, which shows that the matrix inverse is convex in the PSD sense. We give a proof for completeness.

Lemma 4.5. *For any two $m \times m$ matrices $X \succ 0$ and $Y \succ 0$ and any $\alpha \in [0, 1]$, $(\alpha X + (1-\alpha)Y)^{-1} \preceq \alpha X^{-1} + (1-\alpha)Y^{-1}$.*

Proof. Define the matrices

$$U = \begin{pmatrix} X^{-1} & I \\ I & X \end{pmatrix} \qquad V = \begin{pmatrix} Y^{-1} & I \\ I & Y \end{pmatrix}.$$

The Schur complement of X in U is 0, and therefore $U \succeq 0$, and analogously $V \succeq 0$. Therefore $\alpha U + (1-\alpha)V \succeq 0$, and the Schur complement of $\alpha X + (1-\alpha)Y$ in $\alpha U + (1-\alpha)V$ is also positive semidefinite, i.e. $\alpha X^{-1} + (1-\alpha)Y^{-1} - (\alpha X + (1-\alpha)Y)^{-1} \succeq 0$. This completes the proof, after re-arranging terms. $\square$

Consider a matrix $A = (a_j)_{j=1}^n \in \mathbb{R}^{m\times n}$ of rank m. Let us formulate $\|A\|_{E\infty}$ as a convex minimization problem. The problem is defined as follows

$$\text{Minimize } t \text{ s.t.} \tag{4.5}$$

$$X \succ 0 \tag{4.6}$$

$$\forall i \in [m] : e_i^\intercal X^{-1} e_i \leq t \tag{4.7}$$

$$\forall j \in [n] : a_j^\intercal X a_j \leq 1. \tag{4.8}$$

Lemma 4.6. *For a rank m matrix $A = (a_j)_{j=1}^n \in \mathbb{R}^{m\times n}$, the optimal value of the optimization problem (4.5)–(4.8) is equal to $\|A\|_{E\infty}$. Moreover, the objective function (4.5) and constraints (4.7)–(4.8) are convex over $t \in \mathbb{R}$ and $X \succ 0$.*

Proof. Let λ be the optimal value of (4.5)–(4.8). Given a feasible X for (4.5)–(4.8), set $E = X^{-1/2}B_2^m$ (this is well-defined since $X \succ 0$). Then for any $j \in [n]$, $\|a_j\|_E = a_j^\intercal X a_j \leq 1$ by (4.8), and, therefore, $a_j \in E$. Also, by (4.4), $\|E\|_\infty^2 = \max_{i=1}^m e_i^\intercal X e_i \leq t$. This shows that $\|A\|_{E\infty} \leq \lambda$. In the reverse direction, let $E = FB_2^m$ be such that $\forall j \in [n] : a_j \in E$. Then, because A is full rank, F is also full rank and invertible, and we can define $X = (FF^\intercal)^{-1}$ and $t = \|E\|_\infty^2$. Analogously to the calculations above, we can show that X and t are feasible, and therefore $\lambda \leq \|A\|_{E\infty}$.

The objective function and the constraints (4.8) are affine, and therefore convex. To show (4.7) are also convex, let X_1, t_1 and X_2, t_2 be two feasible solutions and let $\alpha \in [0,1]$. Then, Lemma 4.5 implies that for any i, $e_i^\intercal(\alpha X_1 + (1-\alpha)X_2)^{-1}e_i \leq \alpha X_1^{-1} + (1-\alpha)X_2^{-1} \leq \alpha t_1 + (1-\alpha)t_2$, so constraints (4.7) are convex as well. □

The Schatten 1-norm of a matrix M, also known as the trace norm or the nuclear norm, is equal to $\|M\|_{S_1} = \sum_i \sigma_i(M) = \mathrm{tr}((MM^\intercal)^{1/2})$, where $X^{1/2}$ denotes the positive semidefinite root of the matrix $X \succeq 0$. The dual of (4.5)–(4.8) is a problem of maximizing the nuclear norm over re-weightings of the columns and rows of A. Before we prove this fact, let us cite a theorem from convex analysis, which will be used in our proof.

Lemma 4.7 (Corollary 7.5.1. in [124])**.** *Let $f\colon \mathbb{R}^n \to \mathbb{R}\cup\{-\infty\}$ be an upper-semicontinuous concave function. Then for any x and y in the effective domain of f (i.e. $f(x), f(y) >$*

$-\infty$*),*

$$f(x) = \lim_{\lambda \uparrow 1} f(\lambda x + (1-\lambda)y).$$

Lemma 4.7 states that an upper semi-continuous concave function is continuous along any segment in its effective domain. More general statements are known, but we will not need them.

We now state the dual characterization of (4.5)–(4.8).

Theorem 4.9. *Let $A = (a_j)_{j=1}^n \in \mathbb{R}^{m\times n}$ be a rank m matrix. Then,*

$$\|A\|_{E\infty}^2 = \max \|P^{1/2} A Q^{1/2}\|_{S_1}^2 \ \ s.t. \tag{4.9}$$

$$\operatorname{tr}(P) = \operatorname{tr}(Q) = 1 \tag{4.10}$$

$$P, Q \succeq 0; P, Q \ \textit{diagonal}. \tag{4.11}$$

Proof. We shall prove the theorem by showing that the convex optimization problem (4.5)–(4.8) satisfies Slater's condition, and its Lagrange dual is equivalent to (4.9)–(4.11). Let us first verify Slater's condition. We define the domain for constraints (4.7) as the open cone $\{X : X \succ 0\}$, which makes the constraint $X \succ 0$ implicit. Let $d = \|A\|_{1\to 2}$, $X = \frac{1}{d} I$, and $t = d + \varepsilon$ for some $\varepsilon > 0$. Then the affine constraints (4.8) are satisfied exactly, and the constraints (4.7) are satisfied with slack since $\varepsilon > 0$. Moreover, by Lemma 4.6, all the constraints and the objective function are convex. Therefore, (4.5)–(4.8) satisfies Slater's condition, and consequently strong duality holds.

The Lagrange dual function for (4.5)–(4.8) is by definition

$$g(p, r) = \inf_{t, X \succ 0} t + \sum_{i=1}^m p_i (e_i^\intercal X^{-1} e_i - t) + \sum_{j=1}^n r_j (a_j^\intercal X a_j - 1),$$

with dual variables $p \in \mathbb{R}^m$ and $r \in \mathbb{R}^n$, $p, r \geq 0$. Equivalently, writing p as a diagonal matrix $P \in \mathbb{R}^{m\times m}$, $P \succeq 0$, r as a diagonal matrix $R \in \mathbb{R}^{n\times n}$, $R \succeq 0$, we have $g(P, R) = \inf_{t, X\succ 0} t + \operatorname{tr}(PX^{-1}) - \operatorname{tr}(tP) + \operatorname{tr}(ARA^\intercal X) - \operatorname{tr}(R)$. If $\operatorname{tr}(P) \neq 1$, then $g(P, R) = -\infty$, since we can take t to $-\infty$ while keeping X fixed. On the other hand, for $\operatorname{tr}(P) = 1$, the dual function simplifies to

$$g(P, R) = \inf_{X \succ 0} \operatorname{tr}(PX^{-1}) + \operatorname{tr}(ARA^\intercal X) - \operatorname{tr}(R). \tag{4.12}$$

Since $X \succ 0$ implies $X^{-1} \succ 0$, $g(P,R) \geq -\mathrm{tr}(R) > -\infty$ whenever $\mathrm{tr}(P) = 1$. Therefore, the effective domain $\{(P,R) : g(P,R) > -\infty\}$ of $g(P,R)$ is the set of pairs of diagonal non-negative matrices (P,R) such that $\mathrm{tr}(P) = 1$.

Let us first consider the case when P and $ARA^\intercal$ are both invertible. After differentiating the right hand side of (4.12) with respect to X, we get the first-order optimality condition

$$X^{-1}PX^{-1} = ARA^\intercal. \tag{4.13}$$

Multiplying by $P^{1/2}$ on the left and the right and taking square roots gives the equivalent condition $P^{1/2}X^{-1}P^{1/2} = (P^{1/2}ARA^\intercal P^{1/2})^{1/2}$. This equation has a unique solution, since P and $ARA^\intercal$ were both assumed to be invertible. Since $\mathrm{tr}(PX^{-1}) = \mathrm{tr}(P^{1/2}X^{-1}P^{1/2})$ and also, by (4.13), $\mathrm{tr}(ARA^\intercal X) = \mathrm{tr}(X^{-1}P) = \mathrm{tr}(PX^{-1})$, we simplify $g(P,R)$ to

$$g(P,R) = 2\mathrm{tr}((P^{1/2}ARA^\intercal P^{1/2})^{1/2}) - \mathrm{tr}(R) = 2\|P^{1/2}AR^{1/2}\|_{S_1} - \mathrm{tr}(R). \tag{4.14}$$

We will now use Lemma 4.7 to argue that equation (4.14) holds also when P and and $ARA^\intercal$ are not invertible. Fix any non-negative diagonal matrices P and R such that $\mathrm{tr}(P) = 1$ (i.e. any P and R in the domain of g), and for $\lambda \in [0,1]$ define $P(\lambda) \triangleq \lambda P + (1-\lambda)\frac{1}{m}I$ and $R(\lambda) \triangleq \lambda R + (1-\lambda)I$. Observe that for any $\lambda \in [0,1)$, $P(\lambda)$ is invertible, and, because $AA^\intercal \succ 0$ by the assumption that A is of full row-rank m, $AR(\lambda)A^\intercal$ is also invertible. Then, by Lemma 4.7 and (4.14), we have

$$\begin{aligned} g(P,R) = \lim_{\lambda \uparrow 1} g(P(\lambda), R(\lambda)) &= \lim_{\lambda \uparrow 1} \left[2\|P(\lambda)^{1/2}AR(\lambda)^{1/2}\|_{S_1} - \mathrm{tr}(R(\lambda))\right] \\ &= 2\|P^{1/2}AR^{1/2}\|_{S_1} - \mathrm{tr}(R). \end{aligned}$$

where the last equality follows since the nuclear norm and the trace function are continuous.

We showed that (4.5)–(4.8) satisfies Slater's condition and therefore strong duality holds, so by Theorem 4.6 and Lemma 4.6, $\|A\|_{E\infty}^2 = \max\{g(P,R) : \mathrm{tr}(P) = 1, P, R \succeq 0, \text{ diagonal}\}$. Let us define new variables Q and c, where $c = \mathrm{tr}(R)$ and $Q = R/c$. Then we can re-write $g(P,R)$ as

$$g(P,R) = g(P,Q,c) = 2\|P^{1/2}A(cQ)^{1/2}\|_{S_1} - \mathrm{tr}(cQ) = 2\sqrt{c}\|P^{1/2}AQ^{1/2}\|_{S_1} - c.$$

From the first-order optimality condition $\frac{dg}{dc} = 0$, we see that maximum of $g(P, Q, c)$ is achieved when $c = \|P^{1/2}AQ^{1/2}\|_{S_1}^2$ and is equal to $\|P^{1/2}AQ^{1/2}\|_{S_1}^2$. Therefore, maximizing $g(P, R)$ over diagonal positive semidefinite P and R such that $\mathrm{tr}(P) = 1$ is equivalent to the optimization problem (4.9)–(4.11). This completes the proof. □

4.4.2 Spectral Lower Bounds via Restricted Invertibility

In this subsection we relate the dual formulations of the min-ellipsoid problem from Section 4.4.1 to the dual of vector discrepancy. The connection is via the restricted invertibility principle and gives our main lower bounds on hereditary (vector) discrepancy.

Let us first derive a simple lower bound on $\mathrm{vecdisc}(A)$ from the dual (3.12)–(3.15).

Lemma 4.8. *For any $m \times n$ matrix A, and any $m \times m$ diagonal matrix $P \geq 0$ with $\mathrm{tr}(P) = 1$, we have*

$$\mathrm{vecdisc}(A) \geq \sqrt{n}\sigma_{\min}(P^{1/2}A).$$

Proof. Observe that the solution (P, Q), where $Q \triangleq \sigma_{\min}(P^{1/2}A)^2 I$, is feasible for (3.12)–(3.15). By Proposition 3.2, $\mathrm{vecdisc}(A)^2 \geq \mathrm{tr}(Q) = n\sigma_{\min}(P^{1/2}A)^2$. □

We define a *spectral lower bound* based on Lemma 4.8.

$$\mathrm{specLB}(A) \triangleq \max_{k=1}^{n} \max_{J\subseteq[n]:|J|=k} \max_{P} \sqrt{k}\sigma_{\min}(P^{1/2}A_J),$$

where P ranges over positive (i.e. $P \succeq 0$) $m \times m$ diagonal matrices satisfying $\mathrm{tr}(P) = 1$. Lemma 4.8 implies immediately that $\mathrm{hvdisc}(A) \geq \mathrm{specLB}(A)$.

The next lemma relates the dual characterization of $\|A\|_{E\infty}$ to the spectral lower bound

Lemma 4.9. *Let M be an m by n real matrix, and let $W \succeq 0$ be a diagonal matrix such that $\mathrm{tr}(W) = 1$ and $r \triangleq \mathrm{rank}\ MW^{1/2}$. Then there exists a submatrix M_J of M, $|J| \leq r$, such that $|J|\sigma_{\min}(M_J)^2 \geq \frac{c^2\|MW^{1/2}\|_{S_1}^2}{(\log r)^2}$, for a universal constant $c > 0$. Moreover, given M as input, J can be computed in deterministic polynomial time.*

Proof. By homogeneity of the nuclear norm and the smallest singular value, it suffices to show that if $\|MW^{1/2}\|_{S_1}^2 = 1$, then $|J|\sigma_{\min}(M_J)^2 \geq \frac{c^2}{(\log r)^2}$ for a set $J \subseteq [n]$ of size at most r. Let us define $\tilde{M} \triangleq MW^{1/2}$.

Let $K_t \triangleq \{i \in [r] : 2^{-t-1} \leq \sigma_i(\tilde{M}) \leq 2^{-t}\}$ for an integer $0 \leq t \leq \log_2 r$, and $T = \{i \in [r] : 0 < \sigma_i(\tilde{M}) \leq \frac{1}{2r}\}$. Then

$$\sum_{t=0}^{\log_2 r} \sum_{i \in K_t} \sigma_i(\tilde{M}) = 1 - \sum_{i \in T} \sigma_i(\tilde{M}) \geq 1/2,$$

since $|T| \leq r$. Therefore, by averaging, there exists a t^* such that $\sum_{i \in K_{t^*}} \sigma_i(\tilde{M}) \geq \frac{1}{2\log_2 r}$; for convenience, let us define $K \triangleq K_{t^*}$, $k \triangleq |K| \leq r$, and $\alpha \triangleq \frac{1}{2\log_2 r}$.

Next, we define a suitable $k \times n$ matrix with singular values σ_i, $i \in K$. Let $\tilde{M} = U\Sigma V^\intercal$ be the singular value decomposition of $\tilde{M}$, with U and V orthogonal, and Σ diagonal with $\sigma_1(\tilde{M}), \ldots, \sigma_m(\tilde{M})$ on the main diagonal. Set U_K to be the submatrix of U whose columns are the left singular vectors corresponding to $\sigma_i(\tilde{M})$ for $i \in K$, and define the projection matrix $\Pi \triangleq U_K U_K^\intercal$. The nonzero singular values of $\Pi\tilde{M} = U_K \Sigma_K V^\intercal$ are exatly those $\sigma_i(\tilde{M})$ for which $i \in K$, as desired. We have $\|\Pi\tilde{M}\|_{S_1} \geq \alpha$ by the choice of K, and $\|\Pi\tilde{M}\|_2 \leq 2\alpha/k$ because all values of $\Pi\tilde{M}$ are within a factor of 2 from each other. Finally, applying Cauchy-Schwarz to the singular values of $\Pi\tilde{M}$, we have that $\|\Pi\tilde{M}\|_{HS} \geq \alpha/k^{1/2}$. By Theorem 4.5, applied to M and W with $\epsilon = \frac{1}{2}$, there exists a set J of size $r \geq k \geq |J| \geq k/16$ such that $\sigma_{\min}(\Pi M_J)^2 \geq \alpha^2/4k$, implying that

$$|J|\sigma_{\min}(M_J)^2 \geq |J|\sigma_{\min}(\Pi M_J)^2 \geq \frac{1}{64}\alpha^2.$$

Finally, J can be computed in deterministic polynomial time, by Theorem 4.5. □

Theorem 4.10. *For any rank m matrix $A \in \mathbb{R}^{m \times n}$,*

$$\|A\|_{E\infty} = O(\log m)\ \operatorname{hvdisc}(A).$$

Moreover, we can compute in deterministic polynomial time a set $J \subseteq [n]$ such that $\|A\|_{E\infty} = O(\log m) \operatorname{vecdisc}(A_J)$.

Proof. Let P and Q be optimal solutions for (4.9)-(4.11). By Theorem 4.9, $\|A\|_{E\infty} = \|P^{1/2}AQ^{1/2}\|_{S_1}$. Then, by Lemma 4.9, applied to the matrices $M = P^{1/2}A$ and $W = Q$,

there exists a set $J \subseteq [n]$, computable in deterministic polynomial time, such that

$$\mathrm{specLB}(A) \geq \sqrt{|J|}\sigma_{\min}(P^{1/2}A_J) \geq \frac{c\|P^{1/2}AQ^{1/2}\|_{S_1}}{\log m} = \frac{c\|A\|_{E\infty}}{\log m}. \tag{4.15}$$

□

By a similar argument, but using Lemma 4.2 in the place of Theorem 4.5, we show that $\|A\|_{E\infty}$ approximates $\mathrm{detlb}(A)$.

Theorem 4.11. *There exists a constant C such that for any $m \times n$ matrix A of rank r*

$$\mathrm{detlb}(A) \leq \|A\|_{E\infty} = O(\log r) \cdot \mathrm{detlb}(A).$$

Proof. For the inequality $\mathrm{detlb}(A) \leq \|A\|_{E\infty}$, we first observe that if B is a $k \times k$ matrix, then

$$|\det B|^{1/k} \leq \frac{1}{k}\|B\|_{S_1} \tag{4.16}$$

Indeed, the left-hand side is the geometric mean of the singular values of B, while the right-hand side is the arithmetic mean.

Now let $B = A_{I,J}$ be a $k \times k$ submatrix of A, with rows indexed by the set I and columns indexed by the set J, with $\mathrm{detlb}(A) = |\det(B)|^{1/k}$. Define $P = \frac{1}{k}\operatorname{diag}(1_I)$ and $Q = \frac{1}{k}\operatorname{diag}(1_J)$, where 1_I and 1_J are, respectively, the indicator vectors of the sets I and J. By Theorem 4.9,

$$\mathrm{detlb}(A) = |\det(B)|^{1/k} \leq \frac{1}{k}\|B\|_{S_1} = \|P^{1/2}AQ^{1/2}\|_{S_1} \leq \|A\|_{E\infty}.$$

For the second inequality $\|A\|_{E\infty} \leq O(\log m) \cdot \mathrm{detlb}(A)$, we use a strategy analogous to the proof of Lemma 4.9. By homogeneity, we can again assume, without loss of generality, that $\|A\|_{E\infty} = 1$. Let P and Q be optimal solutions to (4.9)-(4.11), so that $\|P^{1/2}AQ^{1/2}\|_{S_1} = 1$ by Theorem 4.9. For brevity, let us write $\tilde{A} \triangleq P^{1/2}AQ^{1/2}$, and let $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ be the nonzero singular values of $\tilde{A}$.

By an argument analogous to the one we used in the proof of Lemma 4.9, there is some integer t such that if we set $K \triangleq \{i \in [m] : 2^{-t-1} \leq \sigma_i < 2^{-t}\}$, then

$$\sum_{i \in K} \sigma_i \geq \alpha \triangleq \frac{1}{2\log_2 r}$$

Let us set $k \triangleq |K|$.

As in Lemma 4.9, we define a $k \times n$ matrix with singular values σ_i, $i \in K$. Let $\tilde{A} = U\Sigma V^\intercal$ be the singular-value decomposition of $\tilde{A}$. Set U_K to be the submatrix of U whose columns are the left singular vectors corresponding to σ_i for $i \in K$. The singular values of $B \triangleq U_K^\intercal \tilde{A} = U_K \Sigma_K V^\intercal$ are exatly those σ_i for which $i \in K$, as desired. As all σ_i for $i \in K$ are within a factor of 2 from each other, we have, by the choice of K,

$$|\det(BB^\intercal)|^{1/2k} = \Big(\prod_{i\in K} \sigma_i\Big)^{1/k} \geq \frac{1}{2k}\sum_{i\in K}\sigma_i \geq \frac{1}{2k}\alpha.$$

It remains to relate $\det BB^\intercal$ to the determinant of a square submatrix of A, and this is where Lemma 4.2 is applied—actually applied twice, once for columns, and once for rows.

First we set $C \triangleq U_K^\intercal P^{1/2} A$; then $B = CQ^{1/2}$. Applying Lemma 4.2 with C in the role of M and Q in the role of W, we obtain a k-element index set $J \subseteq [n]$ such that

$$|\det(C_J)|^{1/k} \geq \sqrt{k/e} \cdot |\det(BB^\intercal)|^{1/2k}.$$

Next, we set $D_J \triangleq P^{1/2} A_J$, and we claim that $\det(D_J^\intercal D_J) \geq (\det C_J)^2$. Indeed, we have $C_J = U_K^\intercal D_J$, and, since U is an orthogonal transformation, $(U^\intercal D_J)^\intercal (U^\intercal D_J) = D_J^\intercal D_J$. Then, by the Binet–Cauchy formula,

$$\begin{aligned}\det(D_J^\intercal D_J) = \det((U^\intercal D)^\intercal (U^\intercal D)) &= \sum_L \det(U_L^\intercal D_J)^2 \\ &\geq \det(U_K^\intercal D_J)^2 = (\det C_J)^2.\end{aligned}$$

The next (and last) step is analogous. We have $D_J^\intercal = A_J^\intercal P^{1/2}$, and so we apply Lemma 4.2 with $A_J^\intercal$ in the role of M and P in the role of W, obtaining a k-element subset $I \subseteq [m]$ with $|\det A_{I,J}|^{1/k} \geq \sqrt{k/e} \cdot |\det(D_J^\intercal D_J)|^{1/2k}$ (where $A_{I,J}$ is the submatrix of A with rows indexed by I and columns by J).

Following the chain of inequalities backwards, we have

$$\begin{aligned}\mathrm{detlb}(A) &\geq |\det(A_{I,J})|^{1/k} \geq \sqrt{k/e} \cdot |\det(D_J^\intercal D_J)|^{1/2k} \geq \sqrt{k/e} \cdot |\det(C_J)|^{1/k} \\ &\geq (k/e)|\det(BB^\intercal)|^{1/2k} \geq \frac{1}{2e}\alpha,\end{aligned}$$

and the theorem is proved. □

4.5 The Approximation Algorithm

We are now ready to give our approximation algorithm for hereditary vector discrepancy and hereditary discrepancy. In fact, the algorithm is a straightforward consequence of the upper and lower bounds we proved in the prior sections.

Theorem 4.12. *Given a real matrix $A \in \mathbb{R}^{m \times n}$, $\|A\|_{E\infty}$ can be approximated to within any degree of accuracy in deterministic polynomial time, and satisfies the inequalities*

$$\frac{1}{O(\log m)}\|A\|_{E\infty} \leq \operatorname{hvdisc}(A) \leq \|A\|_{E\infty},$$
$$\frac{1}{O(\log m)}\|A\|_{E\infty} \leq \operatorname{herdisc}(A) \leq O(\log^{1/2} m) \cdot \|A\|_{E\infty},$$
$$\frac{1}{O(\log m)}\|A\|_{E\infty} \leq \operatorname{detlb}(A) \leq \|A\|_{E\infty}.$$

Moreover, the algorithm finds a submatrix A_J of A, such that $\frac{1}{O(\log m)}\|A\|_{E\infty} \leq \operatorname{vecdisc}(A_J)$.

Proof. We first ensure that the matrix A is of rank m by adding a tiny full rank perturbation to it, and adding extra columns if necessary[1]. By making the perturbation small enough, we can ensure that it affects $\operatorname{herdisc}(A)$ and $\operatorname{hvdisc}(A)$ negligibly. The approximation guarantees follow from Theorems 4.8 and 4.10, and S is computed as in Theorem 4.10.

To compute $\|A\|_{E\infty}$ in polynomial time, we solve (4.5)–(4.8). By Lemma 4.6, this is a convex minimization problem, and as such can be solved using the ellipsoid method up to an ϵ-approximation in time polynomial in the input size and in $\log \epsilon^{-1}$ [72]. The optimal value is equal to $\|A\|_{E\infty}$ by Lemma 4.6, and, therefore, we can compute an arbitrarily good approximation to $\|A\|_{E\infty}$ in polynomial time. □

Observe that Theorem 4.12 implies Theorems 4.1 and 4.2.

Bibliographic Remarks

The first polynomial time approximation algorithm for hereditary discrepancy with a polylogarithmic approximation guarantee was published in [118], and was a corollary

[1]There are other, more numerically stable ways to reduce to the full rank case, e.g. by projecting A onto its range and modifying the norms we consider accordingly. We choose the perturbation approach for simplicity.

of work in differential privacy. The approach in the current chapter is more direct, achieves an improved approximation ratio, and make explicit the central quantity of interest: the ellipsoid infinity norm. The material in the chapter is the result of joint work with Kunal Talwar, and a preliminary version appears at [116].

Chapter 5

More on the Ellipsoid Infinity Norm

5.1 Overview

In this chapter we prove that the ellipsoid infinity norm satisfies a number of nice properties. We show that it is invariant under transposition, satisfies the triangle inequality (and, therefore, is a matrix norm), and is multiplicative with respect to tensor products. Moreover, we prove strengthenings of the triangle inequality in some special cases when the matrices have disjoint support. These properties will be exploited in Chapter 6, where we use them to give remarkably easy proofs of new and classical upper and lower bounds on the discrepancy of natural set systems.

We additionally give examples for which each of the upper and lower bounds in Theorem 4.12 are tight.

5.2 Properties of the Ellipsoid-Infinity Norm

Here we give several useful properties of the ellipsoid infinity norm. These properties make it possible to reason about the the ellipsoid infinity norm of a complicated matrix by decomposing it as the sums of simple matrices. Since the ellipsoid infinity norm approximates hereditary discrepancy, the properties hold approximately for herdisc too, and we will see a number of applications of them in Chapter 6.

5.2.1 Transposition and Triangle Inequality

Two properties of $\|A\|_{E\infty}$ that are not obvious from the definition, but follows easily from Theorem 4.9, are that $\|A^\intercal\|_{E\infty} = \|A\|_{E\infty}$ and $\|A + B\|_{E\infty} \leq \|A\|_{E\infty} + \|B\|_{E\infty}$. We prove both next.

Proposition 5.1. *For any real matrix* $\|A\|_{E\infty} = \|A^\intercal\|_{E\infty}$.

Proof. It is easy to see that the nuclear norms $\|M\|_{S_1}$ and $\|M^{\mathrm{tr}}\|_{S_1}$ are equal. Indeed, M and $M^\intercal$ have the same nonzero singular values, and, therefore, the respective sums of singular values are also equal. Now, given A, let P, and Q be such that $\|P^{1/2}AQ^{1/2}\|_{S_1}$, as in Theorem 4.9. We have

$$\|A\|_{E\infty} = \|P^{1/2}AQ^{1/2}\|_{S_1} = \|(Q^{1/2})^T A^T (P^{1/2})^T\|_{S_1} = \|Q^{1/2}A^T P^{1/2}\|_{S_1} \leq \|A^T\|_{E\infty}.$$

The opposite inequality follows symmetrically. □

Proposition 5.2. *For any two* $m\times n$ *real matrices* A, B, $\|A+B\|_{E\infty} \leq \|A\|_{E\infty}+\|B\|_{E\infty}$

Proof. Let P, Q be such that $\|P^{1/2}(A+B)Q^{1/2}\|_{S_1} = \|A+B\|_{E\infty}$. Since $\|\cdot\|_{S_1}$ is a matrix norm and satisfies the triangle inequality (see [24, Sec. IV.2]), we have

$$\begin{aligned} \|A+B\|_{E\infty} &= \|P^{1/2}(A+B)Q^{1/2}\|_{S_1} = \|P^{1/2}AQ^{1/2} + P^{1/2}BQ^{1/2}\|_{S_1} \\ &\leq \|P^{1/2}AQ^{1/2}\|_{S_1} + \|P^{1/2}BQ^{1/2}\|_{S_1} \leq \|A\|_{E\infty} + \|B\|_{E\infty}. \end{aligned}$$

□

5.2.2 Unions and Direct Sums

The next proposition sometimes strengthens the triangle inequality when the ranges of the matrices A and B lie in orthogonal coordinate subspaces. Unlike the previous two propositions, this one appears easier to prove from the definition of $\|\cdot\|_{E\infty}$, rather than the dual characterization. We recall (4.4), which states that for any ellipsoid $E = FB_2^m$,

$$\|E\|_\infty = \max_{i=1}^{m} \sqrt{e_i^\intercal F F^\intercal e_i},$$

where e_i is the i-th standard basis vector.

Proposition 5.3. *Let* $A_1, \ldots, A_k$ *be real matrices, each with* n *columns. For the matrix*

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix},$$

we have

$$\|A\|_{E\infty} \leq \sqrt{k} \max_{i=1}^{k} \|A_i\|_{E\infty}.$$

Proof. Let each A_i have m_i rows, and let $m \triangleq m_1 + \ldots + m_k$. Let also $E_i = F_i B_2^{m_i}$ be the ellipsoid that achieves $\|A_i\|_{E\infty}$. Define a new ellipsoid E as $E \triangleq F B_2^m$ where

$$F = \sqrt{k} \begin{pmatrix} F_1 & 0 & \ldots & 0 \\ 0 & F_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & F_k \end{pmatrix}.$$

I.e., F is the direct sum of $F_1, \ldots, F_k$. It is clear from (4.4) that $\|E\|_\infty = \sqrt{k} \max_{i=1}^{k} \|E_i\|_\infty$. To finish the proof, we need to show that any column a of A is contained in E. Let Π_i be the projection onto the coordinate subspace corresponding to the rows of A_i. Then, $\Pi_i a$ is a column of A_i. Let us write the ellipsoid E in the form $E = \{x : x^\intercal (FF^\intercal)^{-1} x \leq 1\}$, where

$$(FF^\intercal)^{-1} = \frac{1}{k} \begin{pmatrix} (F_1 F_1^\intercal)^{-1} & 0 & \ldots & 0 \\ 0 & (F_2 F_2^\intercal)^{-1} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & (F_k F_k^\intercal)^{-1} \end{pmatrix},$$

and, analogously, write $E_i = \{x : x^\intercal (F_i F_i^\intercal)^{-1} x \leq 1\}$. By the choice of the E_i, $\Pi_i a \in E_i$, so $a^\intercal \Pi_i (F_i F_i^\intercal)^{-1} \Pi_i a \leq 1$ for all i. It follows that,

$$a^\intercal (FF^\intercal)^{-1} a = \frac{1}{k} \sum_{i=1}^{k} a^\intercal \Pi_i (F_i F_i^\intercal)^{-1} \Pi_i a \leq 1,$$

and, therefore $a \in E$, as desired. This finishes the proof. □

We remark that for the setting of Proposition 5.3, Matoušek [107] proved the stronger bound

$$\|A\|_{E\infty} \leq \sqrt{\|A_1\|_{E\infty}^2 + \ldots + \|A_k\|_{E\infty}^2}.$$

An even stronger bound is possible when we take the direct sum of matrices.

Proposition 5.4. *If A is the block-diagonal matrix*

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix},$$

where A_1 and A_2 are real matrices, then then $\|A\|_{E\infty} = \max(\|A_1\|_{E\infty}, \|A_2\|_{E\infty})$.

Proof. Let, as in the proof of Proposition 5.3, each A_i have m_i rows, and let $m \triangleq m_1 + m_2$. Let also $E_i = F_i B_2^{m_i}$ be the ellipsoid that achieves $\|A_i\|_{E\infty}$. Define a new ellipsoid E as $E \triangleq F B_2^m$ where

$$F = \begin{pmatrix} F_1 & 0 \\ 0 & F_2 \end{pmatrix}.$$

It is clear from (4.4) that $\|E\|_\infty = \max\{\|E_1\|_\infty, \|E_2\|_\infty\}$. To finish the proof, we need to show that any column a of A is contained in E. Let us write the ellipsoid E in the form $E = \{x : x^\intercal (FF^\intercal)^{-1} x \le 1\}$, where

$$(FF^\intercal)^{-1} = \begin{pmatrix} (F_1 F_1^\intercal)^{-1} & \\ 0 & (F_2 F_2^\intercal)^{-1} \end{pmatrix},$$

and, analogously, write $E_i = \{x : x^\intercal (F_i F_i^\intercal)^{-1} x \le 1\}$. Notice that for any column a of A, $a^\intercal (FF^\intercal)^{-1} a$ is equal to $b^\intercal (F_i F_i^\intercal)^{-1} b$ for some column b of A_i, $i \in \{1, 2\}$. By the choice of E_i, $b \in E_i$, and, therefore,

$$a^\intercal (FF^\intercal)^{-1} a = b^\intercal (F_i F_i^\intercal)^{-1} b \le 1.$$

This finishes the proof. □

5.3 Tensor product

Here we show that $\|\cdot\|_{E\infty}$ is multiplicative with respect to tensor products. This fact is going to prove very useful in analyzing the combinatorial discrepancy of axis-aligned boxes in Chapter 6.

5.3.1 Properties of Tensor Products

The tensor product of matrices (a.k.a. Kronecker product) exhibits a number of useful properties with respect to matrix multiplication and addition. In an abstract setting, these properties are often used as the definition of the tensor product. We list them next. The properties hold for any complex matrices for which the operations make

sense given the dimensions; we write them in terms of matrices, but since they hold for single row or single column matrices, they are valid for vectors as well.

1. **Bilinearity**

$$A \otimes B + A \otimes C = A \otimes (B + C);$$
$$A \otimes B + C \otimes B = (A + C) \otimes B.$$

2. **Scaling** For any constant c:

$$c(A \otimes B) = (cA) \otimes B = A \otimes (cB).$$

3. **Conjugate Transpose**

$$(A \otimes B)^* = A^* \otimes B^*.$$

4. **Mixed Products**

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD).$$

The following well-known (and simple) lemma characterizes the singular value decomposition of tensor products.

Lemma 5.1. *For any two matrices $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times q}$, with singular value decompositions $A = U_A \Sigma_A V_A^*$ and $B = U_B \Sigma_B V_B^*$, the matrix $A \otimes B$ has singular value decomposition $A \otimes B = (U_A \otimes U_B)(\Sigma_A \otimes \Sigma_B)(V_A \otimes V_B)^*$.*

Proof. The lemma follows easily from the properties of tensor products. We first verify that the claimed decomposition indeed equals $A \otimes B$:

$$\begin{aligned}(U_A \otimes U_B)(\Sigma_A \otimes \Sigma_B)(V_A \otimes V_B)^* &= ((U_A \Sigma_A) \otimes (U_B \Sigma_B)) \otimes (V_A \otimes V_B)^* \\ &= ((U_A \Sigma_A) \otimes (U_B \Sigma_B)) \otimes (V_A^* \otimes V_B^*) \\ &= (U_A \Sigma_A V_A^*) \otimes (U_B \Sigma_B V_B^*) = A \otimes B.\end{aligned}$$

Then it remains to verify that this is indeed a singular value decomposition. The matrix $\Sigma_A \otimes \Sigma_B$ is easily seen to be diagonal. Also,

$$(U_A \otimes U_B)^*(U_A \otimes U_B) = (U_A^* \otimes U_B^*)(U_A \otimes U_B) = (U_A^* U_A) \otimes (U_B^* U_B) = I \otimes I = I,$$

and, therefore, $U_A \otimes U_B$ is orthonormal. By an analogous argument, $V_A \otimes V_B$ is orthonormal, and this completes the proof. □

5.3.2 Multiplicativity of the Ellipsoid Infinity Norm

Theorem 5.1. *For any $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{r\times s}$, $\|A \otimes B\|_{E\infty} = \|A\|_{E\infty}\|B\|_{E\infty}$.*

Proof. Let us first prove that $\|A \otimes B\|_{E\infty} \leq \|A\|_{E\infty}\|B\|_{E\infty}$. We can approximate A and B arbitrarily well by matrices with ranks m and r, respectively; for the rest of the proof we shall assume that A and B do each have ful row-rank. Let $E_A = F_A B_2^m$ achieve $\|A\|_{E\infty}$, and $E_B = F_B B_2^r$ achieve $\|B\|_{E\infty}$. Consider the matrix $F = F_A \otimes F_B$; using the mixed product property of the tensor product, we can verify that $FF^\intercal = (F_A F_A) \otimes (F_B F_B)$. $(FF^\intercal)^{-1} = (F_A F_A^\intercal)^{-1} \otimes (F_B F_B^\intercal)^{-1}$. Then we can write the ellipsoids E_A, E_B, and $E = FB_2^{mr}$ as

$$E_A = \{x : x^\intercal (F_A F_A^\intercal)^{-1} x \leq 1\},$$
$$E_B = \{x : x^\intercal (F_B F_B^\intercal)^{-1} x \leq 1\},$$
$$E = \{x : x^\intercal (FF^\intercal)^{-1} x \leq 1\}.$$

Each column of $A \otimes B$ is the tensor product $a \otimes b$ of a column a of A and a column b of B; then, using the mixed product property again,

$$(a \otimes b)^\intercal (FF^\intercal)^{-1}(a \otimes b) = (a^\intercal (F_A F_A^\intercal)^{-1} a)(b^\intercal (F_B F_B^\intercal)^{-1} b) \leq 1,$$

where the last inequality follows since $a \in E_A$ and $b \in E_B$. Therefore, the ellipsoid E contains the columns of $A \otimes B$, and has infinity norm, by (4.4),

$$\begin{aligned}
\|E\|_\infty &= \max_{i\in[m], j\in[r]} e_{i,j}^\intercal FF^\intercal e_{i,j} \\
&= \max_{i\in[m], j\in[r]} (e_i \otimes e_j)^\intercal (F_A F_A^\intercal \otimes F_B F_B^\intercal)(e_i \otimes e_j) \\
&= (\max_{i\in[m]} e_i^\intercal F_A F_A^\intercal e_i)(\max_{i\in[r]} e_i^\intercal F_B F_B^\intercal e_i) = \|A\|_{E\infty}^2 \|B\|_{E\infty}^2
\end{aligned}$$

Above $e_{i,j}$ is the standard basis vector corresponding to the pair of coordinates (i, j). Taking square roots, this proves that $\|A \otimes B\|_{E\infty} \leq \|A\|_{E\infty}\|B\|_{E\infty}$.

To prove $\|A \otimes B\|_{E\infty} \geq \|A\|_{E\infty}\|B\|_{E\infty}$, we use the dual characterization in Theorem 4.9. Let P_A, Q_A and P_B, Q_B be such that $\|P_A^{1/2} A Q_A^{1/2}\|_{S_1} = \|A\|_{E\infty}$ and $\|P_B^{1/2} A Q_B^{1/2}\|_{S_1} = \|B\|_{E\infty}$, as in Theorem 4.9. Then $P_A \otimes P_B$ is a non-negative diagonal matrix, and

$$\mathrm{tr}(P_A \otimes P_B) = \sum_{i,j} (p_A)_{ii}(p_B)_{jj} = \mathrm{tr}(P_A)\mathrm{tr}(P_B) = 1.$$

Analogously, $Q_A \otimes Q_B$ is a non-negative diagonal matrix and $\mathrm{tr}(Q_A \otimes Q_B) = \mathrm{tr}(Q_A)\mathrm{tr}(Q_B) = 1$. It is also straightforward to verify that

$$(P_A \otimes P_B)^{1/2}(A \otimes B)(Q_A \otimes Q_B)^{1/2} = (P_A^{1/2} A Q_A^{1/2}) \otimes (P_B^{1/2} B Q_B^{1/2}).$$

Let $C = P_A^{1/2} A Q_A^{1/2}$ and $D = P_B^{1/2} B Q_B^{1/2}$. Then, to bound $\|A \otimes B\|_{E\infty}$ from below, it is enough to show $\|C \otimes D\|_{S_1} = \|C\|_{S_1}\|D\|_{S_1}$. Given Lemma 5.1, this is shown by the following simple calculation

$$\begin{aligned}
\|C \otimes D\|_{S_1} &= \mathrm{tr}(\Sigma_C \otimes \Sigma_D) \\
&= \sum_{i,j} \sigma_i(C)\sigma_j(D) \\
&= \left(\sum_i \sigma_i(C)\right)\left(\sum_i \sigma_i(D)\right) \\
&= \|C\|_{S_1}\|D\|_{S_1}.
\end{aligned}$$

The above proves that $\|A \otimes B\|_{E\infty} \geq \|A\|_{E\infty}\|B\|_{E\infty}$, and finished the proof of the theorem. □

5.4 Tight Examples

In this section we give examples for which our bounds in Theorem 4.12 are tight. Both examples are simple and natural. The lower bounds on discrepancy in terms of the ellipsoid infinity norm are tight for the incidence matrix of prefix intervals of $[n]$. The upper bounds are tight for the incidence matrix of a power set.

5.4.1 The Ellipsoid Infinity Norm of Intervals

Let $\mathcal{I}_n$ be the set system of all initial intervals $\{1, 2, \ldots, i\}$, $i = 1, 2, \ldots, n$, of $[n]$. Its incidence matrix is T_n, the $n \times n$ matrix with 0s above the main diagonal and 1s

everywhere else.

It is well known, and easy to see, that $\operatorname{herdisc}(T_n) = \operatorname{herdisc}(\mathcal{I}_n) = 1$. Indeed, any restriction of $\mathcal{I}_n$ is isomorphic to a subset of $\mathcal{I}_n$, and the coloring $\chi(i) = (-1)^{i \bmod 2}$ achieves discrepancy 1. This implies that $\operatorname{hvdisc}(T_n) = 1$, since vector discrepancy is a relaxation of discrepancy. Since the matrices with hereditary discrepancy 1 are exactly the totally unimodular matrices [66], we also have $\operatorname{detlb}(T_n) = 1$.

We will prove that $\|T_n\|_{E\infty}$ is of order $\log n$. This shows that the ellipsoid-infinity norm can be $\log n$ times larger than the hereditary discrepancy, as well as the hereditary vector discrepancy and the determinant lower bound.

Moreover, this example and Theorem 5.1 are the key ingredients in our near-tight lower bound for Tusnády's problem in Chapter 6.

Proposition 5.5. *For T_n the* 0-1 *matrix with* 1*s on the main diagonal and below, we have $\|T_n\|_{E\infty} = \Theta(\log n)$.*

The lower bound in Proposition 5.5 can be proved by relating T_n to a circulant matrix, whose singular values can be estimated using Fourier analysis. Observe that if we put four copies of T_n together in the following way

$$\begin{pmatrix} T_n & T_n^{\intercal} \\ T_n^T & T_n \end{pmatrix},$$

we obtain a circulant matrix, which we denote by $C_{n+1,2n}$; for example, for $n = 3$, we have

$$C_{4,6} = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

We show the following technical lemma, which will then imply Proposition 5.5.

Lemma 5.2. *For natural numbers $s \leq n$, let $c_{s,n}$ be the vector consisting of s ones followed by $n - s$ zeros, and let $C_{s,n}$ be the $n \times n$ circulant matrix whose jth column*

is the cyclic shift of $c_{s,n}$ by $j-1$ positions to the right. Then, for $\frac{n}{s} \geq \frac{4}{3}$, we have $\|C_{s,n}\|_* = \Omega(n \log s)$.

Proof. Let $\omega = e^{-i2\pi/n}$, where $i = \sqrt{-1}$ is the imaginary unit, and let us write $C := C_{s,n}$ and $c := c_{s,n}$. It is well known that the eigenvalues of a circulant matrix with first column c are the Fourier coefficients $\hat{c}_0, \ldots, \hat{c}_{n-1}$ of c:

$$\hat{c}_j = \sum_{k=0}^{s-1} \omega^{jk} = \frac{\omega^{js} - 1}{\omega^j - 1}.$$

Since C is a normal matrix (i.e., $C^T C = CC^T$), its singular values are equal to the absolute values of its eigenvalues. Therefore, $\|C\|_* = \sum_{j=0}^{n-1} |\hat{c}_j|$, so we need to bound this sum from below. The sum can be estimated analogously to the well-known estimate of the L_1 norm of the Dirichlet kernel. We give the details of the computation next.

To give a lower bound for $|\hat{c}_j|$ (for appropriately chosen values of j), we give a lower bound for $|\omega^{js} - 1|$ and an upper bound for $|\omega^j - 1|$. Let $\{x\}$ be the fractional part of a real number x. If $\frac{1}{8} \leq \{\frac{js}{n}\} \leq \frac{7}{8}$, then $\Re(\omega^{js}) \leq \frac{\sqrt{2}}{2}$, and, therefore, $|\omega^{js} - 1| \geq \frac{2-\sqrt{2}}{2}$. So, we have the implication

$$\frac{1}{8} \leq \left\{\frac{js}{n}\right\} \leq \frac{7}{8} \implies |\hat{c}_j| \geq \frac{2-\sqrt{2}}{2|\omega^j - 1|}. \tag{5.1}$$

We have $\omega^j = \cos\left(\frac{2\pi j}{n}\right) - i \sin\left(\frac{2\pi j}{n}\right)$, and, therefore,

$$\begin{aligned} |\omega^j - 1|^2 &= \left(\cos\left(\frac{2\pi j}{n}\right) - 1\right)^2 + \sin\left(\frac{2\pi j}{n}\right)^2 \\ &= 2\left(1 - \cos\left(\frac{2\pi j}{n}\right)\right). \end{aligned}$$

From the Taylor approximation of the cosine function, for $-\pi/2 \leq \phi \leq \pi/2$, $1 - \cos(\phi) \leq \phi^2/2$. Therefore,

$$|\omega^j - 1| \leq \begin{cases} \frac{2\pi j}{n}, & 0 \leq j \leq \frac{n}{4} \\ \frac{2\pi(n-j)}{n}, & \frac{3n}{4} \leq j \leq n-1 \end{cases}. \tag{5.2}$$

Let S be the set of integers j such that $\frac{1}{8} \leq \left\{\frac{js}{n}\right\} \leq \frac{7}{8}$. Let further $S_1 := S \cap [0, \frac{n}{4}]$ and $S_2 = S \cap [\frac{3n}{4}, n-1]$. By (5.1) and (5.2),

$$\sum_{j=1}^{n} |\hat{c}_i| \geq \frac{(2-\sqrt{2})n}{2\pi} \left(\sum_{j \in S_1} \frac{1}{j} + \sum_{j \in S_2} \frac{1}{n-j}\right).$$

Notice that S is the union of disjoint intervals of size at least $\lfloor \frac{3n}{4s} \rfloor \geq 1$, separated by intervals of size at most $\lceil \frac{n}{4s} \rceil$. Therefore, for any interval $[a, b)$ where $b - a \geq \lceil \frac{n}{s} \rceil + 1$, $|S \cap [a, b)| = \Omega(b - a)$. We have the estimate

$$\begin{aligned}\sum_{j \in S_1} \frac{1}{j} &\geq \sum_{t=\lceil \log_2(n/s) \rceil}^{\lfloor \log_2(n/4) \rfloor - 1} \frac{1}{2^t} \cdot |S \cap [2^t, 2^{t+1})| \\ &= \Omega(\log(n) - \log(n/s)) = \Omega(\log s).\end{aligned}$$

An analogous argument shows that $\sum_{j \in S_2} \frac{1}{n-j} = \Omega(\log s)$, and this completes the proof. □

Proof of Proposition 5.5. The upper bound $\|T_n\|_{E\infty} = O(\log n)$ can be directly proved in a number of ways. Here we simply observe that it follows from Theorem 4.12 and $\operatorname{herdisc}(A) = 1$.

For the lower bound, we can take $P = Q = \frac{1}{n} I$ in Theorem 4.9, and then it suffices to show $\|T_n\|_{S_1} = \Omega(n \log n)$. We prove this by relating $\|T_n\|_{S_1}$ to $\|C_{n+1,2n}\|_{S_1}$: since the nuclear norm is invariant under adding zero rows and columns and transposition, by the triangle inequality we have $\|C_{n+1,2n}\|_{S_1} \leq 4\|T_n\|_{S_1}$. The proposition is then proved by Lemma 5.2. □

We remark that the singular values of T_n are in fact exactly known and are equal to

$$\frac{1}{2 \sin \frac{(2j-1)\pi}{4n+2}},$$

for $j \in [n]$. One way to show this is to observe that the inverse of $T_n T_n^T$ is the matrix of the second difference operator with 1 in the lower right corner:

$$\begin{pmatrix} 2 & -1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{pmatrix}.$$

The singular values of this matrix can be computed by deriving a recurrence for the characteristic polynomial, and observing that it can be as the difference of Chebyshev polynomials of the second kind.

5.4.2 The Ellipsoid Infinity Norm of Power Sets

It is easy to find matrices A for which $\mathrm{detlb}(A) = \mathrm{hvdisc}(A) = \|A\|_{E\infty}$: for example, take the identity matrix or a Hadamard matrix. The next proposition gives a matrix A for which $\mathrm{herdisc}(A) \geq \frac{1}{2}\sqrt{\log_2(m+1)}\|A\|_{E\infty}$.

Proposition 5.6. *Let A be the incidence matrix of the power set $2^{[n]}$, not including the empty set. Then* $\mathrm{herdisc}(A) \geq \frac{n}{2}$*, while* $\|A\|_{E\infty} \leq \sqrt{n}$*.*

Proof. By Proposition 5.1, $\|A\|_{E\infty} = \|A^\intercal\|_{E\infty}$. But every column of A has ℓ_2 norm at most $\sqrt{n}$, so $\|A^\intercal\|_{E\infty} \leq \|\sqrt{n} \cdot B_2^n\| = \sqrt{n}$. On the other hand, for any coloring $x \in \{-1,1\}^n$, one of the sets $\{i : x_i = 1\}$ and $\{i : x_i = -1\}$ has cardinality at least $\frac{n}{2}$. Let the row a of A be the indicator vector of this set; then $|\langle a, x\rangle| \geq \frac{n}{2}$. □

While Propositions 5.5 and 5.6 imply that our analysis of the ellipsoid infinity norm is tight, there are natural strengthenings of this quantity for a better approximation guarantee is plausible. For example, the argument that we used to prove Lemma 4.4 also proves the inequality

$$\mathrm{herdisc}(A) \leq C \max\{\mathbb{E}_g\|Fg\|_\infty : a_1, \ldots a_n \in FB_1^m\}. \tag{5.3}$$

for C an absolute constant, g a standard gaussian in $\mathbb{R}^m$, and A the $m \times n$ matrix $A = (a_i)_{i=1}^n$. It is also straightforward to see that the right hand side is bounded from above by $O(\sqrt{\log m})\|A\|_{E\infty}$ (by the Chernoff bound) and from below by $\Omega(1)\|A\|_{E\infty}$ (by the Jensen inequality). We leave the following question open.

Question 3. *Is the right hand side of* (5.3) *efficiently computable? Is it approximable up to a fixed constant? Does it provide an assymptotically tighter approximation to* $\mathrm{herdisc}(A)$ *than* $\|A\|_{E\infty}$*?*

We remark that any quantity that satisfies Proposition 5.3 (even if the right hand side is multiplied by a constant) cannot approximate hereditary discrepancy better than a factor of $\Omega(\log n)$. The reason is that there exist examples of a pair of set systems $\mathcal{S}_1, \mathcal{S}_2$ each of hereditary discrepancy 1, whose union $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ has hereditary discrepancy $\Omega(\log n)$ [113]. The proof of Proposition 5.3 can be adapted for the quantity in (5.3), so the best approximation we can hope for is $O(\log n)$. These observations suggest the natural conjecture that hereditary discrepancy may be hard to approximate within $o(\log n)$.

Bibliographic Remarks

The results in this chapter come from joint work with Jiří Matoušek. A preliminary version is available as [103]. The Fourier analytic proof of Proposition 5.5 was discovered independently and appears in full for the first time in this thesis.

Chapter 6

Applications to Discrepancy Theory

6.1 Overview

The properties of the ellipsoid-infinity norm established in Chapter 5 can be seen as "composition theorems". I.e. they show us how we can deduce upper and lower bounds on $\|A\|_{E\infty}$ for some matrix A by decomposing A into simpler matrices for which the ellipsoid-infinity norm can be analyzed directly. Since $\|A\|_{E\infty}$ approximates hereditary discrepancy, in effect we show how to estimate the hereditary discrepancy of a matrix (or set system) in terms of the discrepancies of simpler matrices (resp. set systems). In this chapter we provide a number of examples of this technique. Most prominently, we show a near-tight lower bound on the discrepancy of axis-aligned boxes in d dimensions. This nearly settles the high-dimensional version of the Tusnády problem.

6.2 General Results for Discrepancy

The following "composition results" for discrepancy are immediate consequences of the properties established in Chapter 5 and Theorem 4.12.

To state the first general result, let us recall the definition of the *dual set system*.

Definition 6.1. *For a set system $(\mathcal{S}, U)$, the* dual set system $(\mathcal{S}^*, \mathcal{S})$ *has a set S_e^* for each $e \in U$, defined as $S_e^* = \{S \in \mathcal{S} : e \in S\}$.*

An easy observation is that if the incidence matrix of $\mathcal{S}$ is A, then $A^\intercal$ is the incidence matrix of $\mathcal{S}^*$. It follows from Theorem 4.12 and Proposition 5.1 that the hereditary discrepancy of $\mathcal{S}$ and $\mathcal{S}^*$ are the same, up to polylogarithmic factors. We state the result next.

Theorem 6.1. *For any set system $\mathcal{S}$ of m sets on a universe of size n, and its dual $\mathcal{S}^*$,*

$$\operatorname{herdisc}(\mathcal{S}) = O(\log^{3/2} mn) \cdot \operatorname{herdisc}(\mathcal{S}^*).$$

This theorem also follows from Matoušek's result that the determinant lower bound is nearly tight [106]. There exist set systems for which $\operatorname{disc}(\mathcal{S}) = \Omega(\log n) \cdot \operatorname{herdisc}(\mathcal{S}^*)$. For example, for a permutation π of $[n]$, let $\mathcal{I}_\pi$ be the set system of initial intervals $\{\pi_1, \ldots, \pi(i)\}$ for all $i \in [n]$. There exist three permutations π_1, π_2, π_3 such that $\mathcal{S} \triangleq \mathcal{I}_{\pi_1} \cup \mathcal{I}_{\pi_2} \cup \mathcal{I}_{\pi_3}$ has discrepancy $\Omega(\log n)$ [113]. On the other hand, each set in $\mathcal{S}^*$ is isomorphic to the disjoint union of three initial segments of the form $\{1, \ldots, i\}$. Since the set system of initial segments $\mathcal{I}$ has hereditary discrepancy 1 (see Section 5.4.1), it follows that $\mathcal{S}^*$ has hereditary discrepancy at most 3.

The following theorem about unions of set systems was also proved by Matoušek via the determinant lower bound [106]. Here we give a different proof based on the ellipsoid infinity norm.

Theorem 6.2. *Let $\mathcal{S} = \bigcup_{i=1}^k \mathcal{S}_i$, where $\mathcal{S}$ and $\mathcal{S}_1, \ldots, \mathcal{S}_k$ are set systems on the same universe, and $|\mathcal{S}| = m$. Then*

$$\operatorname{herdisc}(\mathcal{S}) \leq O((\log m)^{3/2}) \cdot \sqrt{k} \max_{i=1}^{k} \operatorname{herdisc}(\mathcal{S}_i).$$

Proof. Follows immediately from Theorem 4.12 and Proposition 5.3 applied to the incidence matrices A of $\mathcal{S}$ and A_i of $\mathcal{S}_i$, and the observation that $|\mathcal{S}_i| \leq |\mathcal{S}|$ for all i. □

There is an example of two set systems $\mathcal{S}_1$ and $\mathcal{S}_2$ of hereditary discrepancy 1, such that $\mathcal{S}_1 \cup \mathcal{S}_2$ has discrepancy $\Omega(\log n)$. In fact we can use essentially the same example as the one mentioned above for dual set systems. Let, again π_1, π_2, π_3 be three permutations such that the discrepancy of $\mathcal{S} \triangleq \mathcal{I}_{\pi_1} \cup \mathcal{I}_{\pi_2} \cup \mathcal{I}_{\pi_3}$ is $\Omega(\log n)$, and define $\mathcal{S}_1 \triangleq \mathcal{I}_{\pi_1} \cup \mathcal{I}_{\pi_2}$ and $\mathcal{S}_2 \triangleq \mathcal{I}_{\pi_3}$. The hereditary discrepancy of $\mathcal{I}_{\pi_3}$ is equal to the hereditary discrepancy of the set system $\mathcal{I}$ of initial segments, since the two set systems are isomorphic. As already observed, the hereditary discrepancy of $\mathcal{I}$ is 1. The hereditary discrepancy of $\mathcal{I}_{\pi_1} \cup \mathcal{I}_{\pi_2}$ is 1 for *any* two permutations π_1 and π_2; this was first observed by Beck; see e.g. [136, Lecture 5].

The next general result we state follows from the triangle inequality for the ellipsoid-infinity norm.

Theorem 6.3. *Suppose there exist set systems $\mathcal{S}_1, \ldots, \mathcal{S}_k$ and $\mathcal{T}_1, \ldots, \mathcal{T}_k$ such that each set S in a set system $\mathcal{S}$ can be written as*

$$S = ((S_1 \setminus T_1) \cup S_2) \setminus T_2) \ldots) \cup S_k) \setminus T_k,$$

where $S_i \in \mathcal{S}_i$, $T_i \in \mathcal{T}_i$, each set union is on disjoint sets, and each set difference removes a set from a set that contains it. Then

$$\operatorname{herdisc}(\mathcal{S}) \leq O(\log^{3/2} m) \cdot \sum_{i=1}^{k} (\operatorname{herdisc}(\mathcal{S}_i) + \operatorname{herdisc}(\mathcal{T}_i)),$$

where $m = |\mathcal{S}|$.

Proof. Observe that, since at most one set from each $\mathcal{S}_i$ and $\mathcal{T}_i$ is used in the expression for each $S \in \mathcal{S}$, we can assume that $|\mathcal{S}_i|, |\mathcal{T}_i| \leq m$, for any i. Then, after re-arranging and possibly duplicating or removing some sets, we can write

$$A = A_1 - B_1 + \ldots + A_k - B_k,$$

where A is the incidence matrix of $\mathcal{S}$ and A_i, B_i are the incidence matrices of, respectively, $\mathcal{S}_i$ and $\mathcal{T}_i$. Then the theorem follows from Theorem 4.12, and the triangle inequality in Proposition 5.2. □

The final general result we state gives upper and lower bounds on the hereditary discrepancy of products of set systems. Given two set systems $(\mathcal{S}_1, U_1$ and $(\mathcal{S}_2, U_2)$, the *product set system* $\mathcal{S}_1 \times \mathcal{S}_2$ is a set system on the Cartesian product $U_1 \times U_2$ of the universes, and is defined as

$$\mathcal{S}_1 \times \mathcal{S}_2 \triangleq \{S_1 \times S_2 : S_1 \in \mathcal{S}_1, S_2 \in \mathcal{S}_2\}.$$

It is straightforward to verify that the incidence matrix of $\mathcal{S}_1 \times \mathcal{S}_2$ is the tensor (i.e. Kronecker) product $A_1 \otimes A_2$ of the incidence matrices A_1, A_2 of $\mathcal{S}_1$ and $\mathcal{S}_2$.

Product set systems were considered by Doerr et al. [49], where the authors gave an example in which $\operatorname{disc}(\mathcal{S}_1 \times \mathcal{S}_2) = 0$ for two set systems $\mathcal{S}_1$ and $\mathcal{S}_2$ of nonzero discrepancy.

Therefore, no bound of the form $\operatorname{disc}(\mathcal{S}_1 \times \mathcal{S}_2) \geq \alpha \operatorname{disc}(\mathcal{S}_1) \operatorname{disc}(\mathcal{S}_2)$ is possible in general for $\alpha > 0$. By contrast, using the ellipsoid-infinity norm we can show approximate multiplicativity of hereditary discrepancy with respect to the taking products of set systems. The following theorem is an immediate consequence of Theorems 4.12 and 5.1.

Theorem 6.4. *For any two set systems $\mathcal{S}_1$ and $\mathcal{S}_2$, we have the inequalities*

$$\operatorname{herdisc}(\mathcal{S}_1 \times \mathcal{S}_2) = O(\log^{5/2} m) \cdot \operatorname{herdisc}(\mathcal{S}_1) \operatorname{herdisc}(\mathcal{S}_2),$$

and

$$\operatorname{herdisc}(\mathcal{S}_1 \times \mathcal{S}_2) = \Omega\left(\frac{1}{\log^2 m}\right) \cdot \operatorname{herdisc}(\mathcal{S}_1) \operatorname{herdisc}(\mathcal{S}_2),$$

where $m \triangleq |\mathcal{S}_1| \cdot |\mathcal{S}_2|$.

6.3 Tusnády's Problem

In this section we use the ellipsoid infinity norm to give near tight upper and lower bounds for the higher-dimensional Tusnády's Problem, which asks for the discrepancy of axis-aligned boxes in $\mathbb{R}^d$.

6.3.1 Background

In 1980, Tusnády asked whether every finite set U of points in the plane can be bi-colored so that no axis-aligned rectangle contains more than a constant number of points of one color in excess of the other. Let $\mathcal{B}_2$ be the set of all axis-aligned rectangles $[a_1, b_2) \times [a_2, b2)$ in the plane, and let $\mathcal{B}_2(P) \triangleq \{B \cap P : B \in \mathcal{B}_2\}$, where $P \subset \mathbb{R}^2$ is a finite set of points. Let $\operatorname{disc}(N, \mathcal{B}_2)$ be the maximum of $\operatorname{disc}(\mathcal{B}_2(P))$ over all n-point sets $P \subset \mathbb{R}^2$. Tusnády's problem then asks whether $\operatorname{disc}(N, \mathcal{B}_2)$ remains bounded as N goes to infinity. In 1981, Beck established the transference lemma (Lemma 1.1, which implies that

$$\operatorname{disc}(N, \mathcal{B}_2) = O(1) \cdot D(N, \mathcal{B}_2),$$

where $D(N, \mathcal{B}_2)$ is the Lebesgue-measure discrepancy of N-point sets with respect to axis-aligned rectangles. Classical work by Schmidt [131], improving on a result by Roth [126], shows that $D(N, \mathcal{B}_2) = \Omega(\log N)$, which implies that $\operatorname{disc}(N, \mathcal{B}_2) = \Omega(\log N)$

as well. Beck also showed an upper bound of $O(\log^4 N)$, which has subsequently been improved by Srinivasan [138] to $O(\log^{2.5} N)$.

There is a natural generalization of Tusnády's question to higher dimensions. Let $\mathcal{B}_d$ be the set of all axis-aligned boxes in $\mathbb{R}^d$, i.e. all cross products $[a_1, b_1), \ldots, [a_d, b_d)$, and let $\mathcal{B}_d(P) = \{B \cap P : B \in \mathcal{B}_d\}$ be the set system induced by axis aligned boxes on a finite point set P. The combinatorial discrepancy of axis-aligned boxes in $\mathbb{R}^d$ for size N point sets is denoted $\text{disc}(N, \mathcal{B}_d)$ and is equal to the maximum of $\text{disc}(\mathcal{B}_d(P))$ over sets $P \subset \mathbb{R}^d, |P| = N$. The *quantitative, higher-dimensional* version of Tusnády's problem asks how $\text{disc}(N, \mathcal{B}_d)$ grows as $N \to \infty$. Using the transference lemma, and the results of Roth and Schmidt, Beck showed that for any constant d,

$$\text{disc}(N, \mathcal{B}_d) = \Omega(\max\{\log N, \log^{(d-1)/2} N\}),$$

where the constant hidden in the asymptotic notation depends on d.

The problem of determining the worst-case discrepancy of axis-aligned boxes has been one of the central problems of combinatorial discrepancy theory. After a long line of work, the current best upper bound is $O(\log^{d+0.5} N)$, due to Larsen [89], using a deep result of Banaszczyk. A slightly weaker bound using Beck's partial coloring lemma [20] was proved previously by Matoušek [102]. Nevertheless, the best known lower bound has remained the one due to Beck cited above, which uses lower bounds on Lebesgue-measure discrepancy. Closing the significant gap between upper and lower bounds has been an open problem dating at least as far back as the foundational work of Beck from 1981.

6.3.2 Tight Upper and Lower Bounds

In this paper, we nearly resolve Tusn'ady's problem for any constant dimension d, using an argument that applies directly to combinatorial discrepancy.

Theorem 6.5. *The discrepancy of axis-aligned boxes in* $\mathbb{R}^d$ *is bounded as*

$$\frac{(\log N)^{d-1}}{(Cd)^d} \leq \text{disc}(N, \mathcal{B}_d) \leq (C \log N)^{d+1/2},$$

for a large enough constant C*.*

Our proof of the upper bound is somewhat different, and arguably simpler, than the one given by Larsen.

Proof of Theorem 6.5. Let $\mathcal{A}_d \subseteq \mathcal{B}_d$ be the set of all *anchored* axis-parallel boxes, i.e. boxes of the form $[0, b_1) \times \cdots \times [0, b_d)$. Clearly $\operatorname{disc}(n, \mathcal{A}_d) \leq \operatorname{disc}(n, \mathcal{B}_d)$, and since every box $R \in \mathcal{B}_d$ can be expressed as a signed combination of at most 2^d anchored boxes, we have $\operatorname{disc}(N, \mathcal{B}_d) \leq 2^d \operatorname{disc}(n, \mathcal{A}_d)$.

Let us consider the d-dimensional grid $[n]^d \subset \mathbb{R}^d$ (with $N \triangleq n^d$ points), and let $\mathcal{G}_{d,n} = \mathcal{A}_d([n]^d)$ be the subsets induced on it by anchored boxes. It suffices to prove that $\operatorname{herdisc}(\mathcal{G}_{d,n}) \geq \frac{(c_1 \log n)^d}{C_1 \log N}$, for absolute constants c_1, C_1. For this, in view of Theorem 4.12, it is enough to show that $\|\mathcal{G}_{d,n}\|_{E\infty} \geq (c_1 \log n)^d$.

Now $\mathcal{G}_{d,n}$ is (isomorphic to) the d-fold product $\mathcal{I}_n^d$ of the system of initial segments in $\{1, 2, \ldots, n\}$, and so has incidence matrix $T_n^{\otimes d}$, where T_n is the lower triangular binary matrix with 1s on and below the main diagonal. Then, by Theorem 5.1 and Proposition 5.5,

$$\|\mathcal{G}_{d,n}\|_{E\infty} = \|T_n^{\otimes d}\|_{E\infty} = \|T_n\|_{E\infty}^d \geq (c_1 \log n)^d.$$

This finishes the proof of the lower bound. To prove the upper bound, we consider an arbitrary N-point set $P \subset \mathbb{R}^d$. Since the set system $\mathcal{A}_d(P)$ is not changed by a monotone transformation of each of the coordinates, we may assume $P \subseteq [N]^d$, and $\mathcal{A}_d(P)$ is a subset of at most N sets of $\mathcal{G}_{d,N}$. Hence, by Theorems 4.12 and Theorem 5.1, and Proposition 5.5, there exist constants C_2 and C_3 such that

$$\operatorname{disc}(\mathcal{A}_d(P)) \leq C_2 \sqrt{\log N} \cdot \|\mathcal{G}_{d,N}\|_{E\infty} \leq C_2 \sqrt{\log N} (C_3 \log N)^d.$$

Taking C suitable large with respect to $\frac{1}{c_1}, C_1, C_2, C_3$ finishes the proof. □

6.4 Discrepancy of Boolean Subcubes

Chazelle and Lvov [39, 38] investigated the hereditary discrepancy of the set system $\mathcal{C}_d := \mathcal{B}_d(\{0,1\}^d)$, i.e. the set system induced by axis-parallel boxes on the d-dimensional Boolean cube $\{0,1\}^d$. In other words, the sets in $\mathcal{C}_d$ are subcubes of $\{0,1\}^d$. We can

specify each set $C_v \in \mathcal{C}_d$ by a vector $v \in \{0,1,*\}$, as

$$C_v = \{u \in \{0,1\}^d : v_i \neq * \Rightarrow u_i = v_i\}.$$

Unlike for Tusnády's problem where d was considered fixed, here one is interested in the asymptotic behavior as $d \to \infty$.

Chazelle and Lvov proved $\operatorname{herdisc} \mathcal{C}_d = \Omega(2^{cd})$ for an absolute constant $c \approx 0.0477$, which was later improved to $c = 0.0625$ in [117] (in relation to the hereditary discrepancy of homogeneous arithmetic progressions). Here we obtain an optimal value of the constant c:

Theorem 6.6. *The system $\mathcal{C}_d$ of subcubes of the d-dimensional Boolean cube satisfies*

$$\operatorname{herdisc}(\mathcal{C}_d) = 2^{c_0 d + o(d)},$$

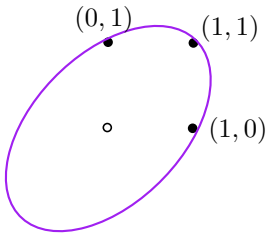
where $c_0 = \log_2(2/\sqrt{3}) \approx 0.2075$. The same bound holds for the system $\mathcal{A}_d(\{0,1\}^d)$ of all subsets of the cube induced by anchored boxes.

Proof. The number of sets in $\mathcal{C}_d$ is 3^d, and so, by Theorem 4.12, it suffices to prove $\|\mathcal{C}_d\|_{E\infty} = \|\mathcal{A}_d(\{0,1\}^d)\|_{E\infty} = 2^{c_0 d}$.

The system $\mathcal{C}_d$ is the d-fold product $\mathcal{C}_1^d$, and so by Theorem 5.1, $\|\mathcal{C}_d\|_{E\infty} = \|\mathcal{C}_1\|_{E\infty}^d$. The incidence matrix of $\mathcal{C}_1$ is

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

To get an upper bound on $\|A\|_{E\infty}$, we exhibit an appropriate ellipsoid; it is more convenient to do it for A^T, since this is a planar problem. The optimal ellipse containing the rows of A is $\{x \in \mathbb{R}^2 : x_1^2 + x_2^2 - x_1 x_2 \leq 1\}$; here is the ellipsoid and the dual matrix:



$$D = \begin{pmatrix} \frac{4}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{4}{3} \end{pmatrix}.$$

Hence $\|A\|_{E\infty} \leq 2/\sqrt{3}$. The same ellipse also works for the incidence matrix of the system $\mathcal{A}_1(\{0,1\})$, which is the familiar lower triangular matrix T_2.

There are several ways of bounding $\|T_2\|_{E\infty} \leq \|A\|_{E\infty}$ from below. For example, we can use Theorem 4.9 with

$$P = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & \frac{2}{3} \end{pmatrix}, \quad Q = \begin{pmatrix} \frac{2}{3} & 0 \\ 0 & \frac{1}{3} \end{pmatrix}.$$

One can compute the characteristic polynomial of $P^{1/2}T_2Q^{1/2}$ and check that the singular values are $\frac{1}{\sqrt{3}} \pm \frac{1}{3}$, and hence the nuclear norm is $2/\sqrt{3}$ as needed.

Alternatively, one can also check the optimality of the ellipse above by elementary geometry. □

6.5 Discrepancy of Arithmetic Progressions

In this section we prove several results on the discrepancy of arithmetic progressions. Irregularities of distribution with respect to arithmetic progressions have been the focus of a long line of research dating back at least as far as van der Waerden's famous theorem from 1927. Van der Waerden's theorem implies that for any k, any n large enough with respect to k, and any coloring $\chi : [n] \to \{-1, 1\}$, there exist an arithmetic progression on $[n]$ of size k which is monochromatic with respect to χ. This is one extreme case of imbalanced arithmetic progressions with respect to colorings: we look for short arithmetic progressions (with respect to n) with maximum discrepancy. At the other end is the problem of analyzing imbalances in long arithmetic progressions. This direction was started by the beautiful work of Roth [127]. Roth's 1/4-theorem shows that the discrepancy of the set system $\mathcal{AP}_n$ of all arithmetic progressions on $[n]$ is $\text{disc}(\mathcal{AP}_n) = \Omega(n^{1/4})$. After pioneering work by Beck [20] and later improvements, Matoušek and Spencer [104] showed that Roth's lower bound is the best possible up to constants.

We complement these classic results and show that $\|\mathcal{AP}_n\|_{E\infty} = \Theta(n^{1/4})$. This implies, via Theorem 4.12, a discrepancy upper bound that is worse than Matoušek and Spencer's by a factor of $O(\sqrt{\log n})$. Arguably, our proof is simpler. Via Theorem 5.1, we also get upper and lower bounds on the hereditary discrepancy of multidimensional arithmetic progressions, which generalize (at the cost of slightly suboptimal bounds)

results of Doerr et al. [49].

We also consider a subset of $\mathcal{AP}_n$: the set system $\mathcal{HAP}_n$ of all homogeneous arithmetic progressions on $[n]$, i.e. all arithmetic progressions of the type $\{a, 2a, \ldots, ka\}$ for $a \leq n$ and $k \leq \lfloor n/a \rfloor$. Circa 1932, Erdős asked whether $\mathrm{disc}(\mathcal{HAP}_n) = \omega(1)$. This problem is now known as the Erdős Discrepancy Problem, and stands as a major open problem in discrepancy theory and combinatorial number theory. Much better discrepancy is possible than for general arithmetic progressions: the coloring χ which takes value $\chi(i) = -1$ if and only if the last nonzero digit of i in ternary representation is 2 has discrepancy $O(\log n)$. As far as lower bounds are concerned, Konev and Lisitsa recently reported [87] that the discrepancy of $\mathcal{HAP}_n$ is at least 3 for large enough n, and this remains the best known lower bound (a lower bound of 2 for $n \geq 12$ was well known).

Via a reduction from the discrepancy of Boolean subcubes we show that the *hereditary* discrepancy of $\mathcal{HAP}_n$ is at least $n^{1/O(\log \log n)}$. This is tight up to the constant in the exponent, as shown by Alon and Kalai [83]. For completeness, we reproduce Alon and Kalai's argument, with a slightly different proof using the ellipsoid infinity norm.

In relation to the above mentioned results, it is worth mentioning that arithmetic progressions are not hereditary, i.e. a restriction of an arithmetic progression on $[n]$ to some $W \subset [n]$ is not necessarily an arithmetic progression. This makes the Erdős discrepancy problem significantly more challenging than the hereditary discrepancy question that we essentially resolve.

6.5.1 General Arithmetic Progressions

We prove the following proposition.

Proposition 6.1. $\|\mathcal{AP}\|_{E\infty} = \Theta(n^{1/4})$.

Before embarking on a proof, let us recall a basic tool in algorithms and combinatorics: the *canonical (dyadic) intervals trick.*

Definition 6.2. *A* canonical interval *is an interval of the form* $(a2^i, (a+1)2^i] \cap \mathbb{N}$*, where* a *and* i *are non-negative integers.*

The following lemma is easy, well-known, and remarkably useful.

Lemma 6.1. *Any initial interval $\{1, \ldots, j\}$, can be written as the disjoint union of at most $1 + \lceil \log_2 j \rceil$ canonical intervals, each of different size.*

Proof. We prove the lemma by induction on j. The base case $j = 1$ is trivial, since $\{1\}$ is a canonical interval. For the inductive step, assume the lemma is true for all $k \leq j-1$; we will prove that the lemma holds for j under this assumption. Let $i = \lfloor \log_2 j \rfloor$. We can write $\{1, \ldots, j\} = \{1, \ldots, 2^i\} \cup \{2^i+1, \ldots, j\}$. Notice that the first set is a canonical interval (for $a = 0$ and i as chosen above). The second set is an interval of size less than 2^i, since i was chosen maximal so that $2^i \leq j$; it follows that $j - 2^i < j/2$. By shifting the integers $\{2^i + 1, \ldots, j\}$ left by 2^i, using the inductive hypothesis, and then shifting right by 2^i, we have that $\{2^i + 1, \ldots, j\}$ can be written as the disjoint union of at most $1 + \lceil \log_2(j - 2^i) \rceil < \lceil \log_2 j \rceil$ canonical intervals, all of different sizes. Moreover, all these intervals must have sizes less than 2^i. This finishes the inductive step. □

With this basic tool in hand, we are ready to prove the proposition.

Proof of Proposition 6.1. The lower bound $\|\mathcal{AP}_n\|_{E\infty} = \Omega(n^{1/4})$ is implied by Lovász's proof of Roths's 1/4-theorem via semidefinite programming [94]. Lovász showed that $\operatorname{vecdisc}(\mathcal{AP}_n) = \Omega(n^{1/4})$. By Theorem 4.12, $\|\mathcal{AP}_n\|_{E\infty} \geq \operatorname{hvdisc} \mathcal{AP}_n = \Omega(n^{1/4})$.

Next, we prove the upper bound. For an interval $I \subseteq [n]$, let $\mathcal{M}$ be the set of all inclusion-maximal arithmetic progressions in $[n]$. We claim that

$$\|\mathcal{M}|_I\|_{E\infty} \leq \sqrt{2}|I|^{1/4}. \tag{6.1}$$

where $|I| \triangleq |b - a|$ is the *size* of the interval $I = [a, b)$.

Before proving (6.1), let us see why it implies $\|\mathcal{AP}\|_{E\infty} = O(n^{1/4})$. Let $\mathcal{M}_i$ be the union of the set systems $\mathcal{M}|_I$ over all canonical intervals I of size 2^i. Since $\mathcal{M}_i$ is a union of set systems with disjoint supports, by Proposition 5.4 and (6.1) $\|\mathcal{M}_i\|_{E\infty} \leq 2^{\frac{i}{4}+\frac{1}{2}}$.

Every arithmetic progression A on $[n]$ can be written as $M \cap J$, where M is a maximal arithmetic progression and J is an interval in $[n]$. J can be written as the set difference of two nested initial intervals $J_1 \subset J_2$. By Lemma 6.1, J_1 and J_2 can each be

written as the disjoint union of canonical intervals of different sizes. Intersecting each of these canonical intervals with M, we have that

$$A = (M_0 \cup \ldots \cup M_k) \setminus (M'_0 \cup \ldots \cup M'_k),$$

where $k \leq 1 + \lceil \log_2 n \rceil$, $M_i, M'_i \in \mathcal{M}_i \cup \{\emptyset\}$, all set unions are disjoint, and $M'_0 \cup \ldots \cup M'_k \subset M_0 \cup \ldots \cup M_k$. The triangle inequality in Proposition 5.2 then gives $\|\mathcal{AP}\|_{E\infty} \leq \sum_{i=0}^{k} 2 \cdot 2^{\frac{i}{4}+\frac{1}{2}} = O(n^{1/4})$.

It remains to prove (6.1). Let us split $\mathcal{M}|_I$ as $\mathcal{M}' \cup \mathcal{M}''$, where the arithmetic progressions in $\mathcal{M}'$ have difference at most $|I|^{1/2}$, and those in $\mathcal{M}''$ have difference larger than $|I|^{1/2}$. By Proposition 5.3, $\|\mathcal{M}|_I\|_{E\infty} \leq \sqrt{2}\max\{\|\mathcal{M}'\|_{E\infty}, \|\mathcal{M}''\|_{E\infty}\}$, so it suffices to show that $\|\mathcal{M}'\|_{E\infty}, \|\mathcal{M}''\|_{E\infty} \leq |I|^{1/4}$.

Given a difference d, each $c \in I$ belongs to exactly one maximal arithmetic progression with difference d, because such an arithmetic progression is entirely determined by the congruence class of $c \bmod d$. Therefore, each integer in I belongs to at most $|I|^{1/2}$ arithmetic progressions in $\mathcal{M}'$, i.e. $\Delta_{\mathcal{M}'} \leq |I|^{1/2}$. It follows that each column of the incidence matrix of $\mathcal{M}'$ has ℓ_2 norm at most $|I|^{1/4}$, and, by the definition of the ellipsoid infinity norm, $\|\mathcal{M}'\|_{E\infty} \leq |I|^{1/4}$.

On the other hand, every arithmetic progression in $\mathcal{M}''$ has size at most $|I|^{1/2}$, so each row of the incidence matrix of $\mathcal{M}''$ has ℓ_2 norm at most $|I|^{1/4}$. Then, by Proposition 5.1 and the definition of the ellipsoid-infinity norm, we have $\|\mathcal{M}''\|_{E\infty} \leq |I|^{1/4}$, as desired. This implies $\|\mathcal{M}|_I\|_{E\infty} \leq \sqrt{2}|I|^{1/4}$, as we argued above, and finishes the proof. □

Proposition 6.1 and Theorem 4.12 imply $\operatorname{herdisc}(\mathcal{AP}_n) = O(n^{1/4}\sqrt{\log n})$, which is a factor $O(\sqrt{\log n})$ larger than the optimal bound.

6.5.2 Multidimensional Arithmetic Progressions

Doerr, Srivastav, and Wehr [49] considered the discrepancy of the system $\mathcal{AP}^d$ of *d-dimensional arithmetic progressions* in $[n]^d$, which are d-fold Cartesian products of arithmetic progressions. They showed that $\operatorname{disc} \mathcal{AP}^d = \Theta(n^{d/4})$.

Their upper bound was proved by a simple product coloring argument, which does not apply to hereditary discrepancy (since the restriction of $\mathcal{AP}^d$ to a subset of $[n]^d$ no longer has the structure of multidimensional arithmetic progressions). By Proposition 6.1 and Theorem 5.1, we have $\|\mathcal{AP}^d\|_{E\infty} = \Theta(n^{d/4})$ for any constant d, and we thus obtain the (probably suboptimal) upper bound $\text{herdisc}(\mathcal{AP}^d) = O(n^{d/4}\sqrt{\log n})$ by Theorem 4.12.

6.5.3 Homogeneous Arithmetic Progressions

In this subsection we characterize the hereditary discrepancy of homogeneous arithmetic progressions.

Theorem 6.7. *We have* $\text{herdisc}(\mathcal{HAP}_n) = n^{1/\Theta(\log\log n)}$.

We first prove the lower bound. The upper bound in was proved by Alon and Kalai; we reproduce a version of their argument at the end of the section.

Proof of the lower bound in Theorem 6.7. For each positive integer d, we will construct a set of integers J_d such that the hereditary discrepancy of homogeneous arithmetic progressions restricted to J_d is lower bounded by the hereditary discrepancy of $\mathcal{C}_d$. Then the lower bound in Theorem 6.7 will follow from Theorem 6.6.

Let $p_{1,0} < p_{1,1} < \ldots < p_{d,0} < p_{d,1}$ be the first $2d$ primes. We define J_d to be the following set of square free integers

$$J_d = \{\prod_{i=1}^{d} p_{i,u_i} : u \in \{0,1\}^d\}.$$

In other words, J_d is the set of all integers that are divisible by exactly one prime $p_{i,b}$ from each pair $(p_{i,0}, p_{i,1})$ and no other primes. By the prime number theorem[1], the largest of these primes satisfies $p_{d,1} = \Theta(d \log d)$. Let $n = n(d)$ be the largest integer in J_d. The crude bound $n(d) = 2^{O(d \log d)}$ will suffice for our purposes. Notice that $d = \Omega(\log n / \log\log n)$.

There is a natural one to one correspondence between the set J_d and the set $\{0,1\}^d$: to each $u \in \{0,1\}^d$ we associate the integer $j_u = \prod_{i=1}^{d} p_{i,u_i}$. By this correspondence, we

[1]Chebyshev's assymptotic estimate suffices.

can think of any coloring $\chi : \{0,1\}^d \to \{-1,+1\}$ as a colorings $\chi : J_d \to \{-1,+1\}$. We also claim that each set in the set system $\mathcal{C}_d$ corresponds to a homogeneous arithmetic progression restricted to J_d. With any $C_v \in \mathcal{C}_d$ (where $v \in \{0,1,*\}^d$) associate the integer $a_v = \prod_{i:v_i \neq *} p_{i,v_i}$. Observe that for any $j_u \in B_d$, a_v divides j_u if and only if $u \in C_v$. We have the following implication for any coloring χ, any $U \subseteq \{0,1\}^d$, and the corresponding $J = \{j_u : u \in U\}$:

$$\exists C_v \in \mathcal{C}_d : |\sum_{u \in C_v \cap U} \chi(u)| \geq D \quad \Leftrightarrow \quad \exists a \in \mathbb{N} : |\sum_{\substack{j \in J \\ a|j}} \chi(j)| \geq D. \tag{6.2}$$

Notice again that we treat χ as a coloring both of the elements of $\{0,1\}^d$ and of the integers in J_d by the correspondence $u \leftrightarrow b_u$. Theorem 6.6 guarantees the existence of some U such that the left hand side of (6.2) is satisfied with $D = 2^{\Omega(d)} = n^{1/O(\log\log n)}$ for any χ. The lower bound in Theorem 6.7 follows from the right hand side of (6.2). □

For completeness, we also give a version of Alon and Kalai's upper bound argument.

Proof of the upper bound in Theorem 6.7. We will represent each set in $\mathcal{HAP}_n$ as the sum of a logarithmic number of sets from small-degree set systems. The ideas here are similar to the ones used in the proof of the upper bound in Proposition 6.1, and canonical intervals will make an appearance again.

Observe that, by Theorem 4.12, it is sufficient to show that $\|\mathcal{HAP}_n\|_{E\infty} \leq n^{C/\log\log n}$ for an absolute constant C. Let $\mathcal{M}$ be the set of all inclusion-maximal *homogeneous* arithmetic progressions on $[n]$. We claim that,

$$\|\mathcal{M}\|_{E\infty} \leq \sqrt{\Delta_{\mathcal{M}}} \leq n^{C_0/2\log\log n}.$$

The first inequality is by the definition of $\|\mathcal{M}\|_{E\infty}$. For the second inequality, observe that the degree $\Delta_{\mathcal{M}}(j)$ for any $j \in [n]$ is equal to the number $d(j)$ of distinct integer divisors of j. It is well-known that $d(j) \leq n^{C_0/\log\log n}$.

Let us define $\mathcal{M}_i$, for $i \in \{0,\ldots,\lceil\log_2 n\rceil\}$, as the union of the restrictions $\bigcup_I \mathcal{M}|_I$, where I runs over canonical intervals of size 2^i. Since the ellipsoid-infinity norm is non-increasing under restrictions, $\|\mathcal{M}|_I\|_{E\infty} \leq \|\mathcal{M}\|_{E\infty} \leq n^{C_0/2\log\log n}$; moreover, $\mathcal{M}_i$ is the union of disjoint restrictions, and so, by Proposition 5.4, $\|\mathcal{M}_i\|_{E\infty} \leq n^{C_0/2\log\log n}$. Each

set H in $\mathcal{HAP}_n$ can be written as $H \cap J$ for some initial interval J. By Lemma 6.1, J can be written as the disjoint union of canonical intervals of different sizes. Intersecting each of these interval with H, we have that H can be written as the the disjoint union of at most one set from each $\mathcal{M}_i$. Therefore, by the triangle inequality (Proposition 5.2), $\|\mathcal{HAP}_n\|_{E\infty} \leq (1 + \lceil \log_2 n \rceil) n^{C_0/2 \log\log n}$, which is bounded by $n^{C/\log\log n}$ for a large enough constant C. □

Bibliographic Remarks

Together with the material in the previous chapter, the near-tight lower bound for the Tusnády problem, the new proof of the upper bound, and the precise bounds on the discrepancy of Boolean subcubes are the result of joint work with Jiří Matoušek and a preliminary version is available as [103]. A weaker lower bound on the discrepancy of Boolean subcubes was proved via Fourier analysis and the determinant lower bound in [117]. The latter paper also proved the tight (up to the constant in the exponent) lower bound on the hereditary discrepancy of homogeneous arithmetic progressions. The general results on discrepancy, and the analysis of the ellipsoid infinity norm of arithmetic progressions are from [103].

Chapter 7

Discrepancy and Differential Privacy

7.1 Overview

7.1.1 The Central Problem of Private Data Analysis

Datasets containing personal information have become common; moreover, as data collection and storage capacity have improved, such datasets have become richer. Some examples include medical studies, census data, marketing or sociological surveys, friendship or followers data from social networking sites. Performing statistical analysis on these datasets holds significant promise: the discovered knowledge can be useful for the life sciences, for policy making, for marketing. Therefore, many such datasets are of interest to the data mining community. However, analyzing them is limited by concerns that private information about individuals represented in the data may be disclosed. Such disclosure can lead to concrete adverse consequences for an individual, for example an increase in insurance premiums, or discriminatory actions. Moreover, the threat of possible disclosure decreases trust in the organization performing the analysis, and discourages participation in studies, thus hurting the validity of the results. Finally, class action law suites prompted by disclosure of private information can pose a legal barrier to conducting further studies. The *central question* of private data analysis then is whether it is possible to perform a reasonably accurate analysis of sensitive data while meaningfully limiting the disclosure of private information.

Many naive approaches to this problem fail due to the abundance of publicly available information about individuals. For example, declaring a subset of data attributes as "personally identifying" and removing them from the data is vulnerable to linkage attacks. A prominent example is the Netflix de-anonymization attack [112], which showed

that the identities of persons in the anonymized Netflix movie ratings dataset can be recovered by linking the dataset with public Internet Movie Database (IMDb) profiles. Protecting from re-identification itself is not sufficient; for example if we know that our neighbor visited a particular doctor's office, and we receive the anonymized records of visitors for that day, then we have a very short list of possible diseases that our neighbor can be suffering from. Yet we have not identified the neighbor in the records. Finally, restricting analysis to aggregate statistics is also not sufficient, because differencing attacks can be used to infer personal information from simple compositions of aggregate information. These examples illustrate the difficulty of the private data analysis problem and the need for formal definitions, that make guarantees in the face of rich and arbitrary auxiliary information.

Differential privacy [50] is a rigorous mathematical definition introduced to address these issues. The definition requires that the result of the analysis, as a probability distribution on possible outputs, remains almost unchanged if a single individual is added or removed from the data. Semantically, this requirement corresponds to the following guarantee to any participant: regardless of what other studies have been performed, and regardless of what public information is available, participating in the study does not pose a significantly larger privacy threat than not participating. Such a guarantee encourages participation in the data collection process.

7.1.2 Characterizing Optimal Error

In a seminal paper, Dinur and Nissim [48] showed that there are limitations to the accuracy of private analyses of statistical databases, even for very weak notions of privacy. Imagine that a database D consists of n private bits, each bit giving information about one individual (e.g. whether the individual tested positive for some disease). Dinur and Nissim showed that answering slightly more than $O(n)$ random subset sum queries on D with additive error $o(\sqrt{n})$ per query allows an attacker to reconstruct an arbitrarily good approximation to D. Thus there is an inherent trade-off between privacy and accuracy when answering a large number of queries, and our main contribution in this chapter is a characterization of this trade-off for linear queries, in essentially all regimes

considered in the literature.

The first step towards our characterization is a new view of Dinur and Nissim's reconstruction attack using discrepancy. Assume again that the private database is a vector of n private bits, one per individual. Assume also that we are given a set system $\mathcal{S}$ of subsets of $[n]$, and each set in the system encodes a subset sum (i.e. counting) query on D. I.e. for a set $S \in \mathcal{S}$, we want to know how many bits in $D|_S$ are equal to 1. We will show that if some algorithm answers all queries in $\mathcal{S}$ with additive error per query upper bounded by an appropriate notion of discrepancy, then we can reconstruct an arbitrarily good approximation to D. The intuition for why this should be the case is that an adversary can "weed out" databases that are far away from D whenever they give sufficiently different answers from D on the queries in $\mathcal{S}$. Therefore, the only reason why an adversary might fail to reconstruct a good approximation to D is that there is a way for the differences between two far-away databases to balance and cancel each other out; this kind cancellation however is limited by the discrepancy of $\mathcal{S}$. This new approach to reconstruction attacks can be used to re-derive Dinur and Nissim's result as well as a number of other known and new results. But more importantly, it gives a general tool to understand privacy-accuracy trade-offs for arbitrary sets of counting queries.

In order to relate this reconstruction attack to more standard discrepancy notions, we need to give slightly more power to the adversary. Assume that for some set $J \subset [n]$, we give the adversary the bits D_j for $j \notin J$. I.e. the adversary is given the knowledge of the private information of a subset of the individuals. Intuitively, under a reasonable notion of privacy, the adversary should not be able to use this auxiliary knowledge to learn the bits D_j for $j \in J$ from the output of a private algorithm. An immediate consequence of the discrepancy-based reconstruction attack we introduce is that if an algorithm answers all queries in $\mathcal{S}$ with additive error per query $o(\text{hvdisc}(\mathcal{S}))$, then there exists some set J such that, given $D|_{[n]\setminus J}$, the adversary can learn most of D_J from the output of the algorithm. Recall that $\text{hvdisc}(\mathcal{S})$ is equal to $\text{herdisc}(\mathcal{S})$ up to polylogarithmic factors (Corollary 3.2).

As a final step, we show that discrepancy in fact characterizes the *necessary and*

sufficient error to answer queries of the above type. As a motivation, let us consider two edge cases. In one extreme, Dinur and Nissim's argument (strengthened in [52]) shows that error $\Omega(\sqrt{n})$ is necessary for $\Theta(n)$ random subset sum queries in order to achieve any reasonable notion of privacy. This argument is tight, and there are differentially private algorithms that achieve error approximately $O(\sqrt{n \log n})$). As a comparison, the (vector) discrepancy of the corresponding set system of $\Theta(n)$ random sets is $\Omega(\sqrt{n})$. In the other extreme, we can achieve significantly better error guarantees and satisfy the stringent restrictions of differential privacy when the set of queries has a lot of structure. For example, if all our queries ask about the number of bits in $D_1, \ldots, D_j$ set to 1 for some $j \in [n]$ (also known as range queries), then we know of private algorithms that achieve error $O(\log^{3/2} n)$ [53, 35, 152]. As a comparison, the corresponding set system has hereditary discrepancy 1.

The discussion above suggests that hereditary discrepancy may characterize the necessary and sufficient error for privately answering subset sum queries up to polylogarithmic factor. In fact, we are able to show a striking threshold behavior:

- If an algorithm answers subset sum queries $\mathcal{S}$ with error $\tilde{\Omega}(1) \cdot \text{herdisc}(\mathcal{S})$, then an adversary can reconstruct most of the private database (given the right auxiliary information).

- There exists an efficient algorithm which answers all queries $\mathcal{S}$ with error $\tilde{O}(1) \cdot \text{herdisc}(\mathcal{S})$, and achieves a strong level of privacy (differential privacy).

The notation $\tilde{O}, \tilde{\Omega}$ above hides factors polylogarithmic in the size of a natural representation of the queries and the database. The efficient algorithm mentioned above is based on computing an ellipsoid E achieving $\|\mathcal{S}\|_{E\infty}$. The algorithm simply computes the true answers to the queries and adds Gaussian noise correlated according to E.

7.2 Preliminaries on Differential Privacy

Here we introduce the basic definitions and results from differential privacy that will be used in the remainder of the chapter.

7.2.1 Basic Definitions and Composition

A *database* is defined as a multiset $D \in U^n$ of n rows from the *data universe* U of size $|U|$. The notation $|D| \triangleq n$ denotes the *size* of the database. Each row represents the information belonging to a single individual. The universe U depends on the domain; a natural example to keep in mind is $U = \{0,1\}^d$, i.e. each row of the database gives the values of d binary attributes for some individual.

Two databases D and D' are *neighboring* if they differ in the data of at most a single individual, i.e. $|D \triangle D'| \leq 1$.

Differential privacy formalizes the notion that an adversary should not learn too much about any individual as a result of a private computation. The formal definition follows.

Definition 7.1 ([50])**.** *A randomized algorithm $\mathcal{M}$ satisfies (ε, δ)-differential privacy if for any two neighboring databases D and D' and any measurable event S in the range of $\mathcal{M}$,*

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \Pr[\mathcal{M}(D') \in S] + \delta.$$

Above, probabilities are taken over the internal coin tosses of $\mathcal{M}$.

Differential privacy guarantees to a data owner that allowing her data to be used for analysis does not risk much more than she would if she did not allow her data to be used.

Let us remark on the parameters. Usually, ε is set to be a small constant so that $e^{\varepsilon} \approx 1 + \varepsilon$, and δ is set to be no bigger than n^{-2} or even $n^{-\omega(1)}$. The case of $\delta = 0$ often requires different techniques from the case $\delta > 0$; as is common in the literature, we shall call the two cases *pure differential privacy* and *approximate differential privacy*.

An important basic property of differential privacy is that the privacy guarantees degrade smoothly under composition and are not affected by post-processing.

Lemma 7.1 ([50, 51])**.** *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ satisfy $(\varepsilon_1, \delta_1)$- and $(\varepsilon_2, \delta_2)$-differential privacy, respectively. Then the algorithm which on input D outputs the tuple $(\mathcal{M}_1(D), \mathcal{M}_2(\mathcal{M}_1(D), D))$ satisfies $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2)$-differential privacy.*

7.2.2 Query Release

In the query release problem we are given a set $\mathcal{Q}$ of queries (called a *workload*), where each $q \in \mathcal{Q}$ is a function $q\colon U^n \to \mathbb{R}$. Our goal is to design a differentially private algorithm $\mathcal{M}$ which takes as input a database D and outputs a list of answers to the queries in $\mathcal{Q}$. We shall call such an algorithm a *(query answering) mechanism*; this motivates our choice of the notation $\mathcal{M}$ for differentially private algorithms. Here we treat the important special case of query release for sets of *linear queries*. A linear query q is specified by a function $q\colon U \to [-1, 1]$; slightly abusing notation, we define the value of the query as $q(D) \triangleq \sum_{e \in D} q(e)$. When $q\colon U \to \{0, 1\}$ is a predicate, $q(D)$ is a *counting query*: it simply counts the number of rows of D that satisfy the predicate.

7.2.3 Histograms and Matrix Notation

It will be convenient to encode the query release problem for linear queries using matrix notation. A common and very useful representation of a database $D \in U^n$ is the *histogram representation*: the histogram of D is a vector $x \in \mathsf{P}^U$ (P is the set of non-negative integers) such that for any $e \in U$, x_e is equal to the number of copies of e in D. Notice that $\|x\|_1 = n$ and also that if x and x' are respectively the histograms of two neighboring databases D and D', then $\|x - x'\|_1 \leq 1$ (here $\|x\|_1 = \sum_e |x_e|$ is the standard ℓ_1 norm). Linear queries are a linear transformation of x. More concretely, let us define the *query matrix* $A \in [-1, 1]^{\mathcal{Q} \times U}$ associated with a set of linear queries $\mathcal{Q}$ by $a_{q,e} = q(e)$. Then it is easy to see that the vector Ax gives the answers to the queries $\mathcal{Q}$ on a database D with histogram x. Notice also that when $\mathcal{Q}$ is a set of counting queries, then A is the incidence matrix of the set system containing the sets $S_q \triangleq \{e \in U : q(e) = 1\}$.

Since this does not lead to any loss in generality, for the remainder of this chapter we will assume that databases are given to mechanisms as histograms, and workloads of linear queries are given as query matrices. We will identify the space of size-n databases with histograms in the scaled ℓ_1 ball $B_1^U(n) \triangleq \{x : \|x\|_1 \leq n\}$, and we will identify neighboring databases with histograms x, x' such that $\|x - x'\|_1 \leq 1$. Definition 7.1 can

be slightly generalized as follows.

Definition 7.2. *A randomized algorithm $\mathcal{M}$ satisfies (ε, δ)-differential privacy if for any two histograms $x, x' \in \mathbb{R}^U$ such that $\|x - x'\|_1 \leq 1$, and any measurable event S in the range of $\mathcal{M}$,*

$$\Pr[\mathcal{M}(D) \in S] \leq e^{\varepsilon} \Pr[\mathcal{M}(D') \in S] + \delta.$$

Probabilities are taken over the internal coin tosses of $\mathcal{M}$.

Definition 7.1 and Definition 7.2 are equivalent when all histograms considered are integral and non-negative. While all our algorithms will work in a more general setting in which they can take fractional histograms, all our negative results (i.e. lower bounds on error) will be for histograms whose coordinates are non-negative integers, and can, therefore, be interpreted as regular databases.

7.2.4 Measures of Error

In this chapter we study the necessary and sufficient error incurred by differentially private mechanisms for approximating workloads of linear queries. As our basic notions of error, we will consider worst-case and average error. Here we define these notions.

For a mechanism $\mathcal{M}$ and a subset $X \subseteq \mathbb{R}^U$, let us define the worst case error with respect to the query matrix $A \in \mathbb{R}^{\mathcal{Q} \times U}$ as

$$\mathrm{err}(\mathcal{M}, X, A) \triangleq \sup_{x \in X} \mathbb{E} \|Ax - \mathcal{M}(A, x)\|_{\infty},$$

where the expectation is taken over the random coins of $\mathcal{M}$. Another notion of interest is average (L_2)-error, defined as

$$\mathrm{err}_2(\mathcal{M}, X, A) \triangleq \sup_{x \in X} \left(\mathbb{E} \frac{1}{|\mathcal{Q}|} \|Ax - \mathcal{M}(A, x)\|_2^2 \right)^{1/2}.$$

We also write $\mathrm{err}(\mathcal{M}, nB_1^U, A)$ as $\mathrm{err}(\mathcal{M}, n, A)$, and $\mathrm{err}(\mathcal{M}, \mathbb{R}^U, A)$ as $\mathrm{err}(\mathcal{M}, A)$. The optimal error achievable by any (ε, δ)-differentially private algorithm for queries A and databases of size up to n is

$$\mathrm{opt}_{\varepsilon,\delta}(n, A) \triangleq \inf_{\mathcal{M}} \mathrm{err}(\mathcal{M}, n, A),$$

where the infimum is taken over all (ε, δ)-differentially private mechanisms. When no restrictions are placed on the size n of the database, the appropriate notion of optimal error is

$$\mathrm{opt}_{\varepsilon,\delta}(A) \triangleq \sup_n \mathrm{opt}_{\varepsilon,\delta}(n, A) = \inf_{\mathcal{M}} \mathrm{err}(\mathcal{M}, A),$$

where the infimum, as before, is over all (ε, δ)-differentially private mechanisms. The optimal average error $\mathrm{opt}^{(2)}_{\varepsilon,\delta}(n, A)$ and $\mathrm{opt}^{(2)}_{\varepsilon,\delta}(A)$ are defined analogously using err_2.

In order to get tight dependence on the privacy parameter ε in our analyses, we will use the following relationship between $\mathrm{opt}_{\varepsilon,\delta}(n, A)$ and $\mathrm{opt}_{\varepsilon',\delta'}(n, A)$.

Lemma 7.2. *For any $\varepsilon > 0$ and $\delta < 1$, any integer k, and for all $\delta' \geq \frac{e^{k\varepsilon}-1}{e^{\varepsilon}-1}\delta$,*

$$\mathrm{opt}_{\varepsilon,\delta}(kn, A) \geq k\,\mathrm{opt}_{k\varepsilon,\delta'}(n, A).$$

The same holds for $\mathrm{opt}^{(2)}_{\varepsilon,\delta}$.

Proof. Let $\mathcal{M}$ be an (ε, δ)-differentially private mechanism. We will use $\mathcal{M}$ as a black box to construct a $(k\varepsilon, \delta')$-differentially private mechanism $\mathcal{M}'$ which satisfies the error guarantee $\mathrm{err}(\mathcal{M}', n, A) \leq \frac{1}{k}\,\mathrm{err}(\mathcal{M}, A, kn)$, which proves the lemma.

On input x satisfying $\|x\|_1 \leq n$, the mechanism $\mathcal{M}'$ outputs $\frac{1}{k}\mathcal{M}(kx)$. We need to show that $\mathcal{M}'$ satisfies $(k\varepsilon, \delta')$-differential privacy. Let x and x' be two neighboring inputs to $\mathcal{M}'$, i.e. $\|x - x'\|_1 \leq 1$, and let S be a measurable subset of the output $\mathcal{M}'$. Denote $p_1 = \Pr[\mathcal{M}'(x) \in S]$ and $p_2 = \Pr[\mathcal{M}'(x') \in S]$. We need to show that $p_1 \leq e^{k\varepsilon}p_2 + \delta'$. To that end, define $x_0 = kx$, $x_1 = kx + (x' - x)$, $x_2 = kx + 2(x' - x)$, …, $x_k = kx'$. Applying the (ε, δ)-privacy guarantee of $\mathcal{M}$ to each of the pairs of neighboring inputs x_0, x_1, x_1, x_2, …, x_{k-1}, x_k in sequence gives us

$$p_1 \leq e^{k\varepsilon}p_2 + (1 + e^{\varepsilon} + \ldots + e^{(k-1)\varepsilon})\delta = e^{k\varepsilon}p_2 + \frac{e^{k\varepsilon} - 1}{e^{\varepsilon} - 1}\delta.$$

This finishes the proof of privacy for $\mathcal{M}'$. It is straightforward to verify that the errors of the mechanisms are related as $\mathrm{err}(\mathcal{M}', n, A) \leq \frac{1}{k^2}\,\mathrm{err}(\mathcal{M}, A, kn)$. □

We emphasize again that, while the definitions above are stated for general real-valued histograms, defining err and opt in terms of integer histograms (i.e. taking $\mathrm{err}(\mathcal{M}, n, A) \triangleq \mathrm{err}(\mathcal{M}, B_1^U \cap \mathbb{P}^U, A)$ and modifying the other definitions accordingly) does not change the asymptotics of our theorems.

7.2.5 The Main Result

The following theorem is our main result of this chapter and shows the existence of an efficient nearly optimal differentially private algorithm.

Theorem 7.1. *There exists an (ε, δ)-differentially private algorithm $\mathcal{M}$ that runs in time polynomial in $|D|$, $|\mathcal{Q}|$, and $|U|$, and has error* $\operatorname{err}(\mathcal{M}, A) = O(\log^{3/2}|\mathcal{Q}|\sqrt{\log 1/\delta}) \cdot \operatorname{opt}_{\varepsilon,\delta}(A)$ *for any query matrix $A \in \mathbb{R}^{\mathcal{Q}\times A}$, any small enough ε, and any δ small enough with respect to ε. Moreover, we have the inequalities*

$$\frac{1}{O(\log|\mathcal{Q}|)}\frac{1}{\varepsilon}\|A\|_{E\infty} \leq \operatorname{opt}_{\varepsilon,\delta}(A) \leq O(\sqrt{(\log|\mathcal{Q}|)(\log 1/\delta)}) \cdot \frac{1}{\varepsilon}\|A\|_{E\infty}.$$

Hardt and Talwar [78] proved a theorem analogous to Theorem 7.1 in the $\delta = 0$ case, which requires somewhat different techniques. Their algorithm is a factor of $O(\log^{3/2}|\mathcal{Q}|)$ away from optimal, assuming the hyperplane conjecture from convex geometry. Subsequently, Bhaskara et al. [23] improved the competitiveness ratio was to $O(\log|\mathcal{Q}|)$, and made the result unconditional, using Klartag's proof of the isomorphic hyperplane conjecture [86]. We note that, at the cost of a small relaxation in the privacy guarantee, the algorithm in Theorem 7.1 is significantly simpler and more efficient than the known nearly optimal algorithms in the $\delta = 0$ case. Indeed, our algorithm simply computes an ellipsoid E that approximately achieves $\|A\|_{E\infty}$, and then adds Gaussian noise correlated to have the shape of E. By contrast, the algorithms of [78, 23] involve sampling from high-dimensional convex bodies, at a minimum.

7.3 Reconstruction Attacks from Discrepancy

In this section we prove a lower bound on $\operatorname{opt}_{\varepsilon,\delta}(A)$ in terms of $\operatorname{hvdisc}(A)$. The main result follows.

Theorem 7.2. *There exists a constant c, such that for any query matrix $A \in \mathbb{R}^{\mathcal{Q}\times U}$ we have*

$$\operatorname{opt}_{\varepsilon,\delta}(A) \geq \frac{c}{\varepsilon}\operatorname{hvdisc}(A),$$

for all small enough ε and any δ sufficiently small with respect to ε.

We prove the theorem in two steps. First we show that the output of any private algorithm must be far away from the input in every coordinate. Then we show that this implies a lower bound on $\mathrm{opt}_{\varepsilon,\delta}(A)$ for small constant ε and δ via a reconstruction attack. We finish the proof for all ε and small enough δ using Lemma 7.2.

The first lemma shows that any private algorithm should fail to guess each coordinate of its input with constant probability.

Lemma 7.3. *Assume $\mathcal{M}$ is an (ε, δ)-differentially private algorithm whose output range is $\mathbb{R}^W$ for some $W \subseteq U$. Let x be uniformly distributed among vectors in $\{0,1\}^U$ supported on W, and define $\tilde{x} \triangleq \mathcal{M}(x)$. Then, for every $e \in W$,*

$$\Pr_{\mathcal{M},x}[\tilde{x}_e \neq x_e] \geq \frac{e^{-\varepsilon} - \delta}{1 + e^{-\varepsilon}}.$$

where the probability is taken over the coin tosses of $\mathcal{M}$ and the choice of x.

Proof. For each $x \in \{0,1\}^U$ and some $e \in U$, define $x^{(e)}$ to by

$$x_f^{(e)} \triangleq \begin{cases} x_f \oplus 1 & f = e \\ x_f & f \neq e \end{cases}.$$

Let $\tilde{x}^{(e)} \triangleq \mathcal{M}(x^{(e)})$. By the definition of (ε, δ)-differential privacy, since $\|x^{(e)} - x\|_1 \leq 1$, for each x we have

$$\Pr_{\mathcal{M}}[\tilde{x}_e = x_e \otimes 1] \geq e^{-\varepsilon} \Pr_{\mathcal{M}}[\tilde{x}_e^{(e)} = x_e \otimes 1] - \delta.$$

Observe that when x is distributed uniformly over the vectors in $\{0,1\}^U$ supported on W, and $e \in W$, $x^{(e)}$ is distributed identically to x. Then, taking probabilities over the choice of x, we have

$$\Pr_{\mathcal{M},x}[\tilde{x}_e \neq x_e] = \Pr_{\mathcal{M},x}[\tilde{x}_e = x_e \otimes 1]$$
$$\Pr_{\mathcal{M},x}[\tilde{x}_e = x_e] = \Pr_{\mathcal{M},x}[\tilde{x}_e^{(e)} = x_e^{(e)}] = \Pr_{\mathcal{M},x}[\tilde{x}_e^{(e)} = x_e \otimes 1]$$

Combining the above equations, by the law of total probability we conclude that

$$\Pr_{\mathcal{M},x}[\tilde{x}_e \neq x_e] \geq e^{-\varepsilon}(1 - \Pr_{\mathcal{M},x}[\tilde{x}_e \neq x_e]) - \delta.$$

Re-arranging the terms gives the claimed bound. □

The second lemma is the reconstruction result: it shows that if a mechanism has error substantially less than $\mathrm{hvdisc}(A)$, then it can be used to guess its input accurately.

Lemma 7.4. *Let $A \in \mathbb{R}^{\mathcal{Q} \times U}$ be a query matrix, let $W \subseteq U$ be such that $\mathrm{vecdisc}(A_W) = \mathrm{hvdisc}(A)$, and define $X \triangleq \{x \in \{0,1\}^U : x_i = 0 \ \forall e \in U \setminus W\}$. Let $\mathcal{M}$ be a mechanism such that $\mathrm{err}(\mathcal{M}, A, X) \leq \alpha \, \mathrm{hvdisc}(A)$. Then, there exists an assignment $q : W \to \mathbb{R}$ of non-negative reals to W, and a deterministic algorithm $\mathcal{R}$ with range $\{0,1\}^U$ such that, for any x supported on W*

$$\mathbb{E} \sqrt{\frac{1}{q(W)} \sum_{e \in W} q(e)(\tilde{x}_e - x_e)^2} \leq 2\alpha, \tag{7.1}$$

where $\tilde{x} \triangleq \mathcal{R}(\mathcal{M}(x))$, $q(W) \triangleq \sum_{e \in W} q(e)$, and the expectation is taken over the randomness of $\mathcal{M}$.

Proof. Recall the dual formulation (3.12)–(3.15) of vector discrepancy. Let $P \in \mathbb{R}^{\mathcal{Q} \times \mathcal{Q}}$ and $Q \in \mathbb{R}^{W \times W}$ give an optimal feasible solution for (3.12)–(3.15), i.e. they are diagonal matrices such that $A_W^\intercal P A_W \succeq Q$, P is non-negative and satisfies $\mathrm{tr}(P) = 1$, and Q satisfies $\mathrm{hvdisc}(A)^2 = \mathrm{vecdisc}(A_W)^2 = \mathrm{tr}(Q)$. We claim that Q is non-negative. Indeed, otherwise we can take W' to be the set $\{e \in U : q_{ee} > 0\}$, and the solution $P, Q_{W',W'}$ (i.e. the submatrix of Q given by the rows and columns in W') is feasible for (3.12)–(3.15) and has strictly higher value than $\mathrm{vecdisc}(A_W)^2 = \mathrm{hvdisc}(A)^2$, a contradiction. Let us then define q to be the function that maps e to q_{ee}.

On input y, we define $\mathcal{R}(y)$ as

$$\mathcal{R}(y) \triangleq \arg\min_{\tilde{x} \in X} \|A\tilde{x} - y\|_\infty.$$

For any $x \in X$, let $y \triangleq \mathcal{M}(x)$ and $\tilde{x} \triangleq \mathcal{R}(y)$. By the triangle inequality, we have the following guarantee:

$$\mathbb{E}\|A\tilde{x} - Ax\|_\infty \leq \mathbb{E}\|A\tilde{x} - y\|_\infty + \mathbb{E}\|y - Ax\|_\infty.$$

The second term on the right hand side is bounded by assumption by $\mathrm{err}(\mathcal{M}, A, \{0,1\}^U) \leq \alpha \, \mathrm{hvdisc}(A)$. The first term satisfies

$$\mathbb{E} \min_{\tilde{x}} \|A_W\tilde{x} - y\|_\infty \leq \mathbb{E}\|Ax - y\|_\infty \leq \mathrm{err}(\mathcal{M}, A, \{0,1\}^U) \leq \alpha \, \mathrm{hvdisc}(A),$$

since x is one of the possible values for $\tilde{x}$ that the minimum is taken over. Define $z \triangleq (\tilde{x} - x)_W$, and observe that $A\tilde{x} - Ax = A_W z$ since $\tilde{x}$ and x are supported on W. Observe further that for any $v \in \mathbb{R}^{\mathcal{Q}}$, since $\operatorname{tr}(P) = 1$ and P is a non-negative diagonal matrix,

$$\sqrt{v^\intercal P v} = \sqrt{\sum_{q \in \mathcal{Q}} p_{qq} v_q^2} \leq \|v\|_\infty.$$

Indeed, the left hand side of the inequality is the square root of an average of the values v_q^2, while the right hand side is the square root of the maximum. Then, using this fact and $Q \preceq A_W^\intercal P A_W$, we have

$$\mathbb{E}\sqrt{z^\intercal Q z} \leq \mathbb{E}\sqrt{z^\intercal A_W^\intercal P A_W z} \leq \mathbb{E}\|A_W z\|_\infty.$$

Combining the inequalities derived so far, we get $\mathbb{E}\sqrt{z^\intercal Q z} \leq 2\alpha \operatorname{hvdisc}(A) = 2\alpha\sqrt{\operatorname{tr}(Q)}$. Expanding the terms on both sides of the inequality proves the lemma. □

Let us give some interpretation to Lemma 7.4. The right-hand side of (7.1) is an L_2 distance (with respect to weights proportional to $q(e)$) between x and $\tilde{x}$. We have that this L_2 distances is bounded by 2α, so, if the error of $\mathcal{M}$ is much less than $\operatorname{hvdisc}(A)$, then x and $\tilde{x}$ would be proportionally close. In this sense the lemma gives a reconstruction attack. The requirement that x be supported on W can be simulated with unrestricted x by giving the reconstruction algorithm $\mathcal{R}$ the coordinates of x not in W.

We are now ready to finish the proof of Theorem 7.2.

Proof of Theorem 7.2. We first show that for $\varepsilon_0 \leq 1$ and $\delta_0 \leq \frac{1}{2e}$, $\operatorname{opt}_{\varepsilon_0,\delta_0}(A) \geq \frac{1}{2(1+e)} \operatorname{hvdisc}(A)$. The theorem will then follow from Lemma 7.2.

Let $\mathcal{M}$ be an arbitrary $(\varepsilon_0, \delta_0)$-differentially private mechanism with $\operatorname{err}(\mathcal{M}, A) \leq \alpha \operatorname{hvdisc}(A)$. Let W,q, and $\mathcal{R}$ be as in Lemma 7.4, and let x be distributed uniformly among vectors in $\{0,1\}^U$ supported on W, and define $\tilde{x} \triangleq \mathcal{R}(x)$. Then, by Lemma 7.4,

$$\mathbb{E}_{\mathcal{M},x} \frac{1}{q(W)} \sum_{e \in W} q(e)|\tilde{x}_e - x_e| \leq \mathbb{E}_{\mathcal{M},x} \sqrt{\frac{1}{q(W)} \sum_{e \in W} q(e)(\tilde{x}_e - x_e)^2} \leq 2\alpha.$$

The first inequality is by the convexity of the square root function (Jensen's inequality). We use the notation $q(W) \triangleq \sum_{e \in W} q(e)$. Because x and $\tilde{x}$ are both binary, $|\tilde{x}_e - x_e|$ is

1 if $\tilde{x}_e = x_e$ and 0 otherwise. By Lemma 7.3 and linearity of expectation,

$$\mathbb{E}_{\mathcal{M},x}\frac{1}{q(W)}\sum_{e\in W} q(e)|\tilde{x}_e - x_e| = \frac{1}{q(W)}\sum_{e\in W} q(e)\Pr[\tilde{x}_e \neq x_e] \geq \frac{e^{-\varepsilon_0}-\delta}{1+e^{-\varepsilon_0}} \geq \frac{1}{2(1+e)}.$$

The last inequality is by the choice of ε_0 and δ_0. It follows that, $\alpha \geq \frac{1}{2(1+e)}$, which implies that $\mathrm{opt}_{\varepsilon_0,\delta_0} \geq \frac{1}{2(1+e)}\,\mathrm{hvdisc}(A)$.

To finish the proof observe that, by Lemma 7.2, $\mathrm{opt}_{\varepsilon,\delta}(A) \geq \lfloor 1/\varepsilon \rfloor \,\mathrm{opt}_{\varepsilon_0,\delta_0}(A)$, as long as $\varepsilon \leq 1$ and $\delta \leq \frac{e-1}{2(e^{1/\varepsilon}-1)}$. □

7.4 Generalized Gaussian Noise Mechanism

In this section we show an efficient mechanism for answering linear queries with error only a polylogarithmic factor larger than the optimal error. The algorithm is a natural modification of the basic Gaussian noise mechanism, once the latter is viewed in a geometric way. Roughly, our algorithm adds correlated Gaussian noise to the true query answers, where the noise is correlated according to the ellipsoid achieving $\|A\|_{E\infty}$ for the query matrix A. The results on $\|A\|_{E\infty}$ from Chapter 4 let us relate the amount of noise added to $\mathrm{hvdisc}(A)$, which is itself related to $\mathrm{opt}_{\varepsilon,\delta}(A)$ via Theorem 7.2.

7.4.1 The Basic Gaussian Mechanism

The Gaussian noise mechanism [48, 57, 50], which adds appropriately scaled independent Gaussian noise to each query answer, is one of the simplest but most useful tools in differential privacy. In this section we recall the formulation of this mechanism and its privacy guarantee. We give a geometric view of the mechanism, which will allow us to generalize it and derive a near-optimal algorithm.

Recall that $\|A\|_{1\to 2}$ equals the largest ℓ_2 norm of the matrix A. Let us use the notation $N(\mu,\sigma^2)^{\mathcal{Q}}$ for the distribution of a vector of identically distributed independent Gaussian random variables indexed by $\mathcal{Q}$, each with mean μ and variance σ^2. The mechanism is given as Algorithm 1.

To give some geometric intuition for the Gaussian mechanism, let us consider the following symmetric convex body, defined for any query matrix A.

Algorithm 1 Gaussian Mechanism

Input: *(Public)* Query matrix $A \in \mathbb{R}^{\mathcal{Q} \times U}$;
Input: *(Private)* Histogram $x \in \mathbb{R}^U$.

Sample a vector w from $N(0, c_{\varepsilon,\delta}^2 \|A\|_{1\to 2}^2)^{\mathcal{Q}}$, where $c_{\varepsilon,\delta} = \frac{0.5\sqrt{\varepsilon}+\sqrt{2\ln(1/\delta)}}{\varepsilon}$;
Output: Vector of query answers $Ax + w$.

Definition 7.3. *The* sensitivity polytope K_A *of a query matrix* $A \in \mathbb{R}^{\mathcal{Q}\times U}$ *is the convex hull of the columns of* A *and the columns of* $-A$. *Equivalently,* $K_A \triangleq AB_1^U$, *i.e. the image of the unit* ℓ_1 *ball in* $\mathbb{R}^U$ *under multiplication by* A.

The crucial property of the sensitivity polytope is that if x and x' are histograms of neighboring databases, i.e. $\|x - x'\|_1 \leq 1$, then $Ax - Ax' \in K_A$. Moreover, it is easy to see that K_A is the smallest convex body with this property. Because a private mechanism must "hide" whether the input is x or x', the mechanism's output must not allow an observer to distinguish too accurately between Ax and Ax'. So, if a mechanism simply adds noise to the true answers, this noise must be "spread out" over K_A.

The Gaussian mechanism is one concrete realization of this intuition. The mechanism adds independent Gaussian noise with standard deviation $c_{\varepsilon,\delta}\sigma$ to each query answer, where σ is such that $K_A \subset B_2^{\mathcal{Q}}(\sigma)$ (the ball centered at 0 with radius σ). The following lemma shows that this is sufficient noise in order to preserve (ε, δ)-differential privacy

Lemma 7.5 ([48, 57, 50])**.** *The Gaussian mechanism in Algorithm 1 is* (ε, δ)*-differentially private.*

Proof. Let $\sigma \triangleq \|A\|_{1\to 2}$, and also let p be the probability density function of $N(0, c_{\varepsilon,\delta}^2\sigma^2)^{\mathcal{Q}}$, and K_A be the sensitivity polytope of A. Define

$$D_v(w) \triangleq \ln\left(\frac{p(w)}{p(w+v)}\right).$$

We will prove that when $w \sim N(0, c_{\varepsilon,\delta}^2\sigma^2)^{\mathcal{Q}}$, for all $v \in K_A$, $\Pr[|D_v(w)| > \varepsilon] \leq \delta$. This suffices to prove (ε, δ)-differential privacy. Indeed, let the algorithm output $Ax + w$ and fix any x' s.t. $\|x - x'\|_1 \leq 1$. Let $v = A(x - x') \in K_A$ and $S = \{w : |D_v(w)| > \varepsilon\}$. For

any measurable $T \subseteq \mathbb{R}^{\mathcal{Q}}$ we have

$$\begin{aligned}\Pr[Ax + w \in T] &= \Pr[w \in T - Ax] \\ &= \Pr[w \in S \cap (T - Ax)] + \Pr[w \in \bar{S} \cap (T - Ax)] \\ &\leq \delta + e^{\varepsilon}\Pr[w \in T - Ax'] = \delta + e^{\varepsilon}\Pr[Ax' + w \in T].\end{aligned}$$

We fix an arbitrary $v \in K_A$ and proceed to prove that $|D_v(w)| \leq \varepsilon$ with probability at least $1 - \delta$. Recall that $p(w) \propto \exp(-\frac{1}{2c_{\varepsilon,\delta}^2\sigma^2}\|w\|_2^2)$. We have

$$D_v(w) = \frac{\|v + w\|_2^2 - \|w\|_2^2}{2c_{\varepsilon,\delta}^2\sigma^2} = \frac{\|v\|^2 + 2v^T w}{2c_{\varepsilon,\delta}^2\sigma^2}. \tag{7.2}$$

To complete the proof, we bound $|v^T w|$, which suffices to bound $|D_v(w)|$. Since K_A is the convex hull of the columns of A and $-A$ by definition, and each of these columns has ℓ_2 norm at most σ, we have $K_A \subset \sigma B_2^{\mathcal{Q}}$. Because $v \in K_A$, it follows that $\|v\|_2 \leq \sigma$. By standard properties of Gaussian random variables, $v^T w \sim N(0, c_{\varepsilon,\delta}^2\|v\|_2^2\sigma^4)$. A Chernoff bound gives us

$$\Pr\left[|v^T w| > c_{\varepsilon,\delta}\sigma^2\sqrt{2\ln(1/\delta)}\right] < \delta.$$

Substituting back into (7.2), we have that with probability at least $1 - \delta$ the following bounds hold

$$-\frac{\sqrt{2\ln(1/\delta)}}{c_{\varepsilon,\delta}} \leq D_v(w) \leq \frac{1}{2c_{\varepsilon,\delta}^2} + \frac{\sqrt{2\ln(1/\delta)}}{c_{\varepsilon,\delta}}.$$

Substituting $c_{\varepsilon,\delta} \geq \frac{0.5\sqrt{\varepsilon}+\sqrt{2\ln(1/\delta)}}{\varepsilon}$ completes the proof. □

For the remainder of the thesis, we fix the notation

$$c_{\varepsilon,\delta} \triangleq \frac{0.5\sqrt{\varepsilon} + \sqrt{2\ln(1/\delta)}}{\varepsilon}.$$

7.4.2 The Generalization

While a very useful tool, the Gaussian mechanism can, in general, however, be rather wasteful when the sensitivity polytope K_A occupies only a small portion of the ball of radius σ. Consider, for example, the all-ones query matrix J. K_J is just a line segment of length $2\sqrt{|\mathcal{Q}|}$; the Gaussian noise mechanism would add noise roughly $\sqrt{|\mathcal{Q}|}$ in every direction, but it is easy to see that noise $O(1)$ is achievable by outputting $Jx + gj$ where

$g \sim N(0, c^2_{\varepsilon,\delta})$ and j is the all-ones vector. This second mechanism adds Gaussian noise that is still well-spread over K_J but fits the shape of the sensitivity polytope much better.

Motivated by this example, our next step is to find a *correlated* Gaussian distribution which fits K_J as tightly as possible. Any correlated Gaussian distribution is "shaped" according to some ellipsoid E, in the sense that it is equivalent to the uniform distribution over E scaled by a random value drawn from the chi distribution. This, and the example of the basic Gaussian mechanism, suggest the Generalized Gaussian Mechanism presented in Algorithm 2.

Algorithm 2 Generalized Gaussian Mechanism $\mathcal{M}_E$

Input: *(Public)* Query matrix A; ellipsoid $E = F \cdot B_2^{\mathcal{Q}}$ such that all columns of A are contained in E.
Input: *(Private)* Histogram x.
Sample a vector w from $N(0, c^2_{\varepsilon,\delta})^{\mathcal{Q}}$.
Output: Vector of query answers $Ax + Fw$.

The following lemma shows that the generalized Gaussian mechanism is (ε, δ)-differentially private by a simple reduction to the standard Gaussian mechanism.

Lemma 7.6. *The generalized Gaussian mechanism $\mathcal{M}_E$ in Algorithm 2 satisfies (ε, δ)-differential privacy for any ellipsoid $E = FB_2^{\mathcal{Q}}$ that contains the columns of A.*

Proof. Define $\tilde{A} = F^+A$, where F^+ is the Moore-Penrose pseudo-inverse of F. The columns of $\tilde{A}$ are contained in $F^+E = F^+FB_2^{\mathcal{Q}} = \Pi B_2^{\mathcal{Q}}$, where Π is the orthogonal projection operator onto the span of the row vectors of F. Since $\Pi B_2^{\mathcal{Q}}$ is also a ball of radius 1, this implies $\|\tilde{A}\|_{1\to 2} \leq 1$, and, by lemma 7.5, the mechanism that outputs $\tilde{A}x + w$ for $w \sim N(0, c^2_{\varepsilon,\delta})^{\mathcal{Q}}$ is (ε, δ)-differentially private. The output of the generalized mechanism in Algorithm 7.6 is distributed identically to $F(\tilde{A}x + w)$, which is a post-processing of the basic Gaussian mechanism and is also (ε, δ)-differentially private by Lemma 7.1. □

An immediate consequence of Lemma 7.6 (and a standard concentration of measure argument) is an upper bound on $\text{opt}_{\varepsilon,\delta}(A)$ in terms of $\|A\|_{E\infty}$. This is excellent news, since $\|A\|_{E\infty}$ approximates $\text{hvdisc}(A)$, which itself gives a lower bound on $\text{opt}_{\varepsilon,\delta}(A)$.

Proof of Theorem 7.1. Let $E = FB_2^{\mathcal{Q}}$ be an ellipsoid that (approximately) achieves $\|A\|_{E\infty}$. As in the proof of Theorem 4.12, such E can be computed in time polynomial in $|\mathcal{Q}|$ and $|U|$ by solving the convex optimization problem (4.5)–(4.8). The generalized Gaussian mechanism $\mathcal{M}_E$ instantiated with E is (ε, δ)-differentially private by Lemma 7.6. Once E is computed, the mechanism only needs to sample $|\mathcal{Q}|$ Gaussian random variables and perform elementary linear algebra operations, so it runs in polynomial time as well. The error of $\mathcal{M}_E$ is equal to $\mathbb{E}\|Fw\|_\infty = \mathbb{E}\max_{q\in\mathcal{Q}} |e_q^\intercal Fw|$, where $w \sim N(0, c_{\varepsilon,\delta}^2)^{\mathcal{Q}}$ and e_q is the standard basis vector corresponding to query q. For any q, $e_q^\intercal Fw$ is a Gaussian random variable with mean 0 and variance equal to

$$\mathbb{E}(e_q^\intercal Fw)^2 = \mathbb{E}e_q^\intercal Fww^\intercal F^\intercal e_q = c_{\varepsilon,\delta}^2 e_q^\intercal FF^\intercal e_q.$$

Therefore, for any $q \in \mathcal{Q}$, the variance of $e_q^\intercal Fw$ is at most $c_{\varepsilon,\delta}^2 \max_{q\in\mathcal{Q}} e_q^\intercal FF^\intercal e_q = c_{\varepsilon,\delta}^2\|E\|_\infty^2 = c_{\varepsilon,\delta}^2\|A\|_{E\infty}^2$, where the first equality is by (4.4). By a Chernoff bound, for any $q \in \mathcal{Q}$ and any $t > 0$, $\Pr[(e_q^\intercal Fw)^2 \geq tc_{\varepsilon,\delta}^2\|A\|_{E\infty}^2] \leq e^{-t/2}$. By the union bound,

$$\Pr[\max_{q\in\mathcal{Q}}(e_q^\intercal Fw)^2 \geq (t + \ln|\mathcal{Q}|)c_{\varepsilon,\delta}^2\|A\|_{E\infty}^2] \leq e^{-t/2},$$

and we can bound the error of $\mathcal{M}_E$ as

$$\begin{aligned}
\mathrm{err}(\mathcal{M}_E, A) &= \mathbb{E}\max_{q\in\mathcal{Q}} |e_q^\intercal Fw| \\
&\leq \left(\mathbb{E}\max_{q\in\mathcal{Q}} (e_q^\intercal Fw)^2\right)^{1/2} \\
&= \left(\int_0^\infty \Pr[\max_{q\in\mathcal{Q}} (e_q^\intercal Fw)^2 \geq x]dx\right)^2 \\
&\leq \left((\ln|\mathcal{Q}|)c_{\varepsilon,\delta}^2\|A\|_{E\infty}^2 + \int_0^\infty \exp(-x/(2c_{\varepsilon,\delta}^2\|A\|_{E\infty}^2))dx\right)^{1/2} \\
&= O(\sqrt{\log|\mathcal{Q}|\log 1/\delta}) \cdot \frac{1}{\varepsilon}\|A\|_{E\infty},
\end{aligned}$$

where the first inequality is by Jensen. On the other hand, by Theorems 4.12 and 7.2, for all sufficiently small ε and δ,

$$\frac{1}{\varepsilon}\|A\|_{E\infty} = O(\log|\mathcal{Q}|) \cdot \frac{1}{\varepsilon}\,\mathrm{hvdisc}(A) = O(\log|\mathcal{Q}|) \cdot \mathrm{opt}_{\varepsilon,\delta}(A).$$

Combining the bounds finishes the proof of the theorem. □

We leave the following question open.

Question 4. *What is the largest gap between* $\frac{1}{\varepsilon}\operatorname{hvdisc}(A)$ *and* $\operatorname{opt}_{\varepsilon,\delta}(A)$*?*

A particularly interesting question is to develop general lower bound techniques that give bounds which grow with $1/\delta$. A first step in this direction was taken by Bun, Ullman, and Vadhan [33], who showed that on a natural query matrix A, $\operatorname{opt}_{\varepsilon,\delta}(A)$ is strictly larger for $\delta = o(|D|^{-1})$ than for $\delta = \Omega(1)$.

The approach of minimizing error over ellipsoids E to use in the generalized Gaussian mechanism is related to work on the Matrix Mechanism [91]. Instead of the geometric presentation we chose, the matrix mechanism is usually given as a matrix factorization: the query matrix A is written as $A = F\tilde{A}$ and the basic Gaussian mechanism is used to compute noisy answers $\tilde{A}x + w$, which are then multiplied by the matrix F. An inspection of the proof of Lemma 7.6 shows that this is equivalent to the generalized Gaussian mechanism $\mathcal{M}_E$ instantiated with the ellipsoid $E = FB_2^{\mathcal{Q}}$. In the matrix mechanism one usually optimizes over factorizations that minimize the L_2 error, while we analyze the stronger worst-case error guarantee. Nevertheless, we can prove an analogue of Theorem 7.1 for L_2 error as well. Thus, Theorem 7.1 can be interpreted as showing that (a natural variant of) the matrix mechanism is nearly optimal among *all* differentially private mechanisms.

7.5 Bounds on Optimal Error for Natural Queries

Theorem 7.1 and estimates of the ellipsoid infinity norm from Chapter 6 give near tight bounds on the optimal error for natural sets of linear queries. Here we give the more interesting results.

Geometric range counting queries are a natural class of linear queries. In a range counting query the data universe is $U \subset \mathbb{R}^d$, i.e. the database is a multiset of d-dimensional points. One may assume, for example that $U = [N]^d$ for a large enough d. The counting queries are given by a collection $\mathcal{S}$ of subsets of U, usually induced by some natural family of geometric sets. Each query q_S for $S \in \mathcal{S}$ asks how many points in the database belong to S. I.e. $q_S(p)$ is the indicator of $p \in S$ for any $p \in U$. For such queries, let us write $\mathcal{Q}_{\mathcal{S}}$ to denote the class of counting queries induced by the set

system $\mathcal{S}$.

For example, when d is fixed, and $U = [N]^d$, the queries given by the family $\mathcal{B}_d$ of axis-aligned boxes $[a_1, b_1), \ldots, [a_d, b_d)$ in $\mathbb{R}^d$ are known as *orthogonal range queries*. Using various methods (Haar wavelets, decompositions into canonical boxes), it is known that $\mathrm{opt}_{\varepsilon,\delta}(\mathcal{Q}_{\mathcal{B}_d}) = O(c_{\varepsilon,\delta} \log^{d+1/2} N)$. In the next theorem gives a different derivation of this upper bound via our more general results, and, more importantly, also shows a nearly-matching lower bound.

Theorem 7.3. *For any constant dimension d, the optimal error for d-dimensional orthogonal range queries is bounded as*

$$\frac{\Omega((\log N)^{d-1})}{\varepsilon} \leq \mathrm{opt}_{\varepsilon,\delta}(Q_{\mathcal{B}_d}) \leq O(c_{\varepsilon,\delta}(\log N)^{d+1/2}),$$

for a large enough constant C.

Proof. Follows immediately from Theorems 7.1 and the fact that $\|\mathcal{B}_d\|_{E\infty} = \Theta((\log N)^d)$ for any constant d. This fact was established in the proof of Theorem 6.5 □

When instead of $\mathcal{B}_d$ we consider the queries induced by the family of boolean subcubes $\mathcal{C}_d$, we get another interesting set of queries: the *marginal queries* $\mathcal{Q}^d_{\mathrm{marg}}$. The data universe U for this class of queries is $\{0,1\}^d$, i.e. each individual is characterized by d binary attributes. A marginal query q_v is specified by a vector $v \in \{1, 0, *\}^d$, and the result is the count of the number of people in the database whose attributes agree with v except at the $*$ coordinates, where they can be arbitrary. Marginal queries are a ubiquitous and important subclass of queries, constituting *contingency tables* in statistics and *OLAP cubes* in databases. Official agencies such as the Census Bureau, the Internal Revenue Service, and the Bureau of Labor Statistics all release certain sets of low dimensional marginals for the data they collect.

From the description above it is clear that the marginal queries are equivalent to the queries $\mathcal{Q}_{\mathcal{C}_d}$ induced by boolean subcubes. From prior work it was known that answering all marginal queries requires error on the order of $2^{\Omega(d)}$. Our methods allow us to determine the precise constant in the exponent up to lower order terms. Moreover, we show that the same error bound holds for the more restricted class of *conjunction*

queries $\mathcal{Q}^d_{\text{conj}}$, which are all queries q_v for $v \in \{1, *\}^d$. Conjunction queries are equivalent to the queries induced by anchored subcubes of the Boolean cube. Our result follows.

Theorem 7.4. *The optimal error for d-dimensional conjunction and marginal queries is bounded as*

$$\frac{1}{\varepsilon} 2^{c_0 d - o(d)} \leq \text{opt}_{\varepsilon,\delta}(\mathcal{Q}^d_{conj}) \leq \text{opt}_{\varepsilon,\delta}(\mathcal{Q}^d_{marg}) \leq c_{\varepsilon,\delta} 2^{c_0 d + o(d)}.$$

where $c_0 = \log_2(2/\sqrt{3}) \approx 0.2075$.

Proof. Follows from the observations that $\mathcal{Q}^d_{\text{conj}} = \mathcal{Q}_{\mathcal{A}_d(\{0,1\}^d)}$ and $\mathcal{Q}^d_{\text{marg}} = \mathcal{Q}_{\mathcal{C}_d}$, Theorems 6.6, 4.12, and 7.1, since $|\mathcal{Q}^d_{\text{conj}}| = 3^d$. Alternatively to using Theorems 6.6 and 4.12, observe that in the proof of Theorem 6.6 we showed that $\|\mathcal{A}_d(\{0,1\}^d)\|_{E\infty} = \|\mathcal{C}_d\|_{E\infty} = 2^{c_0 d}$. □

7.6 Error Lower Bounds for Pan-Privacy

In this section we extend the reconstruction attack-based lower bounds on error to a stronger notion of privacy. We show a strong separation between this stronger notion and ordinary differential privacy.

7.6.1 Pan Privacy: Motivation and Definition

The original definition of differential privacy implicitly assumes that the database itself is secured against intrusion, and the privacy risk comes from publishing query answers. This is a reasonable assumption, as it allows us to separate the issues of security, such as access control and protection against tampering with the data, and privacy issues which cannot be addressed with traditional cryptographic tools because an adversary and a legitimate user of the statistical analysis cannot be separated. However, if the security of the system is compromised, and the sensitive data is leaked in the clear, all hope of further privacy protection is lost. This may happen because a malicious intruder hacks the system. But more subtly, this may happen because an insider with access, such as a systems administrator, may turn curious or crooked; data analysis may be outsourced to far away countries where people and laws are less stringent; or

the contents of the registers may be subpoenaed by law or security officials. Traditional encryption will not work in such cases, because a breach will reveal the hash function or the encrypting key.

To protect privacy in the face of this kind of security breach, we need to store the data itself in a privacy-preserving manner. Dwork, Naor, Pitassi, Rothblum, and Yekhanin introduced *pan-privacy* to formally capture this stronger level of security protection. Intuitively, pan-privacy applies to algorithms that process a stream of data updates, and requires that the pair of internal state of the algorithm and output are jointly differentially private. More formally, we model a streaming algorithm as a pair $\mathcal{M} = (\mathcal{A}, \mathcal{O})$, where $\mathcal{A}$ and $\mathcal{O}$ are each *randomized* algorithms. Assume that the input sequence $\bar{\sigma} \triangleq (\sigma_1, \ldots, \sigma_m)$ arrives online, i.e. at time step t we receive the symbol $\sigma_t \in \Sigma$. We can think of each σ_t as the ID of a user visiting a website, for example. At each time step t, $\mathcal{M}$ is in a state $X_t \in \mathcal{X}$ ($\mathcal{X}$ is called the *state space*); it is initialized to a special *initial state* X_0, and after the symbol σ_t arrives at time t, the state is changed to $\mathcal{M}(X_{t-1}, \sigma_t)$. After the input sequence is processed, $\mathcal{M}$ outputs $\mathcal{O}(X_m)$. (It is possible to modify the definitions so that the algorithm produces multiple outputs; we restrict the discussion to a single output produced at the end for simplicity.) We use the shorthand $\mathcal{A}(X, \bar{\sigma})$ for $X_m = \mathcal{A}(X_{m-1}, \sigma_m)$ where $X_{m-1} = \mathcal{A}(X_{m-2}, \sigma_{m-1})$, ..., $X_1 = \mathcal{A}(X, \sigma_1)$; when X is the initial state X_0, we just write $\mathcal{A}(\bar{sigma})$. In this setting, pan privacy is defined as follows:

Definition 7.4 ([54])**.** *Two input sequences $\bar{\sigma} = (\sigma_1, \ldots, \sigma_m)$ and $\bar{\sigma}' = (\sigma'_1, \ldots, \sigma'_{m'})$ are* neighboring *if there exists a symbol $\sigma \in \Sigma$ so that $\bar{\sigma}'$ can be derived from $\bar{\sigma}$ by only adding and removing occurrences of σ. An algorithm $\mathcal{M} = (\mathcal{A}, \mathcal{O})$ with state space $\mathcal{X}$ is* (ε, δ)-pan private against a single intrusion *if for*

- *all pairs of neighboring streams $\bar{\sigma} = \bar{\sigma}_1 \cdot \bar{\sigma}_2$, and $\bar{\sigma}' = \bar{\sigma}'_1 \cdot \bar{\sigma}'_t$ where $\cdot$ is concatenation and $\bar{\sigma}_1, \bar{\sigma}'_1$ are also neighboring and differ on the same symbol as $\bar{\sigma}, \bar{\sigma}'$,*
- *all sets $X \subseteq \mathcal{X}$,*
- *all subsets S of the range of $\mathcal{O}$,*

we have

$$\Pr[(\mathcal{A}(\bar{\sigma}_1), \mathcal{O}(\mathcal{A}(\bar{\sigma}))) \in X \times S] \leq e^{\varepsilon}\Pr[(\mathcal{A}(\bar{\sigma}'_1), \mathcal{O}(\mathcal{A}(\bar{\sigma}'))) \in X \times S] + \delta,$$

where the probability is taken over the randomness of $\mathcal{A}$ and $\mathcal{O}$.

The intention of the definition is that an intrusion occurs after σ_1 has been processed, after which the intruder can wait to also observe the output of the algorithm. The definition can be generalized to multiple intrusions. However if there are more than two unannounced intrusions, Dwork et al. showed that even simple functions of the stream cannot be approximated with any non-trivial error. On the other hand, if the breach is discovered before a second breach occurs, the algorithm's state can be re-randomized at the cost of a slight increase in error.

While the pan-privacy model is very strong, Dwork et al. designed a number of non-trivial pan-private algorithms for statistics on streams. Recall the notion of a *frequency vector* $f \in \mathbb{N}^{\Sigma}$ associated with a stream $\bar{\sigma}$, and defined by $f_\sigma \triangleq |\{t : \sigma_t = \sigma\}|$. The k-th *frequency moment* of the stream is defined as $F_k \triangleq \sum_{\sigma\in\Sigma} f_\sigma^k$; the 0-th moment, also known as the *distinct count*, is equal to the number of distinct symbols in the stream, i.e. $F_0 \triangleq |\{\sigma \in \Sigma : f_\sigma \neq 0\}|$. Besides the distinct count, all other frequency moments have unbounded sensitivity, i.e. the value of the moment can differ by an arbitrary amount between two neighboring streams. This easily implies that the worst-case error under differential privacy has to be unbounded. To address this issue, Dwork et al. introduced the cropped moments: the k-th moment $F_k(\tau)$ is equal to $F_k(\tau) \triangleq \sum_{\sigma\in\Sigma} \min\{f_\sigma^k, \tau\}$. They gave an $(\varepsilon, 0)$-pan private algorithm for distinct count that achieves additive error $O(\sqrt{|\Sigma|/\varepsilon})$ with constant probability, and an algorithm for $F_1(\tau)$ that achieves error $O(\tau\sqrt{|\Sigma|/\varepsilon})$. In this section we use a discrepancy-based reconstruction attack, similar to the one in Section 7.3, to show that these algorithms are optimal:

Theorem 7.5. *For any small enough ε and any δ small enough with respect to ε, for any (ε, δ)-pan private algorithm $\mathcal{M} = (\mathcal{A}, \mathcal{O})$, there exists a stream $\bar{\sigma}$ of length $m = \Theta(|\Sigma|)$ such that with probability at least α, $|F_0 - \mathcal{O}(\mathcal{A}(\bar{\sigma}))| = \Omega(\sqrt{|\Sigma|\log(1/\alpha)/\varepsilon})$. Similarly, there exists a stream $\bar{\sigma}$ such that with probability at least α $|F_1(\tau) - \mathcal{O}(\mathcal{A}(\bar{\sigma}))| =$*

$\Omega(\tau\sqrt{|\Sigma|\log(1/\alpha)/\varepsilon})$.

Note that under $(\varepsilon, 0)$-differential privacy, F_0 can be computed with error only $O(\frac{1}{\varepsilon})$ and $F_1(\tau)$ can be computed with error $O(\frac{1}{\varepsilon}\tau)$ via the Laplace noise mechanism [50]. Theorem 7.5 therefore shows that pan privacy can have significantly larger cost in terms of error compared with differential privacy, even in the presence of only a single unannounced intrusion.

7.6.2 Reconstruction Attack against Pan- Privacy

Our lower bound argument is based on the observation that the state of a pan-private algorithm can be used to answer many queries privately, by updating it with different continuations of the stream. If the algorithm is accurate with constant probability for any stream, then most of the query answers we get in this manner will be accurate. We can use this to derive a reconstruction attack via a robust notion of discrepancy. The main ideas of the attack itself, together with Fano's inequality, will also be used in Chapter 9 to prove lower bounds in the one-way communication model, as well as to bound the minimax rate of statistical estimators.

Let us introduce the following "norm" for $x \in \mathbb{R}^m$: $\|x\|_{\alpha,\infty} \triangleq \min\{t : |\{i : |x_i| > t\}| \leq \alpha m\}$. Equivalently, if $x_{(1)}, \ldots, x_{(m)}$ are the coordinates of x in non-increasing order with respect to the absolute value, then $\|x\|_{\alpha,\infty} \triangleq |x_{(k)}|$ where $k \triangleq \lceil \alpha m \rceil$. This quantity can be considered a "robust infinity norm", as it ignores the top α fraction of coordinates, which may be outliers. We call it the α-infinity norm. It is not strictly speaking a norm, because, while it is homogeneous, it need not satisfy the triangle inequality. Nevertheless, it satisfies the following relaxed form of the triangle inequality:

$$\|x + y\|_{2\alpha,\infty} \leq \|x\|_{\alpha,\infty} + \|y\|_{\alpha,\infty}. \tag{7.3}$$

We base a notion of *robust discrepancy* of an $m \times n$ matrix on the α-infinity norm:

$$\mathrm{rdisc}_{\alpha,\beta}(A) \triangleq \min_{x \in \{-1,0,1\}^n : \|x\|_1 \geq \beta n} \|Ax\|_{\alpha,\infty}.$$

Notice that robust discrepancy relaxes discrepancy in two ways: it allows for x to have

a constant fraction of zero coordinates, and it also relaxes the infinity norm to the α-infinity norm. Nevertheless, the next lemma shows that we can still prove strong lower bounds on robust discrepancy.

Lemma 7.7. *Let A be the matrix whose rows are the elements of the set $\{-1,1\}^n$. Then $\operatorname{rdisc}_{\alpha,\beta}(A) \geq c\min\{\beta n, \sqrt{\beta n \log(1/\alpha)}\}$ for an absolute constant c.*

Proof. Let us fix an arbitrary $x \in \{-1,0,1\}^n$ such that $\|x\|_1 \geq \beta n$. It suffices to show that there exists a constant c such that $\Pr_a[|\langle a,x\rangle| \geq c\sqrt{\beta n \log(1/\alpha)}] > \alpha$, where $a \in \{-1,1\}^n$ is picked uniformly at random. Let $\ell \triangleq \beta n$. For a fixed x, the random variable $\langle a,x\rangle$ is distributed identically to $\sum_{i=1}^{\ell} s_i$ where each s_i is uniformly sampled from $\{-1,1\}$. Since $\sum_{i=1}^{\ell} s_i = 2(|\{i : s_i = 1\}| - \ell/2)$, we have

$$\Pr_a[|\langle a,x\rangle| \geq t] = 2^{1-n}\sum_{k=1}^{\ell/2-t/2}\binom{\ell}{k} \geq 2^{1-n}\binom{\ell}{\ell/2-t/2} \geq 2^{(H_2(p)-1)\ell - o(\ell)}.$$

Above $H_2(p) \triangleq -p\log_2 p - (1-p)\log_2(1-p)$ is the binary entropy function, $p \triangleq (1-t/\ell)/2$, and the final inequality follows from Sterling's approximation. By the Taylor expansion of binary entropy,

$$H_2(p) - 1 \geq -\frac{1}{2\ln 2}(1-2p)^2 - \frac{(1-2p)^4}{6\ln(2)p^2}.$$

For $p \geq 1/3$, $\frac{(1-2p)^4}{6\ln(2)p^2} \leq \frac{1}{2\sqrt{3}\ln 2}(1-2p)^2$, so we have $H_2(p) - 1 \geq \frac{1+\sqrt{3}}{2\sqrt{3}\ln 2}(1-2p)^2$. Therefore, for $t \leq \ell/3$,

$$\Pr_a[|\langle a,x\rangle| \geq t] \geq 2^{-Ct^2/\ell},$$

where $C = \frac{1+\sqrt{3}}{2\sqrt{3}\ln 2} - o(1)$. Choosing $t = \min\{\ell/3, \sqrt{\ell\log_2(1/\alpha)/C}\}$ completes the proof. □

The lower bound in Lemma 7.7 is optimal up to constants, as shown by a random coloring and Hoeffding's inequality. Armed with this lower bound, and the approximate triangle inequality for the α-infinity norm, we are ready to give our reconstruction attack.

Lemma 7.8. *There exists a deterministic algorithm $\mathcal{R}$, such that for any $A \in \mathbb{R}^{m\times n}$, any $x \in \{0,1\}^n$ and any y such that $\|Ax - y\|_{\alpha,\infty} < \frac{1}{2}\operatorname{rdisc}_{2\alpha,\beta}$, we have $\mathcal{R}(A,y) \in \{0,1\}^n$ and $\|\mathcal{R}(A,y) - x\|_1 < \beta n$.*

Proof. On input y, we define $\mathcal{R}(A, y)$ as

$$\mathcal{R}(A, y) \triangleq \arg \min_{\tilde{x} \in \{0,1\}^n} \|A\tilde{x} - y\|_{\alpha,\infty}.$$

Let $\tilde{x} \triangleq \mathcal{R}(A, y)$ and $D \triangleq \mathrm{rdisc}_{2\alpha,\beta}(A)$. By assumption, $\|A\tilde{x} - y\|_{\alpha,\infty} \leq \|Ax - y\|_{\alpha,\infty} \leq D/2$. By the approximate triangle inequality (7.3), we have the guarantee

$$\|A\tilde{x} - Ax\|_{2\alpha,\infty} \leq \|A\tilde{x} - y\|_{\alpha,\infty} + \|y - Ax\|_{\alpha,\infty} \leq D.$$

Since $\tilde{x}$ and x are binary, $\tilde{x} - x \in \{-1, 0, 1\}^n$, and by the definition of $\mathrm{rdisc}_{2\alpha,\beta}(A)$, we have $\|\tilde{x} - x\|_1 < \beta n$. □

We are now ready to prove our lower bound result. Once again, the intuition is that the state of a pan-private algorithm can be used to answer many queries, and if the algorithm is accurate with large constant probability, then a large fraction of the queries will be answered accurately. Then we can invoke the reconstruction result in Lemma 7.8. This idea is inspired by space lower bound arguments for streaming algorithms, in which one argues that the space complexity of the algorithm must be large because it accurately encodes the answers to too many queries.

Proof of Theorem 7.5. Let ε_0 be a constant to be determined later, and let $k \triangleq \lfloor \varepsilon_0/\varepsilon \rfloor$, and $n \triangleq \lfloor |\Sigma|/k \rfloor$. Let us associate with each $j \in [n]$ a set $\Sigma_j \subseteq \Sigma$ of size k, so that $\Sigma_j \cap \Sigma_\ell = \emptyset$ for $j \neq \ell$. Given a vector $x \in \{0,1\}^n$, we construct a stream $\bar{\sigma}(x)$ as follows: for each $j \in [n]$ such that $x_j = 1$, we insert into $\bar{\sigma}(x)$ all symbols in Σ_j in any order. We will use the state of a pan-private algorithm for distinct counts after processing $\bar{\sigma}(x)$ to answer counting queries on x under differential privacy.

Let β and c' be constants to be determined later, and assume for the sake of contradiction that $\mathcal{M} = (\mathcal{A}, \mathcal{O})$ is an (ε, δ)-pan private algorithm such that for any stream $\bar{\sigma}$, with probability at least $1 - \alpha$, $|F_0 - \mathcal{O}(\mathcal{A}(\bar{\sigma}))| < c'k\sqrt{\beta n \log(1/\alpha)}$. By an argument analogous to the proof of Lemma 7.2, $X(x) \triangleq \mathcal{A}(\bar{\sigma}(x))$ is an $(\varepsilon_0, \delta_0)$-differentially private function of x, for any $\delta_0 \geq \frac{e^{\varepsilon_0}-1}{e^{\varepsilon}-1}\delta$. Let A be a matrix whose rows $a^1, \ldots, a^m$, $m = 2^n$, form the set of all binary vectors $\{0,1\}^n$. For each row a^i we construct a stream $\bar{\sigma}^i$ in which we insert all symbols in Σ_j (in any order)

for each j such that $a^i_j = a_{ij} = 1$. Observe that the distinct count of the stream $\bar{\sigma}(x) \cdot \bar{\sigma}^i$ is equal to $\sum_{j=1}^n k(\{a_{ij} = 1\} \vee \{x_j = 1\})$. Substituting $\{a_{ij} = 1\} \vee \{x_j = 1\} = \frac{1}{2}(a_{ij} + 1) + x_j - \frac{1}{2}(a_{ij} + 1)x_j$, re-arranging the terms and simplifying gives us that the distinct count of $\bar{\sigma}(x) \cdot \bar{\sigma}^i$ is

$$-\frac{k}{2}\langle a^i, x\rangle + \frac{k}{2}\sum_{j=1}^{n}(x_j + a_{ij}) + \frac{kn}{2}.$$

The third term does not depend on x and the second term can be approximated with additive error $\eta \triangleq O(kc_{\varepsilon_0,\delta_0}\sqrt{\log 1/\alpha})$ with probability $1 - \alpha$ under $(\varepsilon_0, \delta_0)$-differential privacy using the Gaussian mechanism (Lemma 7.5). Let us call this approximation w, and compute a vector y by

$$y_i \triangleq -\frac{2}{k}\mathcal{O}(\mathcal{A}(X(x), \bar{\sigma}^i)) + \frac{2}{k}w + n.$$

Because $X(x)$ and w are each $(\varepsilon_0, \delta_0)$-differentially private functions of x, y is an $(2\varepsilon_0, 2\delta_0)$-differentially private function of x by Lemma 7.1. Sample x uniformly at random from $\{0,1\}^n$ and define the random variable $\tilde{x} \triangleq \mathcal{R}(A, y)$, where $\mathcal{R}$ is the reconstruction algorithm from Lemma 7.8. Because $\mathcal{R}$ only accesses the $(2\varepsilon_0, 2\delta_0)$-differentially private y, $\tilde{x}$ is also $(2\varepsilon_0, 2\delta_0)$-differentially private, and by Lemma 7.3 and linearity of expectation, $\mathbb{E}\|\tilde{x} - x\|_1 \geq \frac{e^{-2\varepsilon_0} - 2\delta_0}{1 + e^{-2\varepsilon_0}} n$. On the other hand, by the accuracy guarantee we assumed for $\mathcal{M}$, for each i, with probability at least $1 - 2\alpha$, $|y_i - (Ax)_i| < \frac{2}{k}\eta + 2c'\sqrt{\beta n \log(1/\alpha)} \leq \frac{1}{2}c\sqrt{\beta n \log(1/12\alpha)}$. Since $\frac{2}{k\sqrt{\log(1/\alpha)}}\eta$ is a constant depending only on ε_0, δ_0, setting c' small enough ensures that the last inequality holds for c as in Lemma 7.7 and all big enough n. Then, by Markov's inequality, with probability at least $2/3$, $\|y - Ax\|_{6\alpha,\infty} < \frac{1}{2}c\sqrt{\beta n \log(1/12\alpha)}$, so, by Lemmas 7.7 and 7.8, $\|\tilde{x} - x\|_1 < \beta n$. We can then bound the expected distance of $\tilde{x}$ from x as

$$\mathbb{E}\|\tilde{x} - x\|_1 < \frac{2}{3}\beta n + \frac{1}{3}n = \frac{2\beta + 1}{3}n.$$

Setting ε_0, δ_0 and β small enough so that $\frac{e^{-2\varepsilon_0} - 2\delta_0}{1 + e^{-2\varepsilon_0}} \geq \frac{2\beta+1}{3}$ gives a contradiction and finishes the proof of the lower bound for distinct counts.

The proof for the cropped first moment is analogous, with the modification that any symbol in the streams $\bar{\sigma}(x)$ and $\bar{\sigma}^i$ is included τ times. □

Bibliographic Remarks

The first discrepancy-based reconstruction attack appeared in the paper [110]. In the notation used in this thesis, the result shown there was that $\mathrm{opt}_{\varepsilon_0,\delta_0}(A) = \Omega(1) \cdot \mathrm{rdisc}_{0,1/2}(A) \geq \frac{1}{O(\log |U|))}$ for a query matrix $A \in \mathbb{R}^{\mathcal{Q} \times U}$ and all small enough constant ε_0, δ_0. Theorem 7.2 appears for the first time in this thesis and is a strengthening of the lower bound in [110]. That paper also observed that this implies interesting error lower bounds for orthogonal range queries, but did not report near-tight bounds, because the corresponding discrepancy lower bound was not known. [110] also used a decomposition of bounded shatter function set systems from discrepancy theory to a differentially private algorithm for halfspace counting with tight error (up to constants). Theorem 7.1 was first proved in [118]. The approach via the ellipsoid infinity norm in this chapter is a simplification of the proofs in [118], which used a recursive construction based on computing approximate John-Löwner ellipsoids. The error lower bound for computing distinct counts and cropped first moment in the pan-privacy model was first published in [108]. The argument is basically the same, but the proof in the paper used the reconstruction attack of [52]. Here we instead use a discrepancy-based reconstruction; while our reconstruction algorithm is not efficient, this is not an issue for the lower bound argument, as differential privacy is an information theoretic notion.

Chapter 8

Private Mechanisms for Small Databases

8.1 Overview

In this chapter we consider a setting in which the database size n is significantly smaller than the number of queries $|\mathcal{Q}|$. Since the seminal work Blum, Ligett and Roth [26], a long line of work [55, 58, 125, 76, 73, 77, 74] has shown that in this regime there exist algorithms with error guarantees superior to the general case. In general, there exist (ε, δ)-differentially private mechanisms for linear queries that have error $O(\frac{1}{\sqrt{\varepsilon}}\sqrt{n \log |\mathcal{Q}|} \log^{1/4} |U|)$. Moreover, there exist sets of counting queries for which this bound is tight up to factors polylogarithmic in the size of the database [33].

We extend the results from Chapter 7 and show that there exists an efficient (ε, δ)-differentially private mechanism whose L_2 error is not much larger than $\text{opt}^{(2)}_{\varepsilon,\delta}(A, n)$ on *any* database of size at most n. In other words, if we look at $\text{opt}^{(2)}_{\varepsilon,\delta}(A, n)$ as a function of n, the error of our algorithm approximates this function pointwise, while the error of the algorithm of Theorem 7.1 is only guaranteed to be approximately bounded by the least upper bound of $\text{opt}^{(2)}_{\varepsilon,\delta}(A, n)$. This improved guarantee is important, since in some cases $\text{opt}^{(2)}_{\varepsilon,\delta}(A)$ may be larger than the trivial error n. Giving a similar "strongly optimal" guarantee for the worst-case error $\text{opt}_{\varepsilon,\delta}(A, n)$ is an interesting open problem. Our main result for small databases is summarized in the following theorem.

The following theorem is our main result of this chapter and shows the existence of an efficient nearly optimal differentially private algorithm.

Theorem 8.1. *There exists an (ε, δ)-differentially private algorithm $\mathcal{M}$ that runs in time polynomial in $|D|$, $|\mathcal{Q}|$, and $|U|$, and has error*

$$\text{err}_2(\mathcal{M}, n, A) = O((\log n)(\log 1/\delta)^{1/4}(\log |U|)^{1/4}) \cdot \text{opt}^{(2)}_{\varepsilon,\delta}(n, A)$$

for any n, any query matrix $A \in \mathbb{R}^{\mathcal{Q} \times A}$, any small enough ε, and any $\delta = |U|^{-O(1)}$ small enough with respect to ε.

Question 5. *Prove an analogoue of Theorem 8.1 for the worst-case error* $\mathrm{opt}_{\varepsilon,\delta}(A, n)$.

Our lower bound argument for $\mathrm{opt}^{(2)}_{\varepsilon,\delta}(n, A)$ is analogous to the discrepancy-based reconstruction attack argument from Chapter 7. We simply observe that the hereditary vector discrepancy of any submatrix of A of at most n columns provides a lower bound on the optimal error. The more challenging task is to give an algorithm whose error matches this lower bound. We take the generalized Gaussian mechanism as a basis, and again we instantiate it with a minimal ellipsoid, although with respect to a different objective. By itself this mechanism can have error which is too large when the database is small. Nevertheless, in this case we can use the knowledge that the database is small to reduce the error. Taking an idea from statistics, we perform a regression step: we postprocess the vector $\tilde{y}$ of noisy query answers and find the closest vector that is consistent with the database size bound. This post-processing step is a form of sparse regression, and can be posed as a convex optimization problem using the sensitivity polytope. Indeed, nK_A is easily seen to contain the convex hull of the vectors of query answers produced by databases of size at most n. So we simply need to project $\tilde{y}$ onto nK_A. (In fact our estimator is slightly more complicated and related to the hybrid estimator of Li [155]). Intuitively, when n is small compared to the number of queries, nK_A is small enough that projection cancels the excess error.

8.2 Error Lower Bounds with Small Databases

In this section we discuss how to adapt our lower bounds on the error of differentially private mechanism, so that they hold even when the input is the histogram of a small database. This does not involve any new techniques as much as making observations about the proofs we have already given.

First we give a reformulation of Theorem 7.2. Recall that we use the notation $\mathrm{herdisc}(s, A)$ to denote the maximum discrepancy of all submatrices of A with at most s columns. In this chapter it will be convenient to consider the vector discrepancy

relaxation of the L_2 version of this quantity. Define the L_2 vector discrepancy of a matrix $A \in \mathbb{R}^{m\times n}$ as

$$\operatorname{vecdisc}_2(A) \triangleq \min_{u_1,\ldots,u_n\in\mathbb{S}^{n-1}} \Big(\frac{1}{m}\sum_{i=1}^{m} \big\|\sum_{j=1}^{n} A_{ij}u_j\big\|_2^2\Big)^{1/2}.$$

An analogous argument to the proof of Proposition 3.2 establishes the dual formulation

$$\operatorname{vecdisc}_2(A)^2 = \max \operatorname{tr}(Q) \tag{8.1}$$

$$\text{s.t.}$$

$$Q \preceq \frac{1}{m}A^\intercal A, \tag{8.2}$$

$$Q \text{ diagonal }. \tag{8.3}$$

Now we define the s-hereditary vector discrepancy of A as

$$\operatorname{hvdisc}_2(n, A) \triangleq \max_{J\subseteq[m]:|J|\le s} \operatorname{vecdisc}_2(A_J).$$

The reformulation of Theorem 7.2 follows from the following modification of Lemma 7.4. The proof is exactly analogous to the original proof and we omit it.

Lemma 8.1. *Let $A \in \mathbb{R}^{\mathcal{Q}\times U}$ be a query matrix, let $W \subseteq U$, $|W| \le s$, be such that $\operatorname{vecdisc}_2(A_W) = \operatorname{hvdisc}_2(s, A)$, and define $X \triangleq \{x \in \{0,1\}^U : x_i = 0\ \forall e \in U \setminus W\}$. Let $\mathcal{M}$ be a mechanism such that $\operatorname{err}_2(\mathcal{M}, A, X) \le \alpha \operatorname{hvdisc}_2(s, A)$. Then, there exists an assignment $q : W \to \mathbb{R}$ of non-negative reals to W, and a deterministic algorithm $\mathcal{R}$ with range $\{0,1\}^U$ such that, for any x supported on W*

$$\mathbb{E}\sqrt{\frac{1}{q(W)}\sum_{e\in W} q(e)(\tilde{x}_e - x_e)^2} \le 2\alpha,$$

where $\tilde{x} \triangleq \mathcal{R}(\mathcal{M}(x))$, $q(W) \triangleq \sum_{e\in W} q(e)$, and the expectation is taken over the randomness of $\mathcal{M}$.

We can now state the theorem.

Theorem 8.2. *There exists a constant c, such that for any query matrix $A \in \mathbb{R}^{\mathcal{Q}\times U}$ we have*

$$\operatorname{opt}^{(2)}_{\varepsilon,\delta}(n, A) \ge \frac{c}{\varepsilon}\operatorname{hvdisc}_2(\varepsilon n, A),$$

for all small enough ε and any δ sufficiently small with respect to ε.

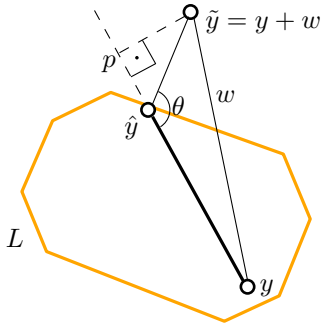


Figure 8.1: A schematic illustration of the key step of the proof of Lemma 8.2. The vector $p - y$ is proportional in length to $|\langle \hat{y} - y, w \rangle|$ and the vector $\hat{y} - p$ is proportional in length to $|\langle \hat{y} - y, \hat{y} - \tilde{y} \rangle|$. Since the angle θ is obtuse, $\|p - y\|_2 \geq \|\hat{y} - p\|_2$.

Proof. Observe that X as defined in Lemma 8.1 satisfies $X \subseteq sB_1^U$, i.e. is a set of databases of size at most s. Using Lemma 7.3 and Lemma 8.1 with $s \triangleq \varepsilon n$, we can use an argument analogous to the one in the proof of Theorem 7.2 to conclude that $\mathrm{opt}^{(2)}_{\varepsilon_0,\delta_0}(\varepsilon n, A) \geq \frac{1}{2(1+e)} \mathrm{hvdisc}_2(\varepsilon n, A)$ for small enough ε and δ. To finish the proof we appeal to Lemma 7.2 to show that $\mathrm{opt}^{(2)}_{\varepsilon,\delta}(n, A) \geq \lfloor 1/\varepsilon \rfloor \, \mathrm{opt}^{(2)}_{\varepsilon_0,\delta_0}(n, A)$. □

8.3 The Projection Mechanism

A key element in our algorithms for the small database case is the use of least squares estimation to reduce error. In this section we introduce and analyze a mechanism based on least squares estimation, similar to the hybrid estimator of [155].

8.3.1 Projection to a Convex Body

Below we present a bound on the error of least squares estimation with respect to symmetric convex bodies. This analysis appears to be standard in the statistics literature; a special case of it appears for example in [123].

For the analysis we will need to recall Hölder's inequality for general norms. If L is a convex body, and L° is its polar body, then for any x and y we have $|\langle x, y \rangle| \leq \|x\|_L \|y\|_{L^\circ}$.

Lemma 8.2. *Let $L \subseteq \mathbb{R}^m$ be a symmetric convex body, let $y \in L, \tilde{y} \in \mathbb{R}^m$, and define $w \triangleq \tilde{y} - y$. Let, finally, $\bar{y} \in L$ be such that $\|\bar{y} - \tilde{y}\|_2^2 \leq \min\{\|z - \tilde{y}\|_2^2 : z \in L\} + \nu$ for*

some $\nu \geq 0$. *We have* $\|\bar{y} - y\|_2^2 \leq \min\{(2\|w\|_2 + \sqrt{\nu})^2, 4\|w\|_{L^\circ} + \nu\}$.

Proof. Let $\hat{y} \triangleq \arg\min\{\|z - \tilde{y}\|_2^2 : z \in L\}$. First we show the easier bound: by the triangle inequality,

$$\|\bar{y} - y\|_2 \leq \|\bar{y} - \tilde{y}\|_2 + \|\tilde{y} - y\|_2 \leq 2\|\tilde{y} - y\|_2 + \sqrt{\nu}.$$

The last inequality above follows from

$$\|\bar{y} - \tilde{y}\|_2 \leq \sqrt{\|\hat{y} - \tilde{y}\|_2^2 + \nu} \leq \|\hat{y} - \tilde{y}\|_2 + \sqrt{\nu} \leq \|y - \tilde{y}\|_2 + \sqrt{\nu}.$$

The bound $\|\bar{y} - y\|_2^2 \leq 4\|w\|_{L^\circ} + \nu$ is based on Hölder's inequality and the following simple but very useful fact, illustrated schematically in Figure 8.1:

$$\begin{aligned} \|\bar{y} - y\|_2^2 &= \langle \bar{y} - y, \tilde{y} - y \rangle + \langle \bar{y} - y, \bar{y} - \tilde{y} \rangle \\ &\leq 2\langle \bar{y} - y, \tilde{y} - y \rangle + \nu. \end{aligned} \tag{8.4}$$

The inequality (8.4) can be proved algebraically:

$$\begin{aligned} \langle \bar{y} - y, \tilde{y} - y \rangle &= \|\tilde{y} - y\|_2^2 + \langle \bar{y} - \tilde{y}, \tilde{y} - y \rangle \\ &\geq \|\bar{y} - \tilde{y}\|_2^2 - \nu + \langle \bar{y} - \tilde{y}, \tilde{y} - y \rangle \\ &= \langle \bar{y} - \tilde{y}, \bar{y} - y \rangle - \nu = \langle \bar{y} - y, \bar{y} - \tilde{y} \rangle - \nu. \end{aligned}$$

Inequality (8.4), $w = \tilde{y} - y$, Hölder's inequality and the triangle inequality imply

$$\|\bar{y} - y\|_2^2 \leq 2\langle \bar{y} - y, w \rangle + \nu \leq 2\|\bar{y} - y\|_L \|w\|_{L^\circ} \leq 4\|w\|_{L^\circ} + \nu,$$

which completes the proof. □

8.3.2 The Mechanism

Lemma 8.2 is the key ingredient in the analysis of the Projection Mechanism, presented as Algorithm 3. This mechanism gives improved L_2 error with respect to the generalized Gaussian mechanism $\mathcal{M}_E$ when the the database size n is smaller than the number of queries: the error is bounded from above roughly by the square root of the sum of squared lengths of the n longest major axes of E.

Algorithm 3 Projection Mechanism $\mathcal{M}_E^{\text{proj}}$

Input: *(Public)* Query matrix A; ellipsoid $E = F \cdot B_2^{\mathcal{Q}}$ such that all columns of A are contained in E.

Input: *(Private)* Histogram x of a database of size $\|x\|_1 \le n$.

1: Run the generalized Gaussian mechanism (Algorithm 2) to compute $\tilde{y} \triangleq \mathcal{M}_E(A, x)$;

2: Let Π be the orthogonal projection operator onto the span of the $\lfloor \varepsilon n \rfloor$ largest major axes of E (equivalently the span of leading $\lfloor \varepsilon n \rfloor$ left singular vectors of F);

3: Compute $\bar{y} \in n(I - \Pi)K_A$, where K_A is the sensitivity polytope of A, and $\bar{y}$ satisfies

$$\|\bar{y} - (I - \Pi)\tilde{y}\|_2^2 \le \min\{\|z - (I - \Pi)\tilde{y}\|_2^2 : z \in n(I - \Pi)K_A\} + \nu,$$

and $\nu \le n c_{\varepsilon,\delta} \sqrt{\log |U|} \|(I - \Pi)A\|_{1\to 2}^2$;

Output: Vector of answers $\Pi\tilde{y} + \bar{y}$.

Lemma 8.3. *The Projection Mechanism $\mathcal{M}_E^{proj}$ in Algorithm 3 is (ε, δ)-differentially private for any ellipsoid $E = FB_2^{\mathcal{Q}}$ that contains the columns of A. Moreover, for $\varepsilon = O(1)$,*

$$\text{err}_2(\mathcal{M}_E^{proj}, n, A) = O\left(c_{\varepsilon,\delta}\Big(1 + \frac{\sqrt{\log |U|}}{\sqrt{\log 1/\delta}}\Big)^{1/2} \right) \cdot \Big(\frac{1}{|\mathcal{Q}|} \sum_{i \le \varepsilon n} \sigma_i^2\Big)^{1/2},$$

where $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_{|\mathcal{Q}|}$ are the singular values of F.

Proof. To prove the privacy guarantee, observe that the output of $\mathcal{M}_E^{\text{proj}}(A, x)$ is just a post-processing of the output of $\mathcal{M}_E(A, x)$, i.e. the algorithm does not access x except to pass it to $\mathcal{M}_E(A, x)$. The privacy then follows from Lemmas 7.5 and 7.1.

Next we bound the error. Let $w \triangleq \tilde{y} - y$ be the random noise introduced by the generalized Gaussian mechanism. Recall that w is distributed identically to Fg, where $g \sim N(0, c_{\varepsilon,\delta}^2)^{\mathcal{Q}}$. By the Pythagorean theorem and linearity of expectation we have

$$\mathbb{E}\|\Pi\tilde{y} + \bar{y} - y\|_2^2 = \mathbb{E}\|\Pi\tilde{y} - \Pi y\|_2^2 + \mathbb{E}\|\bar{y} - (I - \Pi)y\|_2^2.$$

Above and in the remainder of the proof the expectations are taken is with respect to the randomness of the choice of w. We bound the two terms on the right hand side separately. For the first term, observe that $\Pi\tilde{y} - \Pi y = \Pi w$ is distributed identically to $\Pi F g$, with g distributed as above. Since, by the definition of Π, the non-zero singular values of ΠF are $\sigma_1, \ldots, \sigma_k$ where $k \triangleq \lfloor \varepsilon n \rfloor$, we have

$$\mathbb{E}\|\Pi\tilde{y} - \Pi y\|_2^2 = \mathbb{E}\text{tr}(\Pi F g g^\intercal F^\intercal \Pi) = c_{\varepsilon,\delta}^2 \text{tr}(\Pi F F^\intercal \Pi) = c_{\varepsilon,\delta}^2 \sum_{i=1}^{k} \sigma_i^2.$$

To bound the second term we appeal to Lemma 8.2. Define $\tilde{K} \triangleq (I-\Pi)K_A$. With $n\tilde{K}$ in the place of L, the lemma implies that

$$\mathbb{E}\|\bar{y} - (I-\Pi)y\|_2^2 \le 4\mathbb{E}\|(I-\Pi)w\|_{(n\tilde{K})^\circ} + \nu \le 4n\mathbb{E}\|(I-\Pi)w\|_{\tilde{K}^\circ} + \nu, \tag{8.5}$$

where we used the simple fact

$$\|(I-\Pi)w\|_{(n\tilde{K})^\circ} = \sup_{z \in n\tilde{K}} \langle (I-\Pi)w, z\rangle = n \sup_{z\in\tilde{K}} \langle (I-\Pi)w, z\rangle = n\|(I-\Pi)w\|_{\tilde{K}^\circ}.$$

Since $\tilde{K} \subseteq (I-\Pi)E$, and $(I-\Pi)E$ is contained in a Euclidean ball of radius bounded above by $\sigma_{k+1} \le \sigma_k$ by the choice of Π, we have that any point $z \in \tilde{K}$ has length bounded as $\|z\|_2 \le \sigma_k$. Moreover, $\tilde{K}$ is the convex hull of at most $N \le 2|U|$ vertices: it is the convex hull of the $2|U|$ vertices of K_A (the columns of A and $-A$) projected by the operator $I-\Pi$. Call these vertices $z_1, \ldots, z_N$. Since a linear functional is always maximized at a vertex of a polytope, we have $\|(I-\Pi)w\|_{\tilde{K}^\circ} = \sup_{z\in\tilde{K}} \langle (I-\Pi)w, z\rangle = \max_{i=1}^N \langle (I-\Pi)w, z_i\rangle$. Each inner product $\langle (I-\Pi)w, z_i\rangle$ is a zero mean Gaussian random variable with variance

$$\mathbb{E}\langle (I-\Pi)w, z_i\rangle^2 = z_i^\intercal (I-\Pi)\mathbb{E}[ww^\intercal](I-\Pi)z_i = c_{\varepsilon,\delta}^2 z_i^\intercal (I-\Pi)FF^\intercal(I-\Pi)z_i.$$

By the choice of Π, the largest singular value of $(I-\Pi)FF^\intercal(I-\Pi)$ is $\sigma_{k+1} \le \sigma_k$. Therefore, since the Euclidean norm of z_i is also at most σ_k, we have that the variance of $\langle (I-\Pi)w, z_i\rangle$ is at most $c_{\varepsilon,\delta}^2\sigma_k^4$. By an argument analogous to the one in the proof of Theorem 7.1, we can bound the expectation of the maximum of the inner products as

$$\mathbb{E}\|(I-\Pi)w\|_{\tilde{K}^\circ} = \mathbb{E}\max_{i=1}^N \langle (I-\Pi)w, z_i\rangle = O(\sqrt{\log N})c_{\varepsilon,\delta}\sigma_k^2.$$

Plugging this into (8.5) and using that $\|(I-\Pi)A\|_{1\to 2} = \max_{i=1}^N \|z_i\|_2 \le \sigma_k$, we get

$$\mathbb{E}\|\bar{y} - (I-\Pi)y\|_2^2 = O(\sqrt{\log N})c_{\varepsilon,\delta} n\sigma_k^2.$$

Observe that $c_{\varepsilon,\delta}n\sigma_k^2 \le \frac{c_{\varepsilon,\delta}n}{k}\sum_{i=1}^k \sigma_i^2$. Since $k = \lfloor \varepsilon n \rfloor$, $\frac{c_{\varepsilon,\delta}n}{k} = O\Big(\frac{c_{\varepsilon,\delta}^2}{\sqrt{\log 1/\delta}}\Big)$. This finishes the proof. □

8.3.3 Efficient Implementation: Frank-Wolfe

Computing $\bar{y}$ in Algorithm 3 requires approximately solving a convex optimization problem. Any standard tool for convex optimization, such as the ellipsoid algorithm

can be used. We recall an algorithm of Frank and Wolfe which has slower convergence rate than the ellipsoid method, but may be more practical since we only require a very rough approximation. Moreover, the algorithm allows reducing the problem to solving linear programs over $(I - \Pi)K_A$. The algorithm is presented as Algorithm 4.

Algorithm 4 Frank-Wolfe Algorithm

Input: convex body $L \subseteq \mathbb{R}^m$; point $r \in \mathbb{R}^m$; number of iterations T
  Let $q^{(0)} \in L$ be arbitrary.
  for $t = 1$ **to** T **do**
    Let $v^{(t)} = \arg\max_{v \in L} \langle r - q^{(t-1)}, v \rangle$.
    Let $\alpha^{(t)} = \arg\min_{\alpha \in [0,1]} \|r - \alpha q^{(t-1)} - (1-\alpha)v^{(t)}\|_2^2$.
    Set $q^{(t)} = \alpha^{(t)} q^{(t-1)} + (1 - \alpha^{(t)})v^{(t)}$.
  end for
  Output $q^{(T)}$.

The expensive step in each iteration of Algorithm 4 is computing $v^{(t)}$, which requires solving a linear optimization problem over L. Computing $\alpha^{(t)}$ is a quadratic optimization problem in a single variable, and has a closed form solution.

We use the following bound on the convergence rate of the Frank-Wolfe algorithm. It is a refinement of the original analysis of Frank and Wolfe, due to Clarkson.

Theorem 8.3 ([63, 45])**.** *The point $q^{(T)}$ computed by T iterations of Algorithm 4 satisfies*

$$\|r - q^{(T)}\|_2^2 \leq \min\{\|r - q\|_2^2 : q \in L\} + \frac{4\mathrm{diam}(L)^2}{T+3}.$$

In Algorithm 3, we can apply the Frank-Wolfe algorithm to the body $L = n(I - \Pi)K_A$ and the point $r = (I - \Pi)\tilde{y}$. The diameter of L is at most $n\|(I - \Pi)A\|_{1\to 2}$, so to achieve the required approximation ν we need to set the number of iterations T to $4\frac{n}{c_{\varepsilon,\delta}\sqrt{\log|U|}}$.

Another useful feature of the Frank-Wolfe algorithm is that $q^{(T)}$ is in the convex hull of $v^{(0)}, \ldots, v^{(T)}$, which allows for a concise representation of its output.

8.4 Optimality of the Projection Mechanism

In this section we show that we can choose an ellipsoid E so that $\mathcal{M}_E^{\mathrm{proj}}$ has nearly optimal error. Once again we optimize over ellipsoids and use convex duality and

the restricted invertibility principle to relate the optimal ellipsoid to the appropriate notion of discrepancy, which itself bounds from below the error necessary for privacy. The optimization problem over ellipsoids is different, but closely related, to the one used to define the ellipsoid infinity norm.

8.4.1 Minimizing Ky Fan Norm over Containing Ellipsoids

Given an ellipsoid $E = FB_2^m$, define $f_k(E) = \left(\sum_{i=1}^k \sigma_i^2\right)^{1/2}$, where $\sigma_1 \geq \ldots \geq \sigma_m$ are the singular values of F. Define $\|M\|_{(k)}$ to be the Ky Fan k-norm, i.e. the sum of the top k singular values of M. The already familiar nuclear norm $\|M\|_{S_1}$ is equal to $\|M\|_{(r)}$ where r is the rank of M. An equivalent way to define $f_k(E)$ then is as $f_k(E) \triangleq \|FF^\intercal\|_k^{1/2}$.

The ellipsoid we use in the projection mechanism will be the one achieving $\min\{f_k(E) : a_e \in E \ \forall e \in U\}$, where a_e is the column of the query matrix A associated with the universe element e. This choice is directly motivated by Lemma 8.3. We can write this optimization problem in the following way.

$$\text{Minimize } \|X^{-1}\|_{(k)} \text{ s.t.} \tag{8.6}$$

$$X \succ 0 \tag{8.7}$$

$$\forall e \in U : a_e^\intercal X a_e \leq 1. \tag{8.8}$$

To show that the above program is convex we will need the following well-known result of Fan.

Lemma 8.4 ([62])**.** *For any $m \times m$ real symmetric matrix M,*

$$\|M\|_{(k)} = \max_{U \in \mathbb{R}^{m \times k} : U^\intercal U = I} \operatorname{tr}(U^\intercal M U).$$

With this result in hand, we can prove that (8.6)–(8.8) captures the optimization problem we are after analogously to the proof of Lemma 4.6.

Lemma 8.5. *For a rank $|\mathcal{Q}|$ query matrix $A = (a_e)_{e \in U} \in \mathbb{R}^{\mathcal{Q} \times univ}$, the optimal value of the optimization problem* (8.6)–(8.8) *is equal to* $\min\{f_k(E)^2 : a_e \in E \ \forall e \in U\}$. *Moreover, the objective function* (8.6) *and constraints* (8.8) *are convex over $X \succ 0$.*

Proof. Let λ be the optimal value of (8.6)–(8.8) and let $\mu = \min\{f_k(E)^2 : a_e \in E \ \forall e \in U\}$. Given a feasible X for (8.6)–(8.8), set $E = X^{-1/2}B_2^{\mathcal{Q}}$ (this is well-defined since $X \succ 0$). Then for any $j \in [n]$, $\|a_j\|_E = a_j^\intercal X a_j \leq 1$ by (4.8), and, therefore, $a_j \in E$. Also, by (4.4), $f_k(E)^2 = \|X^{-1}\|_{(k)}$ by definition. This shows that $\mu \leq \lambda$. In the reverse direction, let $E = FB_2^{\mathcal{Q}}$ be such that $\forall e \in U : a_e \in E$. Then, because A is full rank, F is also full rank and invertible, and we can define $X = (FF^\intercal)^{-1}$. Analogously to the calculations above, we can show that X is feasible, and therefore $\lambda \leq \mu$.

The objective function and the constraints (8.8) are affine, and therefore convex. It remains to show that the objective (8.6) is also convex. Let X_1 and X_2 be two feasible solutions and define $Y = \alpha X_1 + (1-\alpha)X_2$ for some $\alpha \in [0,1]$. By Lemma 4.5, $Y^{-1} \preceq \alpha X_1^{-1} + (1-\alpha)X_2^{-1}$. Let U be such that $\mathrm{tr}(U^\intercal Y^{-1} U) = \|Y^{-1}\|_{(k)}$ and $U^\intercal U = I$; then, by Lemma 8.4

$$\begin{aligned}\|Y^{-1}\|_{(k)} = \mathrm{tr}(U^\intercal Y^{-1} U) &\leq \alpha \mathrm{tr}(U^\intercal X_1^{-1} U) + (1-\alpha)\mathrm{tr}(U^\intercal X_2^{-1} U) \\ &\leq \alpha \|X_1^{-1}\|_{(k)} + (1-\alpha)\|X_2^{-1}\|_{(k)}.\end{aligned}$$

This finishes the proof. □

Since the program (8.6)–(8.8) is convex, its optimal solution can be approximated to any given degree of accuracy using the ellipsoid algorithm [72]. Analogously to the ellipsoid infinity norm, we can define the *ellipsoid Ky Fan k norm* of an $m \times n$ matrix $A = (a_i)_{i=1}^n$ by $\|A\|_{E(k)} \triangleq \min\{f_k(E) : a_i \in E \ \forall i \in [n]\}$. An argument analogous to the one in the proof of Lemma 4.4 using Lemma 8.4 proves that $\mathrm{herdisc}_2(s, A) \leq O(1)\|A\|_{E(s)}$. We shall not pursue this direction further here.

8.4.2 The Dual of the Ellipsoid Problem

Our next goal is derive a dual characterization of (8.6)–(8.8). Before we can do that, we need to define a somewhat complicated function of the singular values of a matrix. The next lemma is needed to argue that this function is well-defined.

Lemma 8.6. *Let $\sigma_1 \geq \ldots \sigma_m \geq 0$ be non-negative reals, and let $k \leq m$ be a positive*

integer. There exists an integer t, $0 \leq t \leq k-1$, such that

$$\sigma_t > \frac{\sum_{i>t} \sigma_i}{k-t} \geq \sigma_{t+1}, \tag{8.9}$$

with the convention $\sigma_0 = \infty$.

Proof. Define $\sigma_{>t} \triangleq \sum_{i>t} \sigma_i$. If $\sigma_{>0} \geq k\sigma_1$ holds, then (8.9) is satisfied for $t = 0$, and we are done. So let us assume that $\sigma_{>0} < k\sigma_1$. Then $\sigma_{>1} = \sigma_{>0} - \sigma_1 < (k-1)\sigma_1$, and the first inequality in (8.9) is satisfied for $t = 1$. If the second inequality is also satisfied we are done, so let us assume that $\sigma_{>1} < (k-1)\sigma_2$, which implies the first inequality in (8.9) for $t = 2$. Continuing in this manner, we see that if the inequalities (8.9) are not satisfied for any $t \in \{0, \ldots, k-2\}$, then we must have $\sigma_{>k-1} < \sigma_{k-1}$. But the second inequality for $t = k-1$, i.e. $\sigma_{>k-1} \geq \sigma_k$ is always satisfied because all the σ_i are non-negative, so we have that if (8.9) does not hold for $t \leq k-2$, then it must hold for $t = k-1$. This finishes the proof. □

We now introduce a function which will be used in formulating a dual characterization of (8.6)–(8.8).

Definition 8.1. *Let $M \succeq 0$ be an $m \times m$ positive semidefinite matrix with singular values $\sigma_1 \geq \ldots \geq \sigma_m$, and let $k \leq m$ be a positive integer. The function $h_k(M)$ is defined as*

$$h_k(M) \triangleq \sum_{i=1}^{t} \sigma_i^{1/2} + \sqrt{k-t} \left(\sum_{i>t} \sigma_i \right)^{1/2},$$

where t is the smallest integer such that $\sigma_t > \frac{\sum_{i>t} \sigma_i}{k-t} \geq \sigma_{t+1}$.

Lemma 8.6 guarantees that $h_k(M)$ is a well-defined real-valued function. The next lemma shows that it is a continuous function.

Lemma 8.7. *The function h_k is continuous over positive semidefinite matrices with respect to the operator norm.*

Proof. Let M be a $m \times m$ positive semidefinite matrix with singular values $\sigma_1 \geq \ldots \geq \sigma_m$ and let t be the smallest integer so that $\sigma_t > \frac{\sum_{i>t} \sigma_i}{k-t} \geq \sigma_{t+1}$. If $\frac{\sum_{i>t} \sigma_i}{k-t} > \sigma_{t+1}$, then setting δ small enough ensures that, for any M' such that $\|M - M'\|_2 < \delta$, $h_k(M)$ and

$h_k(M')$ are computed with the same value of t, in which case the proof of continuity follows from the continuity of the square root function. Let us therefore assume that $\frac{\sum_{i>t}\sigma_i}{k-t} = \sigma_{t+1} = \ldots = \sigma_{t'} > \sigma_{t'+1}$ for some $t' \geq t+1$. Then for any integer $s \in [t, t']$,

$$\sum_{i>s} \sigma_i = \sum_{i>t} \sigma_i - (s-t)\sigma_{t+1} = (k-s)\sigma_{t+1}.$$

We then have

$$\begin{aligned}
\sum_{i=1}^{t} \sigma_i^{1/2} + \sqrt{k-t}\left(\sum_{i>t}\sigma_i\right)^{1/2} &= \sum_{i=1}^{t} \sigma_i^{1/2} + (k-t)\sigma_{t+1}^{1/2} \\
&= \sum_{i=1}^{s} \sigma_i^{1/2} + (k-s)\sigma_{t+1}^{1/2} \\
&= \sum_{i=1}^{s} \sigma_i^{1/2} + \sqrt{k-s}\left(\sum_{i>s}\sigma_i\right)^{1/2} \qquad (8.10)
\end{aligned}$$

For any M' such that $\|M' - M\|_2 < \delta$ for a small enough δ, we have

$$h_k(M') = \sum_{i=1}^{s} \sigma_i(M')^{1/2} + \sqrt{k-s}\left(\sum_{i>s}\sigma_i(M')\right)^{1/2},$$

where s is an integer in $[t, t']$. Continuity then follows from (8.10), and the continuity of the square root function. □

Since the objective of (8.6)–(8.8) is not necessarily differentiable, in order to analyze the dual we need to recall the concepts of subgradients and subdifferentials. A *subgradient* of a function $f\colon S \to \mathbb{R}$ at $x \in S$, where S is some open subset of $\mathbb{R}^d$, is a vector $y \in \mathbb{R}^d$ so that for every $z \in S$ we have

$$f(z) \geq f(x) + \langle x - z, y\rangle.$$

The set of subgradients of f at x is denoted $\partial f(x)$ and is known as the *subdifferential*. When f is differentiable at x, the subdifferential is a singleton set containing only the gradient $\nabla f(x)$. If f is defined by $f(x) = f_1(x) + f_2(x)$, where $f_1, f_2\colon S \to \mathbb{R}$, then $\partial f(x) = \partial f_1(x) + \partial f_2(x)$. A basic fact in convex analysis is that f achieves its minimum at x if and only if $0 \in \partial f(x)$. For more information on subgradients and subdifferentials, see the classical text of Rockafellar [124].

Overton and Womersley [120] analyzed the subgradients of functions which are a composition of a differentiable matrix-valued function with a Ky Fan norm. The special case we need also follows from the results of Lewis [90].

Lemma 8.8 ([120],[90])**.** *Let $g_k(X) \triangleq \|X^{-1}\|_{(k)}$ for a positive definite matrix $X \in \mathbb{R}^{m\times m}$. Let $\sigma_1 \geq \ldots \geq \sigma_m$ be the singular values of X^{-1} and let Σ be the diagonal matrix with the σ_i on the diagonal. Assume that for some $r \geq k$, $\sigma_k = \ldots = \sigma_r$. Then the subgradients of g_k are given by*

$$\partial g_k(X) = \mathrm{conv}\{U_S U_S^\intercal X^{-2} U_S U_S^\intercal : U \, orthogonal, U\Sigma U^\intercal = X^{-1}, S \subseteq [r]\},$$

where U_S is the submatrix of U indexed by S.

We use the following well-known characterization of the convex hull of boolean vectors of Hamming weight k.

Lemma 8.9. *Let $V_{k,n} \triangleq \mathrm{conv}\{v \in \{0,1\}^n : \|v\|_1 = k\}$. Then $V_{k,n} = \{v : \|v\|_1 = k, 0 \leq v_i \leq 1 \ \forall i\}$.*

Proof. Let $P_{k,n}$ be the polytope $\{v : \|v\|_1 = k, 0 \leq v_i \leq 1 \ \forall i\}$. We need to show that $V_{k,n} = P_{k,n}$. The containment $V_{k,n} \subseteq P_{k,n}$ is easy to verify, as all extreme points of $V_{k,n}$ satisfy the constraints defining $P_{k,n}$. In the other direction, observe that $P_{k,n} = \{v : b \leq Av \leq c\}$, where b and c are vectors with integer coordinates, and A is a matrix with first row the all ones vector, followed by the $n \times n$ identity matrix I. It is easy to verify that A is totally unimodular, either directly or by observing that $\mathrm{herdisc}(A) = 1$, which is equivalent to total unimodularity by the Ghouila-Houri characterization [66]. It follows that $P_{k,n}$ is a polytope with integral vertices, and it is easy to verify that any integral point in $P_{k,n}$ is a boolean vector of Hamming weight k, and, therefore, lies in $V_{k,n}$. Then $P_{k,n} \subseteq V_{k,n}$, implying $P_{k,n} = V_{k,n}$, as desired.

This characterization of $V_{k,n}$ is a part of the more general theory of basis polytopes of matroids. In particular, $V_{k,n}$ is the basis polytope of the rank k uniform matroid. For more details, see [132]. □

Before we give our dual characterization, we need one more technical lemma.

Lemma 8.10. *Let M be an $m \times m$ positive semidefinite matrix of rank at least k. Then there exists an $m \times m$ positive definite matrix X such that $M \in \partial g_k(X)$, and $\|X^{-1}\|_{(k)} = g_k(X) = h_k(M)$.*

Proof. Let $r = \text{rank } M$, and let $\sigma_1 \geq \ldots \geq \sigma_r$ be the non-zero singular values of M. Let $U\Sigma U^\intercal = M$ be some singular value decomposition of M: U is an orthonormal matrix and Σ is a diagonal matrix with the σ_i on the diagonal, followed by 0s.

Assume that t is the smallest integer such that $\sigma_t > \frac{\sum_{i>t}\sigma_i}{k-t} \geq \sigma_{t+1}$ and define $\alpha \triangleq \frac{\sum_{i>t}\sigma_i}{k-t}$. A choice of $t \leq k-1$ exists by Lemma 8.6. Let the diagonal matrix Σ' be defined by

$$\sigma'_{ii} \triangleq \begin{cases} \sigma_i & i \leq t \\ \alpha & t < i \leq r \\ \alpha - \epsilon & i > r \end{cases}.$$

We set ϵ to be an arbitrary number satisfying $\alpha > \epsilon > 0$. Let us set $X \triangleq (U\Sigma' U^\intercal)^{-1/2}$. By Lemma 8.9 and and the choice of t, the vector $(\sigma_{t+1}, \ldots, \sigma_r)$ is an element of the polytope $\alpha V_{k-t,r}$. Then M is an element of $\text{conv}\{U_S U_S^\intercal X^{-2} U_S U_S^\intercal : S = [t] \cup T, T \subseteq \{t+1, \ldots, r\}, |T| = k-t\}$. Since this set is a subset of $\partial g_k(X)$, we have $M \in \partial g_k(X)$. A calculation shows that $\|X^{-1}\|_{(k)} = \|(U\Sigma' U^\intercal)^{1/2}\|_{(k)} = \sum_{i \leq t} \sigma_i^{1/2} + (k-t)\alpha^{1/2} = h_k(M)$. This completes the proof.

□

The following theorem is our dual characterization of (8.6)–(8.8).

Theorem 8.4. *Let $A = (a_e)_{e \in U} \in \mathbb{R}^{\mathcal{Q} \times U}$ be a rank $|\mathcal{Q}|$ matrix, and let $\mu = \min\{f_k(E) : a_e \in E \ \forall e \in U\}$. Then,*

$$\mu^2 = \max h_k(AQA^\intercal)^2 \ \ s.t. \tag{8.11}$$

$$Q \succeq 0, \ \ \textit{diagonal}, \text{tr}(Q) = 1 \tag{8.12}$$

Proof. The proof of this theorem is similar to the proof of Theorem 4.9. Let us define $\{X : X \succ 0\}$ to be the domain for the constraints (8.8) and the objective function (8.6). This makes the constraint $X \succ 0$ implicit. The optimization problem is convex

by Lemma 8.5. Is is also always feasible: for example for $r = \|A\|_{1\to 2}$, $\frac{1}{r}I$ is a feasible solution. Slater's condition is therefore satisfied, since the constraints are affine, and strong duality holds.

The Lagrange dual function for (8.6)–(8.8) is

$$g(p) = \inf_{X\succ 0} \|X^{-1}\|_{(k)} + \sum_{e\in U} p_e(a_e^\intercal X a_e - 1),$$

with dual variables $p \in \mathbb{R}^U$, $p \geq 0$. Equivalently, writing p as a diagonal matrix $P \in \mathbb{R}^{U\times U}$, $P \succeq 0$, with entries $p_{ee} = p_e$, we have

$$g(P) = \inf_{X\succ 0} \|X^{-1}\|_{(k)} + \mathrm{tr}(APA^\intercal X) - \mathrm{tr}(P) \tag{8.13}$$

Since $X \succ 0$ implies $X^{-1} \succ 0$, $g(P) \geq -\mathrm{tr}(P) > -\infty$. Therefore, the effective domain $\{P : g(P) > -\infty\}$ of $g(P)$ is $\{P : P \succeq 0,\ \text{diagonal}\}$. Since we have strong duality, and, by Lemma 8.5, μ^2 is equal to the optimal value of (8.6)–(8.8), we have $\mu^2 = \max\{g(P) : P \succeq 0,\ \text{diagonal}\}$.

By the additivity of subgradients, a matrix X achieves the minimum in (8.13) if and only if $APA^\intercal \in \partial g_k(X)$, where $g_k(X) = \|X^{-1}\|_{(k)}$. Consider first the case in which $APA^\intercal$ has rank at least k. Then, by Lemma 8.10, there exists an X such that $APA^\intercal \in \partial g_k(X)$ and $\|X^{-1}\|_{(k)} = h_k(APA^\intercal)$. Observe that, if U is an $m \times k$ matrix such that $U^\intercal U = I$ and $\mathrm{tr}(U^\intercal X^{-1} U) = \|X^{-1}\|_{(k)}$, then

$$\mathrm{tr}(UU^\intercal X^{-2} UU^\intercal X) = \mathrm{tr}((U^\intercal X^{-2} U)(U^\intercal X U)) = \mathrm{tr}(U^\intercal X^{-1} U) = \|X^{-1}\|_{(k)}.$$

Since $APA^\intercal$ is a convex combination of matrices $UU^\intercal X^{-2} UU^\intercal$ for U as above, it follows that $\mathrm{tr}(APA^\intercal X) = \|X^{-1}\|_{(k)}$. Then we have

$$\begin{aligned} g(P) &= \|X^{-1}\|_{(k)} + \mathrm{tr}(APA^\intercal X) - \mathrm{tr}(P) \\ &= 2\|X^{-1}\|_{(k)} - \mathrm{tr}(P) = 2h_k(APA^\intercal) - \mathrm{tr}(P). \end{aligned} \tag{8.14}$$

If P is such that $APA^\intercal$ has rank less than k, we can reduce to the rank k case by a continuity argument as in the proof of Theorem 4.9. Fix any non-negative diagonal matrix P and for $\lambda \in [0,1]$ define $P(\lambda) \triangleq \lambda P + (1-\lambda)I$. For any $\lambda \in [0,1)$, $AP(\lambda)A^\intercal$

has rank $|\mathcal{Q}|$, since $AA^\intercal$ has rank $|\mathcal{Q}|$ by assumption, and, therefore, $AP(\lambda)A^\intercal \succeq (1-\lambda)AA^\intercal \succ 0$. Then, by Lemma 4.7 and (8.14), we have

$$\begin{aligned} g(P) = \lim_{\lambda \uparrow 1} g(P(\lambda)) &= \lim_{\lambda \uparrow 1} \left[2h_k(AP(\lambda)^\intercal) - \lambda \mathrm{tr}(P) - (1-\lambda)|U|\right] \\ &= 2h_k(APA^\intercal) - \mathrm{tr}(P). \end{aligned}$$

The final equality follows from the continuity of h_k, proved in Lemma 8.7, and standard perturbation bounds.

Defining new variables Q and c with $c = \mathrm{tr}(P)$, $Q = P/c$, and optimizing over c as in Theorem 4.9 finishes the proof. □

8.4.3 Proof of the Main Theorem

We use the dual formulation in Theorem 8.4 and the restricted invertibility principle to give lower bounds on $\mathrm{hvdisc}_2(s, A)$.

Let us first give a variant of the spectral lower bound for $\mathrm{hvdisc}_2(s, A)$. We define

$$\mathrm{specLB}_2(s, A) \triangleq \max_{k=1}^{s} \max_{J \subseteq [n] : |J| = k} \sqrt{\frac{k}{m}} \sigma_{\min}(A_J).$$

Analogously to Lemma 4.8, it follows from (8.1)–(8.3) that $\mathrm{hvdisc}_2(s, A) \geq \mathrm{specLB}_2(s, A)$.

Proof of Theorem 8.1. Given a database size n and a query matrix A, the near optimal algorithm is the projection algorithm $\mathcal{M}_E^{\mathrm{proj}}$ instantiated with an ellipsoid E that (approximately) achieves $\min\{f_k(E) : a_e \in E \ \forall e \in U\}$ for $k \triangleq \lfloor \varepsilon n \rfloor$, where a_e is the column of A corresponding to the universe element e. By Lemma 8.5, E can be computed by solving the program (8.6)–(8.8), which is a convex minimization problem and can be arbitrarily well approximated using the ellipsoid method [72], or the algorithm of Overton and Womersley [120].

By Lemma 8.3,

$$\mathrm{err}_2(\mathcal{M}_E^{\mathrm{proj}}, n, A) = O\left(c_{\varepsilon,\delta} \left(1 + \frac{\sqrt{\log |U|}}{\sqrt{\log 1/\delta}}\right)^{1/2} \right) \cdot \frac{1}{\sqrt{|\mathcal{Q}|}} f_k(E). \tag{8.15}$$

By Theorem 4.9, the optimal solution Q of (8.11)–(8.12) satisfies

$$f_k(E) = h_k(AQA^\intercal) = \sum_{i=1}^{t} \sigma_i^{1/2} + \sqrt{k-t} \left(\sum_{i>t} \sigma_i \right)^{1/2},$$

where $\sigma_1 \geq \ldots \geq \sigma_m$ are the singular values of $AQA^\intercal$ and t is such that $(k-t)\sigma_t > \sum_{i>t}\sigma_i \geq (k-t)\sigma_{t+1}$. At least one of $\sum_{i=1}^t \sigma_i^{1/2}$ and $\sqrt{k-t}\left(\sum_{i>t}\sigma_i\right)^{1/2}$ must be bounded from below by $\frac{1}{2}f_k(E)$. Next we consider these two cases separately.

Assume first that $\sum_{i=1}^t \sigma_i^{1/2} \geq \frac{1}{2}f_k(E)$. Let Π be the orthogonal projection operator onto the span of the singular vectors of $AQA^\intercal$ corresponding to $\sigma_1, \ldots, \sigma_t$. Then, $\|\Pi AQ^{1/2}\|_{S_1} = \sum_{i=1}^t \sigma_i^{1/2}$, and by Lemma 4.9 applied to the matrices $M = \Pi A$ and $W = Q$, there exists a set $S \subseteq U$ of size at most $|S| \leq \text{rank}\,\Pi AQ^{1/2} \leq \varepsilon n$, such that

$$\begin{aligned}\text{specLB}_2(\varepsilon n, A) &\geq \sqrt{\frac{|S|}{|\mathcal{Q}|}}\sigma_{\min}(A_S) \\ &\geq \sqrt{\frac{|S|}{|\mathcal{Q}|}}\sigma_{\min}(\Pi A_S) \geq \frac{c\|\Pi AQ^{1/2}\|_{S_1}}{(\log \varepsilon n)\sqrt{|\mathcal{Q}|}} = \frac{cf_k(E)}{2(\log \varepsilon n)\sqrt{|\mathcal{Q}|}} \qquad (8.16)\end{aligned}$$

for an absolute constant c.

For the second case, assume that $\sqrt{k-t}\left(\sum_{i>t}\sigma_i\right)^{1/2} \geq \frac{1}{2}f_k(E)$. Let Π be an orthogonal projection operator onto the span of the singular vectors of $AQA^\intercal$ corresponding to $\sigma_{t+1}, \ldots, \sigma_m$. By the choice of t, we have

$$\frac{\|\Pi AQ^{1/2}\|_{HS}^2}{\|\Pi AQ^{1/2}\|_2^2} = \frac{\sum_{i>t}\sigma_i}{\sigma_{t+1}} \geq k - t.$$

By the Restricted Invertibility Principle (Theorem 4.5), applied with $M = \Pi A$, $W = Q$, and $\varepsilon = \frac{1}{2}$, there exists a set $S \subseteq U$ of size $\frac{1}{4}(k-t)$ so that

$$\begin{aligned}\text{specLB}_2(\varepsilon n, A) &\geq \sqrt{\frac{|S|}{|\mathcal{Q}|}}\sigma_{\min}(A_S) \\ &\geq \sqrt{\frac{|S|}{|\mathcal{Q}|}}\sigma_{\min}(\Pi A_S) \geq \frac{\sqrt{k-t}\left(\sum_{i>t}\sigma_i\right)^{1/2}}{4\sqrt{|\mathcal{Q}|}} \geq \frac{f_k(E)}{8\sqrt{|\mathcal{Q}|}}. \qquad (8.17)\end{aligned}$$

The theorem follows from (8.15), (8.16), (8.17), and Theorem 8.2. □

Bibliographic Remarks

A variant of Theorem 8.1, with a weaker bound, was proved in [118]. The approach there was somewhat different: the main algorithmic tool was again the Projection Mechanism; however, the same recursively computed Gaussian noise was used as in the large database case. Here we take the alternative approach of optimizing the noise distribution with respect to the specific guarantee achieved by the projection mechanism.

With any set of counting queries and IID Gaussian noise (i.e. spherical noise), the projection mechanism achieves average error $O(\sqrt{n}\log^{1/4}|U|)$ (ignoring the dependence on ε and δ). Moreover, this holds when the average error is computed with respect to any distribution on queries, and via private boosting [58] we can also get a comparable worst-case error guarantee. Using the Frank-Wolfe algorithm, the Projection mechanism can be implemented in time sublinear in the universe size, and polynomial in the number of queries and the database size, if linear programs over the sensitivity polytope can be optimized in time polynomial in the dimension. In fact, it is enough to be able to optimize linear programs over any convex body that contains the sensitivity polytope and is not much wider on average. These observations were used in [56] to give the first algorithm for answering 2-wise marginal queries that has asymptotically optimal error and runs in polynomial time in the number of attributes.

Chapter 9

Reconstruction and Communication Complexity

9.1 Overview

In this chapter we give several further applications of reconstruction attacks in the style of Chapter 7 to topics in computer science. The applications combine the reconstruction algorithms with simple arguments from information theory to prove results on communication complexity. To give some intuition for the applications, assume we have a method to produce some data structure $D(x)$ that allows us to compute an accurate approximation of Ax for any binary vector x. If the additive error of the approximation is less than the appropriate notion of discrepancy, we can nearly reconstruct x, so the mutual information between $D(x)$ and x must be large. We then have a lower bound on the expected size of $D(x)$ in terms of the mutual information. We formalize this general argument, and use it to derive a lower bound in the one-way communication model. As applications, we give space lower bounds for density estimation problems and we strengthen a lower bound of Woodruff [151] for approximating the Hamming distance in the one-way communication model. The latter result implies a new proof of Jayram and Woodruff's [82] space lower bound for computing distinct count in the streaming model.

9.2 The One-way Communication Model

In Yao's communication model [154], we have two parties, Alice and Bob, respectively holding inputs $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, who want to compute the value of a function $f(x, y)$, while minimizing their communication. More generally, they could also compute a relation $F \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$, in which case any z such that $(x, y, z) \in F$ is an admissible

output. Here we consider the more restricted one-way communication model, in which Alice sends a single message to Bob, who must produce the output.

Definition 9.1. *A* deterministic one-way protocol *is a pair of functions* $\Pi = (\Pi_A, \Pi_B)$, $\Pi_A \colon \mathcal{X} \to \{0,1\}^t$, $\Pi_B \colon \{0,1\}^t \times \mathcal{Y} \to \mathcal{Z}$. *The cost of* Π *is equal to* t.

The next definition captures the notion of communication complexity we will be interested in.

Definition 9.2. *Let* $F \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ *be a relation, and let* μ *be a distribution on* $\mathcal{X} \times \mathcal{Y}$. *The* distributional one-way communication complexity $D^{1\text{-}way}_{\mu,\delta}(F)$ *of* F *for error probability* δ *and distribution* μ *is equal to the smallest cost of any deterministic one-way protocol* $\Pi = (\Pi_A, \Pi_B)$ *such that* $\Pr_{(x,y)\sim\mu}[(x,y,z) \in F] \geq 1 - \delta$ *for* $z \triangleq \Pi_B(\Pi_A(x), y)$.

The main application of distributional one-way communication complexity is to prove lower bounds on randomized one-way communication complexity via Yao's minimax principle [153]. While we only work with distributional communication complexity, for completeness we define randomized protocols and randomized communication complexity next.

Definition 9.3. *A* randomized one-way protocol *in the public coin model is a pair of functions* $\Pi = (\Pi_A, \Pi_B)$, $\Pi_A \colon \mathcal{X} \times \{0,1\}^r \to \{0,1\}^t$, $\Pi_B \colon \{0,1\}^t \times \mathcal{Y} \times \{0,1\}^r \to \mathcal{Z}$. *The cost of* Π *is equal to* t, *and the randomness complexity is equal to* r.

Definition 9.4. *Let* $F \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ *be a relation. The* randomized (public coin) one-way communication complexity $R^{1\text{-}way}_{\delta}(F)$ *of* F *for error probability* δ *is equal to the smallest upper bound on the cost of any randomized one-way protocol* $\Pi = (\Pi_A, \Pi_B)$ *such that* $\Pr_b[(x,y,z) \in F] \geq 1 - \delta$ *for* $z \triangleq \Pi_B(\Pi_A(x,b), y, b)$ *and probability taken over* b *sampled uniformly from* $\{0,1\}^r$.

Yao's minimax principle [153] asserts that $R^{1\text{-way}}_{\delta}(F) = \max_\mu D^{1\text{-way}}_{\mu,\delta}(F)$, where the maximum is over all probability distributions μ on $\mathcal{X} \times \mathcal{Y}$. Therefore, any lower bound on $D^{1\text{-way}}_{\mu,\delta}(F)$ for any μ is also a lower bound on $R^{1\text{-way}}_{\delta}(F)$.

9.3 Reconstruction and Fano's Inequality

Starting with the pioneering work of Bar-Yossef, Jayram, Kumar, and Sivakumar [14], information theory has proved to be a powerful tool for proving communication complexity lower bounds. Many applications of information theory can be quite sophisticated; here we only use elementary arguments, and all the necessary background can be found in Chapter 2 of Cover and Joy's text [47]. We make the connection between reconstruction and information lower bounds precise via Fano's inequality. In the following section, we will use this fact to prove communication lower bounds in the one-way model.

All logarithms in this chapter will be base 2. We use $H(X) \triangleq -\mathbb{E}[\ln p(X)]$ for the *entropy* of a random variable X with density function p. The *conditional entropy* of the random variable Y given a random variable X is equal to the expectation of $H(Y|X = x)$ over x sampled according to X. The *binary entropy function* $H_2(p)$ is equal to the entropy of a Bernoulli random variable with success probability p, i.e. $H_2(p) \triangleq -p \log p - (1-p)\log(1-p)$.

Lemma 9.1 (Fano's inequality)**.** *Let X be a random variable taking values in some finite set $\mathcal{X}$. Let Y be another random variable, and let $\hat{X} = g(Y)$ for a (deterministic) function g with range $\mathcal{X}$. Then, for $p_e = \Pr[\hat{X} \neq X]$,*

$$H(X|Y) \leq H_2(p_e) + p_e \log(|X| - 1).$$

Let us use the notation $d_H(x, y) = \|x - y\|_1$ for the *Hamming distance* between two vectors $x \in \{0,1\}^n$. We will need the definition of the *mutual information* $I(X;Y)$ of two random variables:

$$I(X;Y) \triangleq H(X) - H(X|Y) = H(Y) - H(Y|X).$$

Notice that the mutual information is a lower bound on both $H(X)$ and $H(Y)$, because entropy is non-negative. The conditional mutual information $I(X;Y|Z)$ is equal to the mutual information with all entropy functions conditioned on Z, i.e. $H(X|Z) - H(X|Y,Z)$. The chain rule for entropy implies the *chain rule for mutual information*, which is $I((X_1, X_2); Y) = I(X_1;Y) + I(X_2;Y|X_1)$.

The following lemma is a relatively easy consequence of Fano's inequality. It can be seen as a version of the inequality for approximate recovery.

Lemma 9.2. *Let X be a random vector sampled uniformly from $\{0,1\}^n$, and let Y be a random variable. If there exists a deterministic function $\mathcal{R}$ with range $\{0,1\}^n$ such that, with probability at least $2/3$, $d_H(X, \mathcal{R}(Y)) \leq \beta n$, then $I(X;Y) \geq (2/3 - \beta \log(e/\beta))n - H_2(1/3)$.*

Proof. Let $\tilde{X} \triangleq \mathcal{R}(Y)$ and let us define a new random variable $Z \triangleq (Y, S)$, where $S \triangleq \{i : X_i \neq \tilde{X}_i\}$ if $d_H(X, \tilde{X}) \leq \beta n$ and $S \triangleq \emptyset$ otherwise. Whenever $d_H(X, \tilde{X}) \leq \beta n$, we can recover X from Z exactly by flipping all bits of Y indexed by S. So, by Lemma 9.1, $I(X;Z) = H(X) - H(X|Z) \geq 2n/3 - H_2(1/3)$. By the chain rule for mutual information, $I(X;Z) = I(X;Y) + I(X;S|Y)$. We have

$$I(X;S|Y) \leq H(S|Y) \leq H(S) \leq \log \binom{n}{\beta n},$$

where the first inequality is by the non-negativity of entropy, the second holds because conditioning does not increase entropy, and the final inequality holds because entropy is always bounded above by the logarithm of the range. Putting everything together and using the estimate $\binom{n}{\beta n} \leq \left(\frac{e}{\beta}\right)^{\beta n}$, we have

$$I(X;Y) \geq \frac{2}{3}n - \log \binom{n}{\beta n} - H_2(1/3) \geq \left(\frac{2}{3} - \beta \log \frac{e}{\beta}\right) n - H_2(1/3).$$

This completes the proof. □

9.4 Communication Lower Bounds via Robust Discrepancy

Robust discrepancy and the reconstruction algorithm in Lemma 7.8 allow us to prove a lower bound in the one-way communication model for a problem of approximating dot products. Consider a relation $F_t(A) \subseteq \{0,1\}^n \times \mathbb{Z}^n \times \mathbb{Z}$ defined for an integer matrix A to include the tuple (x, a, z) if a is a row of A and $|\langle x, a \rangle - z| \leq t$. The problem of computing $F_t(A)$ in the one-way communication model captures the communication complexity of approximating various problems in which Alice holds a set P and Bob holds a set S, and their goal is approximate $|P \cap S|$. With P a set of points in $\mathbb{R}^d$

and S is a geometric set, we get natural approximate range counting problems, and the communication lower bounds imply lower bounds on the size of data structures. With P and S general, we get a problem equivalent to approximating Hamming distance, for which we prove a lower bound that generalizes a result of Woodruff, arguably with a simpler proof.

Our general result on the one-way communication complexity of $F_t(A)$ follows.

Theorem 9.1. *For μ the uniform distribution on $\{0,1\}^n \times \{a^1, \ldots, a^m\}$, where a^i is the i-th row of the matrix A, and any $t \leq \frac{1}{2}\operatorname{rdisc}_{6\delta,1/8}(A)$, we have $D^{1\text{-}way}_{\mu,\delta}(F_t) = \Omega(n)$.*

Proof. Let $\Pi = (\Pi_A, \Pi_B)$ be a one-way deterministic protocol such that, when (x, a) is sampled from μ, and $z \triangleq \Pi_B(\Pi_A(x), a)$, $\Pr[|\langle x, a\rangle - z| > t] \leq \delta]$. Let $w \in \mathbb{Z}^m$ be the vector defined by $w_i \triangleq \Pi_B(\Pi_A(x), a^i)$. Then, by Markov's inequality applied to the random variable $|\{i : |\langle x, a^i\rangle - w_i| > t\}|$, we have that with probability at least $2/3$ over the choice of x, $\|Ax - w\|_{3\delta,\infty} \leq t$. By Lemma 7.8, there is a deterministic procedure $\mathcal{R}$ so that $d_H(x, \mathcal{R}(w)) \leq n/8$. Since w is itself a deterministic function of $\Pi_A(x)$, we have a deterministic procedure $\mathcal{R}'$ that, with probability at least $2/3$ over the random choice of x, satisfies $d_H(x, \mathcal{R}'(\Pi_A(x))) \leq n/8$. By Lemma 9.2,

$$I(x; \Pi_A(x)) \geq \left(\frac{2}{3} - \frac{1}{8}\log(8e)\right) n - H_2(1/3) = \Omega(n).$$

Since $I(x; \Pi_A(x)) \leq H(\Pi_A(x))$, and the entropy $H(\Pi_A(x))$ is at most the number of bits needed to write $\Pi_A(x)$, i.e. the cost of Π, the theorem is proved. □

9.5 Density Estimation

Let us define a hereditary version of robust discrepancy as

$$\operatorname{hrdisc}_{\delta,\beta}(s, A) \triangleq \max_{J \subseteq [n] : |J| \leq s} \operatorname{rdisc}_{\delta,\beta}(A_J),$$

where $A \in \mathbb{R}^{m \times n}$. We define $\operatorname{hrdisc}_{\delta,\beta}(s, \mathcal{S})$ for a set system $(\mathcal{S}, U)$ with incidence matrix A as $\operatorname{hrdisc}_{\delta,\beta}(s, A)$. We use hereditary robust discrepancy to give lower bounds on the communication complexity of *density estimation problems*. Let us define a relation $G_\varepsilon(\mathcal{S}) \subseteq 2^U \times \mathcal{S} \times \mathbb{R}$ which includes (P, S, ϕ) if $\left|\frac{|P \cap S|}{|P|} - \phi\right| < \varepsilon$. The classical construction of ε-approximations (see e.g. [105]) shows that for any P we can take a subset

$Q \subseteq P$ of size $O(s)$, where s is the smallest integer such that $\text{hrdisc}_{0,1/8}(s, A) \leq \varepsilon s$; then $\left|\frac{|Q\cap S|}{|Q|} - \frac{|P\cap S|}{|P|}\right| \leq \varepsilon$. This gives a data structure of bit size $O(s \log(|P|/s))$ to approximate densities deterministically. The next corollary of Theorem 9.1 shows that the bit size of this construction cannot be improved by more than a factor of $\log(|P|/s)$ even if we allow an arbitrary data structure rather than a subset of P.

Corollary 9.2. *For any s such that* $\text{hrdisc}_{6\delta,1/8}(s, \mathcal{S}) \geq \varepsilon s$, $D_\delta^{1\text{-way}}(G_\varepsilon(\mathcal{S})) = \Omega(s)$.

In the special case when $(\mathcal{S}, U)$ is the set system $\mathcal{B}_2(U)$ induced by axis-aligned rectangles on $U \subseteq \mathbb{R}^2$, Wei and Yi [147] showed a nearly tight lower bound of $\Omega\left(\frac{1}{\varepsilon}\log\frac{1}{\varepsilon}\log|P|\right)$. They also used discrepancy theory techniques, but exploited the fact that set systems induced by axis-aligned rectangles have many restrictions with high discrepancy. Giving similarly near-tight space lower bounds for any density estimation problem is an intriguing open problem.

Question 6. *Give tight bounds on the smallest bit size of a data structure for any given density estimation problem.*

A recent paper by Huang and Yi [80] shows communication lower bounds for computing ε-approximations in a distributed setting, also using discrepancy. The problem the consider is different from ours, even only for two parties: in their setting, the point set P is split between the parties, while in ours Alice holds the entire pointset, while Bob evaluates density queries.

Note that a multiplicative approximation of $|P \cap S|$ within factor $(1 \pm \varepsilon)$ allows an ε-approximation of the density $\frac{|P\cap S|}{|P|}$, so we can interpret the above corollary as also giving a lower bound for multiplicative approximations.

9.6 Approximating Hamming Distance

Let us define the relation $H_{t,n} \subseteq \{0,1\}^n \times \{0,1\}^n \times \mathbb{N}$ to include (x, y, z) for all x, y, and z such that $|d_H(x, y) - z| \leq t$. In other words, $H_{t,n}$ captures the problem of approximating the Hamming distance up to additive error t on strings of length n. Woodruff [151] proved that $D_{\mu,\delta}^{1\text{-way}}(H_{t,n}) = \Omega(n)$ when $t \leq c\sqrt{n}$, μ is the uniform

distribution, and δ is a small enough constant. In fact he showed this lower bound for the promise problem of distinguishing between $d_H(x,y) > n/2 + c\sqrt{n}$ and $d_H(x,y) < n/2 - c\sqrt{n}$. This is known as the Gap Hamming problem. While our techniques can be adapted to this promise version as well with some additional effort, for simplicity we will focus on the approximation problem directly. The next theorem is our main lower bound for approximating Hamming distance. It extends Woodruff's lower bound to also give a tight dependence on the failure probability δ. The theorem follows easily from Theorem 9.1.

Theorem 9.3. *For μ the uniform distribution on $\{0,1\}^n \times \{0,1\}^n$ and $t \leq c'\sqrt{n\log(1/\delta)}$ for small enough constant c', $D^{1\text{-}way}_{\mu,\delta}(H_{t,n}) = \Omega(n)$.*

Proof. Let A be the matrix whose rows are all elements of the set $\{-1,1\}^n$. By Theorem 9.1 and Lemma 7.7, $D^{1\text{-way}}_{\mu',\delta}(F_t(A)) = \Omega(n)$ for $t \leq \frac{1}{2}\operatorname{rdisc}_{6\delta,1/8}(A) \leq \frac{c}{2}\sqrt{n\log(6/\delta)/8}$ and μ' the uniform distribution on $\{0,1\}^n \times \{-1,1\}^n$. Then the theorem follows from the inequality $D^{1\text{-way}}_{\mu,\delta}(H_{t,n}) \geq D^{1\text{-way}}_{\mu',\delta}(F_t(A))$, which holds for any any δ and t. Indeed, given an input $(x,a) \in \{0,1\}^n, \times\{-1,1\}^n$, let us define y by $y_i = (a_i+1)/2$. This transformation is a bijection, so the uniform distribution on $\{0,1\}^n \times \{-1,1\}^n$ induces the uniform distribution on $\{0,1\}^n \times \{0,1\}^n$. Then, because x and y are binary,

$$\begin{aligned} d_H(x,y) = \|x-y\|_1 = \|x-y\|_2^2 & \\ &= \sum_{i=1}^n x_i^2 + \sum_{i=1}^n y_i^2 - 2\langle x,y\rangle \\ &= \sum_{i=1}^n x_i + \sum_{i=1}^n y_i - 2\langle x,y\rangle, \end{aligned}$$

and $\langle x,a\rangle = 2\langle x,y\rangle - \sum_{i=1}^n x_i$. Therefore, $\langle x,a\rangle = \sum_{i=1}^n y_i - d_H(x,y)$. It follows that Bob can approximate $\langle x,a\rangle$ from an approximation to $d_H(x,y)$ only with access to his own input and without degrading the quality of the approximation. □

Theorem 9.3 has implications for space complexity lower bounds in the streaming model [109]. The model is essentially the same as the pan-privacy model from Chapter 7, ignoring the privacy guarantees; indeed, the pan-privacy model was inspired by data streams theory. A stream processing algorithm $\mathcal{A}$ is given as input a sequence of symbols

$\bar{\sigma} \triangleq (\sigma_1, \ldots, \sigma_m)$, $\sigma_t \in \Sigma$, which arrives online. At each time step a symbol arrives and the algorithm updates its memory state. At any point the algorithm is required to be able to output an approximation to some statistic on $\bar{\sigma}$ with high probability. One of the most basic statistics is the *distinct count* F_0: the number of distinct symbols in $\bar{\sigma}$. Kane, Nelson, and Woodruff [84] give an optimal algorithm with space complexity $O(\varepsilon^{-2} + \log n)$ that outputs an $(1+\varepsilon)$-approximation to F_0 with constant probability. By running $O(\log 1/\delta)$ copies of the algorithm in parallel and taking the median answer from all instances, the error probability can be brought down to δ at the cost of increasing the space complexity by a factor of $O(\log 1/\delta)$. Jayram and Woodruff [82] show a lower bound of $\Omega(\varepsilon^{-2} \log(1/\delta))$, which is optimal when $\varepsilon \leq \sqrt{\frac{\log 1/\delta}{\log n}}$. The same lower bound is implied by Theorem 9.3 via standard reductions [150], which we briefly sketch next. We choose $n = \varepsilon^{-2} \log(1/\delta)$ so that $D^{\text{1-way}}_{\mu,\delta}(H_{\varepsilon n, n}) = \Omega(n)$. Then we argue that we can use a streaming algorithm for the distinct count problem to approximate the Hamming distance between arbitrary $x, y \in \{0,1\}^n$; the argument is similar to the proof of Theorem 7.5. The stream is split into two halves, where the first half contains all j so that $x_j = 1$ and the second half contains all j such that $y_j = 1$; Alice constructs the first half of the stream from x, and after processing the stream sends the memory state of the algorithm to Bob, who finishes the computation with the second half of the stream and computes the output. It is easy to argue that, with a small additional message from Alice, Bob can compute the Hamming distance from the distinct count approximation. See [150, 151] for the detailed argument.

It is an interesting research direction to find other natural problems in the streaming model for which Theorem 9.1 implies tight or near-tight space lower bounds.

Chapter 10

Avenues to Further Applications of Discrepancy

10.1 Overview

In prior chapter we posed research questions directly related to the material in these chapters. In this chapter we outline some directions for research in other areas where we believe discrepancy theory can be useful in making progress on important questions. We first present a discrepancy theory view of expander graphs. Then we describe a discrepancy-based approach to compressed sensing, motivated by the reconstruction algorithms of prior chapters. We finish with applications of discrepancy theory to the design of approximation algorithms.

10.2 Expander Graphs and Sparsification

The impact of expander graphs on mathematics and theoretical computer science can hardly be overstated. Here we we give an interpretation of some basic definitions and facts from the viewpoint of discrepancy theory. Essentially all observations we make are widely known.

10.2.1 Spectral Expansion as Discrepancy

We start by recalling several basic definitions from spectral graph theory. All graphs will be simple, unweighted, and undirected. Recall that the *adjacency matrix* $A \in \{0,1\}^{V \times V}$ of a graph $G = (V, E)$ is defined by $a_{u,v} = 1 \Leftrightarrow (u, v) \in E\}$. The *graph Laplacian* L_G of G is the matrix $L \triangleq D - A$, where $D \in \mathbb{N}^{V \times V}$ is a diagonal matrix with the degree sequence of G on the main diagonal, i.e. $d_{uu} \triangleq \deg_G(u) \triangleq |\{v \in V : (u, v) \in E\}|$. Equivalently, $L = \sum_{(u,v)\in E} (e_u - e_v)(e_u - e_v)^\intercal$, where $e_u \in \mathbb{R}^V$ is the

standard basis vector corresponding to u. Therefore, L is positive semidefinite, and $x^\intercal Lx = \sum_{(u,v)\in E}(x_u - x_v)^2$ for every $x \in \mathbb{R}^V$. Letting $\lambda_1 \leq \ldots \leq \lambda_n$ be the eigenvalues of L (with multiplicities), we see that $\lambda_1 = 0$ with eigenvector the all-ones vector e. The dimension of the nullspace of L is equal to the number of components of G.

Recall the Erdős-Renyi random graph model $G_{n,p}$, in which each edge $(u,v) \in \binom{V}{2}$ is sampled independently with probability p. The *expected Laplacian* of a graph $G \sim G_{n,p}$ is

$$\mathbb{E}_{G\sim G_{n,p}} L_G = \sum_{(u,v)\in E} p(e_u - e_v)(e_u - e_v)^\intercal = p(n-1)I - p(J-I) = p(nI - J),$$

where J is the all-ones matrix. We define a (n,p)-expander as a graph that approximates this expected Laplacian in the spectral norm.

Definition 10.1. *A graph G is an (n,p,α)-expander if $\|L_G - p(nI-J)\|_2 \leq \alpha$, where $\|\cdot\|_2$ is the spectral norm.*

We emphasize that this is not quite the standard definition, but we will soon relate it to more standard ones.

The notion of an expander graphs, as we defined it, is related to linear discrepancy. Let us define a linear operator $T\colon \mathbb{R}^{\binom{V}{2}} \to \mathbb{R}^{V\times V}$ by

$$Tx \triangleq \sum_{(u,v)\in\binom{V}{2}} x_{u,v}(e_u - e_v)(e_u - e_v)^\intercal.$$

Then if $x(G)$ is the indicator vector of the edge set $E(G)$ of a graph G, G is an (n,p,α) expander if and only if $\|Tx - T(px(K_n))\|_2 \leq \alpha$. Since any finite dimensional norm can be embedded linearly into a finite dimensional subspace of ℓ_∞ with arbitrarily small distortion, there exists a linear operator $T'\colon \mathbb{R}^{\binom{V}{2}} \to \mathbb{R}^N$ so that we can characterize (n,p,α)-expanders as those graphs for which $\|T'x(G) - T'(px(K_n))\|_\infty \leq \alpha'$, for some α' arbitrarily close to α. Thus the problem of constructing an (n,p,α)-expander graph is equivalent to minimizing the linear discrepancy of T' with respect to a fixed vector $px(K_n)$, i.e. the all-ones vector scaled by p.

Let us clarify the relationship of (n,p,α)-expanders to more standard definitions. In their survey, Hoory, Linial, and Wigderson [79] define a degree-d regular graph G

on n vertices to be an expander with parameter σ if $\sigma_2(A) \leq \sigma$, where $\sigma_2(A)$ is the second largest singular value of the adjacency matrix A of G. (Note that, since A is a symmetric matrix, its singular values are equal to the absolute values of the eigenvalues.) This is equivalent to G being an (n, p, α)-expander for $p = \frac{d}{n}$. To see this, notice that $L_G - p(nI - J) = pJ - A$. The only non-zero eigenvalue of the matrix pJ is $pn = d$ and corresponds to the eigenvector $\frac{1}{\sqrt{n}}e$, for e being the all-ones vector. This vector is also an eigenvector of A with the same eigenvalue, and, therefore, the singular values of $pJ - A$ are $\sigma_2(A), \ldots, \sigma_n(A), 0$. It follows that $\|L_G - p(nI - J)\|_2 = \|pJ - A\|_2 = \sigma_2(A)$, which is what we wanted to show.

An important property of expander graphs is captured by the *expander mixing lemma*, which states that any (n, p, α)-expander graph $G = (V, E)$ satisfies

$$\left|E(S,T) - p|S||T|\right| \leq \alpha\sqrt{|S||T|}, \tag{10.1}$$

for any two disjoint sets $S, T \subseteq V$. Here we use the notation $E(S,T) \triangleq |\{(u,v) \in E : u \in S, v \in T\}|$ for the size of the cut between S and T. This property is fairly easy to verify from our definition. Let x be the indicator vector of the set S and y the indicator of T. By the definition of the Laplacian,

$$-x^\intercal L_G y = \sum_{(u,v)\in E} (x_v - x_u)(y_u - y_v) = E(S,T).$$

The last equality follows because each term in the sum is nonzero only if both $|\{u,v\} \cap S| = 1$ and $|\{u,v\} \cap T| = 1$, and, because S and T are disjoint, this happens only if $u \in S$ and $v \in T$ or vice versa. On the other hand, by an analogous argument,

$$-x^\intercal p(nI - J)y = \sum_{(u,v)\in\binom{V}{2}} p(x_v - x_u)(y_u - y_v) = p|S||T|.$$

It then follows from the definitions of the operator norm and Cauchy-Schwarz that

$$\left|E(S,T) - p|S||T|\right| = |x^\intercal(L_G - p(nI - J))y| \leq \|x\|_2\|y\|_2\|L_G - p(nI - J)\|_2.$$

(10.1) follows from the inequality above and the definition of an (n, p, α)-expander.

The bound (10.1) is a typical discrepancy property: it says that the number of edges of any cut in an expander graph is not very different from the expected number

of edges in the same cut in the $G_{n,p}$ model. This property is key in many applications of expanders, e.g. in randomness reduction [1, 46, 81] and hardness of approximation [4]. It resembles, but is different from, another combinatorial notion of discrepancy of graphs, introduced in the work of Erdős and Spencer [60] and Erdős, Goldberg, Pach, and Spencer [61], and more closely related to Ramsey's theorem. That notion compares the density of subgraphs to the expected density in $G_{n,p}$. It would be interesting to explore explicit constructions and algorithmic applications of this notion of low-discrepancy graphs as well.

It turns out that the discrepancy property (10.1) nearly characterizes expanders: Bilu and Linial [25] showed that if a d-regular graph G on n vertices satisfies (10.1) for all S, then G is an $(n, \frac{d}{n}, O(\alpha \log(2d/\alpha)))$-expander. They used this fact to construct infinite families of regular expander graphs of any degree d with nearly optimal parameters. Let us clarify what the optimal parameters are. Alon and Boppana (see [119, 64]) showed that any d-regular $(n, \frac{d}{n}, \alpha)$-expander satisfies $\alpha \geq \sqrt{d-1} - o(1)$, where the asymptotic notation assumes d stays fixed and $n \to \infty$. Any graph matching this bound is called a *Ramanujan graph.* Bilu and Linial constructed d-regular $(n, \frac{d}{n}, O(\sqrt{d \log^3 d}))$-expanders for any integer d and infinitely many n. Very recently, Marcus, Spielman and Srivastava [96] showed that there exist infinite families of *bipartite* Ramanujan graphs of any degree. The analogous result for families of non-bipartite graphs remains open. These advances suggest the following question.

Question 7. *Can the definition of (n, p, α)-expander as a low-discrepancy object be used to construct an infinite family (non-bipartite) Ramanujan graphs of any degree via discrepancy theory techniques? Can this view be used to give* deterministic polynomial time *constructions of Ramanujan families?*

The result of Bilu and Linial is efficient: a graph of size n can be constructed in deterministic polynomial in n time. On the other hand, the result of Marcus, Spielman, and Srivastava is only existential.

The connection between the combinatorial discrepancy property (10.1) and expanders proved by Bilu and Linial is tight. Therefore, in order to make progress on

Question 7, we need to work directly with the more linear-algebraic definition.

10.2.2 Sparsification

Marcus, Spielman, and Srivastava's recent resolution of the Kadison-Singer problem [97] makes some progress on Question 7. Their result implies the following discrepancy bound. This observation was made, for example, in the weblog post [139].

Theorem 10.1 ([97])**.** *Let $M = \sum_{i=1}^m v_i v_i^\intercal$, where $v_1, \ldots, v_m \in \mathbb{R}^n$. If $v_i^\intercal M^+ v_i \leq \alpha$ for all i and M^+ denoting the pseudoinverse of M, then there exist signs $\varepsilon_1, \ldots, \varepsilon_m \in \{-1, 1\}$ such that for all $x \in \mathbb{R}^n$,*

$$\left| \sum_{i=1}^m \varepsilon_i \langle v_i, x \rangle^2 \right| \leq 10\sqrt{\alpha} \sum_{i=1}^m \langle v_i, x \rangle^2.$$

In particular, $\| \sum_{i=1}^m \varepsilon_i v_i v_i^\intercal \|_2 \leq 10\sqrt{\alpha} \|M\|_2$, where $\| \cdot \|_2$ is the spectral norm.

Theorem 10.1 is a vector-balancing result for "small" rank-1 matrices with respect to the spectral norm. The values $v_i^\intercal M^+ v_i$ are known as *leverage scores*. If V is the matrix whose columns are $v_1, \ldots, v_m$, and Π is the orthogonal projection matrix onto the row-span of V, then the leverage scores are equal to the diagonal entries of Π. The condition that the leverage scores are bounded by α is related to the notion of *coherence*; when α is small, it implies that no strict subset of $v_1 v_1^\intercal, \ldots, v_m v_m^\intercal$ has too large of a contribution to the total energy $\operatorname{tr}(M)$.

In the context of expander graph constructions, Theorem 10.1 and the classical "halving" construction in combinatorial discrepancy theory can be applied to construct (n, p, α)-expanders. The halving construction itself is outlined in [139] and is very closely related to the proof of Beck's transference lemma (Lemma 1.1). Let us take $M \triangleq L_{K_n} = \sum_{(u,v) \in \binom{V}{2}} (e_u - e_v)(e_u - e_v)^\intercal = nI - J$. All leverage scores are equal to $2/n$, and, by Theorem 10.1, there exist signs $\{\varepsilon_{u,v}\}_{(u,v) \in \binom{V}{2}}$ such that

$$\| \sum_{(u,v) \in \binom{V}{2}} \varepsilon_{u,v} (e_u - e_v)(e_u - e_v)^\intercal \|_2 = O(\sqrt{n}).$$

We can then take the graph $G = (V, E)$ where E is the smaller of the two edge sets $E_+ = \{(u, v) : \varepsilon_{u,v} = +1\}$ and $E_- = \{(u, v) : \varepsilon_{u,v} = -1\}$. We have

$$\|L_G - \frac{1}{2} L_{K_n}\|_2 = \|L_G - \frac{1}{2}(nI - J)\| \leq \frac{1}{2} \cdot O(\sqrt{n}).$$

I.e. G is an $(n, \frac{1}{2}, \frac{1}{2} \cdot O(\sqrt{n}))$-expander. We can then apply the same technique to $M = L_G$ to get a $(n, \frac{1}{4}, \frac{1}{4} \cdot O(\sqrt{n}))$-expander, and so on recursively, until we have a $(n, \frac{d}{n}, O(\sqrt{d}))$-expander. This is close to optimal, but to resolve Question 7, we would need to get tighter constant factors and adapt the construction to produce regular graphs of any degree. It is also an interesting question whether this construction can be done in polynomial time, as the known proof of Theorem 10.1 is existential.

The "sparsification by halving" argument above can be applied to any graph H in order to derive a sparser spectral approximation G. Here, by *spectral approximation*, we mean that, for some $p < 1$, $\|L_G - pL_H\|_2$ is bounded. The quality of the sparsification will depend on the leverage scores, which in the case of graph Laplacians are equal to the *effective resistances* of the graph edges. A similar sparsification result was proved by Batson, Spielman, and Srivastava [17]. In the setting of Theorem 10.1, they proved that there exists a set of scalars $x_1, \ldots, x_m$, at most dn of them nonzero, so that

$$\left(1 - \frac{1}{\sqrt{d}}\right)^2 M \preceq \sum_{i=1}^{m} x_i v_i v_i^\intercal \preceq \left(1 + \frac{1}{\sqrt{d}}\right)^2 M.$$

In fact, this result does not require any condition on the leverage scores. It is proved via a deterministic polynomial-time algorithm, but it requires that the sparsified graph be weighted. As there has been substantial recent progress on constructive methods in discrepancy theory, we are prompted to ask the following question.

Question 8. *Can constructive discrepancy minimization techniques be applied to efficiently produce, given a graph H, an unweighted sparse graph G that is a spectral approximation to G?*

We also note that there are other notions of graph sparsification. For one closely related example, *cut sparsifiers* [22] relax the spectral approximation requirement and require that $x^\intercal(L_G - pL_H)x$ is bounded only for binary vectors x. One can also define sparsifiers with respect to a measure approximation based on subgraph densities: G approximates H scaled down by p if the density of any induced subgraph of G is close to the density of the corresponding subgraph of H scaled down by p. This approximation notion is closely related to the discrepancy quantity for pairs of graphs defined by Bollobás and Scott [29].

10.3 Compressed Sensing

A basic observation in signal processing is that real-life signals are often *sparse* in some basis, or at least well-approximated by a sparse signal. A popular example is digital images, which tend to be sparse in the wavelet basis. This fact is traditionally exploited for *compression*: after an image is acquired, only the largest coefficients are retained, while those that fall below some threshold are dropped; once the remaining coefficients are transformed back into an image, we get an image that visually looks very close to the original, but can be stored in smaller space. *Compressed sensing* is a new framework in which the first two-steps of the traditional approach are combined into one: the measurements are carefully designed so that we directly acquire a compressed image. Moreover, the number of measurements is comparable to the size of the image *after compression*. Compressed sensing has revolutionized signal processing and is now an active field which has also crossed over into computer science and statistics. For a recent survey of results, we recommend the book [59], and in particular the introductory chapter by Davenport, Duarte, Eldar, and Kutyniok.

In this section we offer a more combinatorial perspective on compressed sensing, inspired by the reconstruction algorithms in Chapter 7. These connections are preliminary, and we do not aim to reconstruct the best results in compressed sensing. Our goal is rather to offer a different perspective, which can hopefully lead to further advances.

We represent a *signal* as a vector $x \in \mathbb{R}^n$. We assume that the vector is *k-sparse* in the standard basis, i.e. has at most k non-zero entries. This comes without loss of generality: if the signal is sparse in another basis, we can perform a change of basis in order to make sure the assumption is satisfied. The goal in compressed sensing is to design a *measurement matrix* $A \in \mathbb{R}^{m \times n}$, so that any k-sparse x can be efficiently reconstructed from Ax. Moreover, it is desirable that the reconstruction is robust in a number of ways: we would like a good approximation $\tilde{x}$ of x when we only observe noisy measurements, and when x is not exactly k-sparse but only close to a k-sparse vector. This class of problems are collectively known as *sparse recovery*.

The following proposition shows a connection between sparse recover and the concept of robust discrepancy defined in Chapter 7. We recall that we use d_H as the Hamming distance function.

Proposition 10.1. *There exists an algorithm $\mathcal{R}$ such that for any real matrix $A \in \mathbb{R}^{m \times n}$, any k-sparse $x \in \{0,1\}^n$ and any y such that*

$$\|y - Ax\|_{\alpha,\infty} \leq \frac{1}{2} \min_{J \subset [n] : |J| = 2k} \operatorname{rdisc}_{2\alpha,\beta}(A_J), \tag{10.2}$$

$\tilde{x} \triangleq \mathcal{R}(A, y)$ satisfies $d_H(\tilde{x}, x) \leq \beta k$.

Proof. The proof is very similar to that of Lemma 7.8. We define $\mathcal{R}(A, y)$ as

$$\mathcal{R}(A, y) \triangleq \arg \min_{\tilde{x} \in \{0,1\}^n, k\text{-sparse}} \|A\tilde{x} - y\|_{\alpha,\infty}.$$

Let $\tilde{x} \triangleq \mathcal{R}(A, y)$ and $D \triangleq \min_{J \subseteq [n] : |J| = 2k} \operatorname{rdisc}_{\alpha,\beta}(A_J)$. By assumption, $\|A\tilde{x} - y\|_{\alpha,\infty} \leq \|Ax - y\|_{\alpha,\infty} \leq D/2$. By the approximate triangle inequality (7.3), we have the guarantee

$$\|A\tilde{x} - Ax\|_{2\alpha,\infty} \leq \|A\tilde{x} - y\|_{\alpha,\infty} + \|y - Ax\|_{\alpha,\infty} \leq D.$$

Since $\tilde{x}$ and x are binary, $\tilde{x} - x \in \{-1, 0, 1\}^n$. Moreover, because both vectors are k-sparse, the union of their supports is contained in some set $J \subseteq [n]$ of size $2k$, so $A(\tilde{x} - x) = A_J(\tilde{x} - x)$. Then, by the definition of $\operatorname{rdisc}_{\alpha,\beta}(A)$, we have $d_H(\tilde{x}, x) = \|\tilde{x} - x\|_1 \leq \beta k$. □

The quantity $\min_{J \subseteq [n] : |J| = k} \operatorname{rdisc}_{\alpha,\beta}(A_J)$ can be seen as a combinatorial analogue of the *restricted isometry property (RIP)* of order k, which requires that for any submatrix A_J for $|J| = k$, the ratio between the largest singular value of A_J and the smallest nonzero singular value is bounded by $1 + \epsilon$. The correspondence would be closer if we were to replace the $\|\cdot\|_{\alpha,\infty}$ norm in the definition of robust discrepancy with the ℓ_2^m-norm.

Proposition 10.1 shows that sparse reconstruction is possible in the presence of an α-fraction of *gross (unbounded) errors*, and the other errors bounded as in the right hand side of (10.2). In this sense it gives a *robust* reconstruction guarantee. This mixed error setting is similar to the one in [52, 34]. These papers do not consider the

sparse setting but they do propose *efficient* reconstruction algorithms. We suggest the following question.

Question 9. *Under what conditions can the reconstruction algorithm $\mathcal{R}$ in Proposition 10.1 be made efficient?*

We have not addressed several other issues which are important in compressed sensing. For example, usually the signal x is arbitrary, rather than binary. This issue can be addressed by appropriately strengthening the definition of discrepancy; we will not pursue this further in this section. A very important issue is the number of measurements m. It can be shown that for $m = \Theta(k \log(n/k))$ random linear measurements drawn from the Rademacher distribution, the right hand side of (10.2) is, with overwhelming probability, $\Omega(\sqrt{\beta k})$ for any constant α.

Proposition 10.2. *Let the matrix A be picked uniformly at random from $\{-1,1\}^{m\times n}$. There exists a constant C such that for $m \geq Ck\log(n/k)$, with probability $1 - e^{-\Omega(n)}$ we have that for any set $J \subseteq [n]$ of size $|J| = k$, $\mathrm{rdisc}_{\alpha,beta}(A_J) = \Omega(\sqrt{\beta k \log(1/\alpha)})$.*

Proof. Let P be the matrix whose rows are the set $\{-1,1\}^n$. Let $\beta_0 \triangleq \beta k/n$. For any $J \subseteq [n]$ and any $x \in \{-1,0,1\}^J$, we define its *extension* $x' \in \{-1,0,1\}^n$ to agree with x on J and have entries 0 everywhere else. Then, there exists a constant c such that for any α, any $J \subseteq [n]$ of size k, and any $x \in \{-1,0,1\}^J$ such that $\|x\|_1 \geq \beta k$, by Lemma 7.7,

$$\|P_J x\|_{2\alpha,\infty} = \|Px'\|_{2\alpha,\infty} \geq \mathrm{rdisc}_{2\alpha,\beta_0}(P) \geq c\sqrt{\beta_0 n \log(1/2\alpha_0)} = c\sqrt{\beta k \log(1/2\alpha)}.$$

Let A be the random matrix we get by sampling m rows uniformly and independently from P. For any fixed J and x as above, $\mathbb{E}[|\{i : |(A_J x)| > c\sqrt{\beta k \log(1/2\alpha)}\}|] \geq 2\alpha m$, and, by the Chernoff bound,

$$\Pr[\|A_J x\|_{\alpha,\infty} < c\sqrt{\beta k \log(1/\alpha_0)}] \leq \exp(-c'm),$$

for a constant c'. Setting $m > n + \frac{1}{c'} \ln\left(3^k \binom{n}{k}\right)$ and taking a union bound over all choices of J and x completes the proof. □

The bound in Proposition 10.2 is of the same order of magnitude as the size of random matrices with the restricted isometry property.

An interesting question is whether sparse reconstruction is possible with more restricted measurements. If the measurements have some nice geometric structure, it is possible that designing the sensing hardware would be less costly. Discrepancy theory seems like a well-suited tool to address this problem, since it provides discrepancy estimates for many classes of structured matrices A. However, while Proposition 10.2 shows that the quantity on the right hand side of (10.2) can be nicely bounded from below for random matrices, this is in general a very strong property, and it is not clear if it holds for any family of structured matrices. On the other hand, it is natural to also assume that the signal x has some nice structure, and it seems plausible that under such an assumption reconstruction is possible even with restricted measurements. As a motivating example, we have the following proposition.

Proposition 10.3. *Let $P \subseteq [n]^2$ be a $O(1)$-spread set (see Definition 2.3) of k points in the plane. Let $\mathcal{H}$ be the set of halfplanes that have non-zero intersection with $[n]^2$, and let $y \in \mathbb{R}^{\mathcal{H}}$ be such that for any $H \in \mathcal{H}$, $|y_H - |H \cap P|| = o(k^{1/4})$. There exists an algorithm $\mathcal{R}$ such that $|\mathcal{R}(y) \triangle P| = o(k)$.*

Proof Sketch. The reconstruction algorithm $\mathcal{R}$ outputs a c-spread point set $\tilde{P}$ that minimizes $\max_H |y_H - |H \cap \tilde{P}||$. Let A be the incidence matrix of the set system induced by $\mathcal{H}$ on $[n]^2$, and let x be the indicator vector of P. By an argument analogous to the one in Proposition 10.1, it is enough to show that $\|A(x - \tilde{x})\|_\infty = \Omega(k^{1/4})$ for any indicator vector $\tilde{x}$ of a c-spread set such that $\|x - \tilde{x}\|_1 = \Omega(k)$. Notice that a c-spread set is contained in a disc of radius $c\sqrt{n}$. Let $\tilde{P}$ be the set of points for which $\tilde{x}$ is an indicator vector. If we can draw two discs of radius $c\sqrt{n}$, one containing P and one containing $\tilde{P}$, such that the discs intersect, then $P \cup \tilde{P}$ is $2c$-spread and the claim follows from Lemma 2.7. Otherwise, there is a line separating P and $\tilde{P}$, and for any halfplane H bounded by this line, $|(A(x - \tilde{x}))_H| = ||P \cap H| - |\tilde{P} \cap H|| = \Omega(k)$. This completes the proof sketch. □

Proposition 10.3 bounds the amount of information needed for reconstruction in a

different way from the usual reconstruction results: by putting a restriction on the expressiveness of measurements rather than on their number. Also, the restriction on the signal combines a geometric assumption (well-spreadedness) and a sparsity assumption. This is similar to *model-based compressed sensing*, see e.g. [15].

Nevertheless, it is interesting to explore whether the number of measurements in Proposition 10.3 (which *apriori* is $O(n^4)$ since this is the number of distinct sets induced by halfplanes on $[n]^2$) can be reduced. A possible direction is to consider a limited number of adaptive measurements to "weed out" most of the grid $[n]^2$, followed by $O(k^2)$ non-adaptive halfplane measurements. Another important question is whether the reconstruction algorithm can be made to run in polynomial time.

We finish the section with the following general question.

Question 10. *Under what natural assumptions on the signal x is reconstruction from a restricted class of structured measurements possible? What structured measurements are important in practice, e.g. for reducing the cost of compressed sensing hardware?*

10.4 Approximation Algorithms

Many combinatorial optimization problems can be posed as an *integer program* (IP) $\min\{c^\intercal x : Ax \geq b, x \in \mathbb{Z}^n\}$. In general, such formulations are NP-hard, as are many interesting examples. As a basic example, consider the NP-hard SetCover problem, in which we are given m subsets $S_1, \ldots, S_m \subseteq [n]$, and our goal is to find a set $I \subseteq [m]$ of the smallest size such that $\bigcup_{i \in I} S_i = [n]$. As an integer program, SetCover can be formulated as $\min\{e^\intercal x : A^T x \geq e, x \in \{0,1\}^m\}$, where $e = (1, \ldots, 1) \in \mathbb{R}^m$ and A is the incidence matrix of the input set system $\{S_1, \ldots, S_m\}$.

While exactly solving an NP-hard problem in polynomial time is implausible, it is often possible to design an efficient approximation algorithm. One of the most powerful strategies for doing this is to *relax* an integer programming formulation of an optimization problem to a linear program (LP) by simply dropping the integrality constraints. I.e., in our general formulation above, the LP relaxation would be $\min\{c^\intercal x : Ax \geq b\}$, and for the SetCover problem the relaxation would be $\min\{e^\intercal x : A^T x \geq e, x \in$

$[0,1]^m\}$. Clearly, for a minimization problem, the value of the LP relaxation is no larger than the value of the IP. The challenge then is to use the LP to compute a feasible IP solution whose value is not much larger than the optimal value of the LP (and therefore not much larger than the optimal value of the IP as well). One common way to do this is to design a *rounding algorithm* which takes a feasible LP solution x as input and outputs a feasible IP solution $\bar{x}$ so that $c^\intercal x \geq \alpha c^\intercal \bar{x}$. This guarantee then implies an approximation factor of α^{-1}. For general background and more information on the design of approximation algorithms we refer the reader to the books by Williamson and Shmoys [149] and by Vazirani [145].

The connection between rounding algorithms and discrepancy theory is via *linear discrepancy*. Recall that we define the linear discrepancy $\operatorname{lindisc}(A)$ of a matrix A as

$$\operatorname{lindisc}(A) \triangleq \max_{c\in[-1,1]^n} \min_{x\in\{-1,1\}^n} \|Ax - Ac\|_\infty.$$

Recall also that, by Theorem 1.1, $\operatorname{lindisc}(A) \leq 2\operatorname{herdisc}(A)$.

Proposition 10.4. *Let* $v_{IP} \triangleq \min\{c^\intercal x : Ax \geq b, x \in \mathbb{Z}^n\}$ *and* $v_{LP} \triangleq \min\{c^\intercal x : Ax \geq b\}$. *Define the matrix* $D \triangleq \begin{pmatrix} c^\intercal \\ A \end{pmatrix}$. *There exists a solution* $\bar{x} \in \mathbb{Z}^n$ *such that*

$$c^\intercal x - v_{LP} \leq \frac{1}{2}\operatorname{lindisc}(D)$$
$$\|Ax - b\|_\infty \leq \frac{1}{2}\operatorname{lindisc}(D).$$

Proof. Let x^* be the optimal solution of the LP $\min\{c^\intercal x : Ax \geq b\}$, and let x_0 be the vector consisting of the integer parts of each coordinate of x^*. Let $x_1 \triangleq \arg\min_{x\in\{-1,1\}^n} \|Dx - Df\|_\infty$ for $f \triangleq e - 2(x^* - x_0) \in [-1,1]^n$ and e the all-ones vector. By the definition of linear discrepancy, $\|Dx_1 - Df\|_\infty \leq \operatorname{lindisc}(D)$. Let $\bar{x} \triangleq x_0 + \frac{1}{2}(e - x_1)$, and observe that

$$x^* - \bar{x} = x^* - x_0 - \frac{1}{2}(e - x_1) = \frac{1}{2}(x_1 - f),$$

and, therefore, $\|Dx^* - D\bar{x}\|_\infty = \|Dx_1 - Df\|_\infty$, and the proposition follows. □

An important note to make here is that if we can minimize linear discrepancy in polynomial time (for the given matrix), then the integer solution $\bar{x}$ can also be found in

polynomial time. Moreover, the proof of Theorem 1.1 is constructive, in the sense that if we can find a coloring in polynomial time that achieves discrepancy bounded by the hereditary discrepancy, then we can compute a coloring that achieves linear discrepancy bounded by at most twice the hereditary discrepancy. It is also not necessary to exactly minimize discrepancy and linear discrepancy: whatever value we can achieve efficiently will give a corresponding bound in Proposition 10.4.

Proposition 10.4 does not immediately imply an approximation guarantee, because the integer solution $\bar{x}$ is not necessarily feasible. However, in special cases, it may be possible to "fix" $\bar{x}$ to make it feasible, while incurring only a small cost in terms of the objective value $c^\intercal x$. One simple strategy, which works when A, b, c are non-negative, is to *scale* the vector b in the linear program by a large enough number K so that $\|Kb - b\|_\infty \geq \frac{1}{2}\operatorname{lindisc}(D)$. The new linear program $\min\{c^\intercal x : Ax \geq Kb\}$ has value at most Kv_{LP}, and if we apply Proposition 10.4 to it, we get an integral $\bar{x}$ which is feasible for the original IP and has objective function value at most $Kv_{LP} + \frac{1}{2v_{LP}}\operatorname{lindisc}(D)$.

As an example, let us apply the above observation to the SetCover problem. It is easy to see that the linear discrepancy of the matrix $\begin{pmatrix} e^\intercal \\ A^\intercal \end{pmatrix}$ is at most the degree $\Delta_{\mathcal{S}}$ of the input set system $\mathcal{S} = \{S_1, \ldots, S_m\}$: any coloring $x \in \{-1, 1\}^m$ that satisfies $|e^\intercal x| \leq 1$, for example, achieves this bound. Therefore, we can approximate SetCover up to a factor of $(1 + \frac{1}{v_{LP}})\frac{1}{2}\Delta_{\mathcal{S}} + 1$. For example, when $\Delta_{\mathcal{S}} = 2$, we have the VertexCover problem, and for large enough v_{LP} we nearly recover the best known approximation ratio of 2. (When the optimal vertex cover is of constant size, it can be found in polynomial time.) This approach is similar to the scaling strategy proposed by Raghavan and Thompson [122] for randomized rounding.

For any particular problem there may be a more efficient way to make the integer solution $\bar{x}$ feasible. Eisenbrand, Pálvölgyi, and Rothvoß showed how to do this for the BinPacking problem. In BinPacking we are given a set of n items with sizes $s_1, \ldots, s_n \in [0, 1]$. The goal is to pack the items into the smallest number of bins, each of size at most 1. BinPacking can be relaxed to the Gilmore-Gomory linear program [68] $\min\{e^\intercal x : A^\intercal x \geq e\}$, where the rows of the matrix A are the indicator vectors of all ways

to pack the items into a bin of size 1. In fact, this is a special case of the SetCover problem, but the sets are exponentially many, and are given implicitly. Karmarkar and Karp [85] showed that this linear program can be efficiently approximated to any given degree, and can then be rounded to get a packing that uses at most $O(\log^2 n)$ more bins than the optimal solution. In the interesting special case where all item sizes are bounded from below by a constant, Karmarkar and Karp's algorithm gives additive approximation $O(\log n)$. Eisenbrand et al. presented a discrepancy-based approach to improve on Karmarkar and Karp's algorithm for this special case. Assuming that $s_1 \geq \ldots \geq s_n$, they substitute the constraint $A^\intercal x \geq b$ with $LA^\intercal x \geq Lb$, where L is the $n \times n$ lower triangular matrix with 1s on the main diagonal and below it. Hall's marriage theorem can be used to show that this new constraint is equivalent to the original one. However, the new constraint has the benefit that it allows for an easy method of fixing "slightly infeasible" solutions $\bar{x}$: if $Lb - LA\bar{x} \leq de$ for some value d, then we can make $\bar{x}$ feasible by only opening d new bins. Eisenbrand et al. showed that when the item sizes in the BinPacking instance are bounded below by a constant, the discrepancy of LA is equal, up to constants, to the discrepancy of a set system of initial intervals of $O(1)$ permutations on $[n]$.

Unfortunately [113] proved the existence of 3 permutations on $[n]$ so that the set system of their initial intervals is $\Omega(\log n)$, showing that the original approach of Eisenbrand at al. could not improve on the Karmakar and Karp algorithm ([113] also showed that the same holds for a larger natural class of rounding algorithms). Nevertheless, this does not mean that discrepancy-based rounding, together with other methods, could not lead to an improved approximation gaurantee for the BinPacking problem. A powerful illustration of this argument is the recent work by Rothvoß [128], who improved on Karmakar and Karp's algorithm and showed that for general BinPacking instances, the optimal solution can be approximated to within $O(\log n \log\log n)$ bins. His algorithm, on a very high level, transforms the constraint matrix via gluing and grouping operations (without changing the optimal value of the LP relaxation much) so that the discrepancy becomes very low.

In the reverse direction, assume we have an integer program $\min\{c^\intercal x : Ax \geq b, x \in$

$\{0,1\}^n\}$. We have that there exists some vector x so that any integer vector $\bar{x}$ satisfies $\|Ax - A\bar{x}\|_\infty \geq \text{lindisc}(A)$. While this does not imply a gap between the integer program and its linear relaxation, it is plausible that, for specific problems, such a connection can be made. This is especially interesting for BinPacking, where the largest known additive gap between the Gilmore-Gomory linear program and the smallest achievable number of bins is 1.

Question 11. *Can linear discrepancy be used to prove a super-constant additive integrality gap for the Gilmore-Gomory relaxation of bin packing? For other interesting problems? Can discrepancy-based rounding be used to give improved approximation algorithms for interesting problems.*

We note that discrepancy techniques were successfully used to give approximation algorithms and integrality gaps for the broadcast scheduling problem [12].

10.5 Conclusion

Many questions in computer science can be phrased as questions about how well a "simple" (discrete) structure can mimic a "complex" (continuous) structure. Techniques to address such problems have been developed in parallel in discrepancy theory and computer science. There have been many interesting examples of interaction between the two fields, some presented in this thesis, and we can expect more such examples in the future. Moreover, while discrepancy theory is already a mature field, we only recently began to understand the computational challenges associated with it. Until a few years ago, many positive results in discrepancy were not constructive, and thus not available for the design of efficient algorithms. Furthermore, prior to the results of this thesis, no efficient non-trivial algorithms were known to accurately estimate the fundamental measures of combinatorial discrepancy. As we understand these computational discrepancy theory questions better, we can expect that the relevance of discrepancy to computer science and related fields will only grow.

Vita

Aleksandar Nikolov

2014 Ph. D. in Computer Science, Rutgers University

2004-08 B. Sc. in Computer Science from Saint Peter's University

2012-2014 Simons Graduate Fellow, Dept. of Computer Science, Rutgers University

2008-2013 Graduate Assistant, Dept. of Computer Science, Rutgers University

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. Deterministic simulation in LOGSPACE. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 132–140, 1987.

[2] R. Alexander. Geometric methods in the study of irregularities of distribution. *Combinatorica*, 10(2):115–136, 1990.

[3] Noga Alon, László Babai, and Alon Itai. A fast and simple randomized parallel algorithm for the maximal independent set problem. *J. Algorithms*, 7(4):567–583, 1986.

[4] Noga Alon, Uriel Feige, Avi Wigderson, and David Zuckerman. Derandomized graph products. *Computational Complexity*, 5(1):60–75, 1995.

[5] Noga Alon and Yishay Mansour. ϵ-discrepancy sets and their application for interpolation of sparse polynomials. *Inform. Process. Lett.*, 54(6):337–342, 1995.

[6] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2008.

[7] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *J. ACM*, 42(4):844–856, July 1995.

[8] Per Austrin, Venkatesan Guruswami, and Johan Håstad. $(2+\epsilon)$-SAT is NP-hard. In *ECCC*, 2013.

[9] W. Banaszczyk. Balancing vectors and gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

[10] Wojciech Banaszczyk. Balancing vectors and convex bodies. *Studia Math.*, 106(1):93–100, 1993.

[11] Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on*, pages 3–10. IEEE, 2010.

[12] Nikhil Bansal, Moses Charikar, Ravishankar Krishnaswamy, and Shi Li. Better algorithms and hardness for broadcast scheduling via a discrepancy approach. In *SODA*, pages 55–71, 2014.

[13] Nikhil Bansal and Joel Spencer. Deterministic discrepancy minimization. *Algorithmica*, 67(4):451–471, 2013.

[14] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. Information theory methods in communication complexity. In *Proceedings of the 17th Annual IEEE Conference on Computational Complexity, Montréal, Québec, Canada, May 21-24, 2002*, pages 93–102, 2002.

[15] Richard G. Baraniuk, Volkan Cevher, Marco F. Duarte, and Chinmay Hegde. Model-based compressive sensing. *IEEE Trans. Inform. Theory*, 56(4):1982–2001, 2010.

[16] I. Bárány and VS Grinberg. On some combinatorial questions in finite-dimensional spaces. *Linear Algebra and its Applications*, 41:1–9, 1981.

[17] Joshua D. Batson, Daniel A. Spielman, and Nikhil Srivastava. Twice-ramanujan sparsifiers. *SIAM Review*, 56(2):315–334, 2014.

[18] J. Beck and T. Fiala. integer-making theorems. *Discrete Applied Mathematics*, 3(1):1–8, 1981.

[19] József Beck. Balanced two-colorings of finite sets in the square i. *Combinatorica*, 1(4):327–335, 1981.

[20] József Beck. Roth's estimate of the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1(4):319–325, 1981.

[21] József Beck and Vera T. Sós. Discrepancy theory. In *Handbook of combinatorics, Vol. 1, 2*, pages 1405–1446. Elsevier, Amsterdam, 1995.

[22] András A. Benczúr and David R. Karger. Approximating s-t minimum cuts in $\tilde{O}(n^2)$ time. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on the Theory of Computing, Philadelphia, Pennsylvania, USA, May 22-24, 1996*, pages 47–55, 1996.

[23] Aditya Bhaskara, Daniel Dadush, Ravishankar Krishnaswamy, and Kunal Talwar. Unconditional differentially private mechanisms for linear queries. In *Proceedings of the 44th symposium on Theory of Computing*, STOC '12, pages 1269–1284, New York, NY, USA, 2012. ACM.

[24] Rajendra Bhatia. *Matrix analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1997.

[25] Yonatan Bilu and Nathan Linial. Lifts, discrepancy and nearly optimal spectral gap. *Combinatorica*, 26(5):495–519, 2006.

[26] Avrim Blum, Katrina Ligett, and Aaron Roth. A learning theory approach to non-interactive database privacy. In *STOC '08: Proceedings of the 40th annual ACM symposium on Theory of computing*, pages 609–618, New York, NY, USA, 2008. ACM.

[27] Manuel Blum, Vaughan Pratt, Robert E. Tarjan, Robert W. Floyd, and Ronald L. Rivest. Time bounds for selection. *J. Comput. System Sci.*, 7:448–461, 1973. Fourth Annual ACM Symposium on the Theory of Computing (Denver, Colo., 1972).

[28] Géza Bohus. On the discrepancy of 3 permutations. *Random Structures Algorithms*, 1(2):215–220, 1990.

[29] Béla Bollobás and Alex Scott. Intersections of graphs. *J. Graph Theory*, 66(4):261–282, 2011.

[30] J. Bourgain and L. Tzafriri. Invertibility of large submatrices with applications to the geometry of banach spaces and harmonic analysis. *Israel journal of mathematics*, 57(2):137–224, 1987.

[31] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, 2004.

[32] Andrei Z. Broder, Moses Charikar, Alan M. Frieze, and Michael Mitzenmacher. Min-wise independent permutations. *J. Comput. Syst. Sci.*, 60(3):630–659, 2000.

[33] Mark Bun, Jonathan Ullman, and Salil Vadhan. Fingerprinting codes and the price of approximate differential privacy. *arXiv preprint arXiv:1311.3158*, 2013.

[34] Emmanuel J. Candès and Paige A. Randall. Highly robust error correction by convex programming. *IEEE Trans. Inform. Theory*, 54(7):2829–2840, 2008.

[35] T-H. Hubert Chan, Elaine Shi, and Dawn Song. Private and continual release of statistics. In *ICALP*, 2010.

[36] Moses Charikar, Venkatesan Guruswami, and Anthony Wirth. Clustering with qualitative information. *J. Comput. Syst. Sci.*, 71(3):360–383, 2005.

[37] Moses Charikar, Alantha Newman, and Aleksandar Nikolov. Tight hardness results for minimizing discrepancy. In *SODA '11: Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1607–1614. SIAM, 2011.

[38] B. Chazelle and A. Lvov. The discrepancy of boxes in higher dimension. *Discrete Comput. Geom.*, 25(4):519–524, 2001. The Micha Sharir birthday issue.

[39] B. Chazelle and A. Lvov. A trace bound for the hereditary discrepancy. *Discrete Comput. Geom.*, 26(2):221–231, 2001. ACM Symposium on Computational Geometry (Hong Kong, 2000).

[40] B. Chazelle, J. Matoušek, and M. Sharir. An elementary approach to lower bounds in geometric discrepancy. *Discrete and Computational Geometry*, 13(1):363–381, 1995.

[41] Bernard Chazelle. *The Discrepancy Method*. Cambridge University Press, 1991.

[42] Bernard Chazelle. A spectral approach to lower bounds with applications to geometric searching. *SIAM J. Comput.*, 27(2):545–556, 1998.

[43] Bernard Chazelle. A minimum spanning tree algorithm with inverse-ackermann type complexity. *Journal of the ACM (JACM)*, 47(6):1028–1047, 2000.

[44] Fan R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997.

[45] K.L. Clarkson. Coresets, sparse greedy approximation, and the frank-wolfe algorithm. *ACM Transactions on Algorithms (TALG)*, 6(4):63, 2010.

[46] Aviad Cohen and Avi Wigderson. Dispersers, deterministic amplification, and weak random sources (extended abstract). In *30th Annual Symposium on Foundations of Computer Science, Research Triangle Park, North Carolina, USA, 30 October - 1 November 1989*, pages 14–19, 1989.

[47] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second edition, 2006.

[48] Irit Dinur and Kobbi Nissim. Revealing information while preserving privacy. pages 202–210, 2003.

[49] B. Doerr, A. Srivastav, and P. Wehr. Discrepancy of Cartesian products of arithmetic progressions. *Electron. J. Combin.*, 11:Research Paper 5, 16 pp. (electronic), 2004.

[50] C. Dwork, F. Mcsherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *TCC*, 2006.

[51] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. 4004:486–503, 2006.

[52] Cynthia Dwork, Frank McSherry, and Kunal Talwar. The price of privacy and the limits of lp decoding. In *STOC*, pages 85–94, 2007.

[53] Cynthia Dwork, Moni Naor, Toniann Pitassi, and Guy N. Rothblum. Differential privacy under continual observation. In Leonard J. Schulman, editor, *STOC*, pages 715–724. ACM, 2010.

[54] Cynthia Dwork, Moni Naor, Toniann Pitassi, Guy N. Rothblum, and Sergey Yekhanin. Pan-private streaming algorithms. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 5-7, 2010. Proceedings*, pages 66–80, 2010.

[55] Cynthia Dwork, Moni Naor, Omer Reingold, Guy N Rothblum, and Salil Vadhan. On the complexity of differentially private data release: efficient algorithms and hardness results. In *Proceedings of the 41st annual ACM symposium on Theory of computing*, pages 381–390. ACM, 2009.

[56] Cynthia Dwork, Aleksandar Nikolov, and Kunal Talwar. Efficient algorithms for privately releasing marginals via convex relaxations. In *Proceedings of the 30th Annual Symposium on Computational Geometry*, Kyoto, Japan, 2014.

[57] Cynthia Dwork and Kobbi Nissim. Privacy-preserving datamining on vertically partitioned databases. In *CRYPTO*, pages 528–544, 2004.

[58] Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan. Boosting and differential privacy. In *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, FOCS '10, pages 51–60, Washington, DC, USA, 2010. IEEE Computer Society.

[59] Yonina C. Eldar and Gitta Kutyniok, editors. *Compressed sensing*. Cambridge University Press, Cambridge, 2012. Theory and applications.

[60] P. Erdős and J. Spencer. Imbalances in k-colorations. *Networks*, 1:379–385, 1971/72.

[61] Paul Erdős, Mark Goldberg, János Pach, and Joel Spencer. Cutting a graph into two dissimilar halves. *J. Graph Theory*, 12(1):121–131, 1988.

[62] Ky Fan. On a theorem of Weyl concerning eigenvalues of linear transformations. I. *Proc. Nat. Acad. Sci. U. S. A.*, 35:652–655, 1949.

[63] M. Frank and P. Wolfe. An algorithm for quadratic programming. *Naval research logistics quarterly*, 3(1-2):95–110, 1956.

[64] Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Mem. Amer. Math. Soc.*, 195(910):viii+100, 2008.

[65] Bernd Gärtner and Jiří Matoušek. *Approximation algorithms and semidefinite programming*. Springer, Heidelberg, 2012.

[66] Alain Ghouila-Houri. Caractérisation des matrices totalement unimodulaires. *C. R. Acad. Sci. Paris*, 254:1192–1194, 1962.

[67] Apostolos A Giannopoulos. On some vector balancing problems. *Studia Mathematica*, 122(3):225–234, 1997.

[68] P.C. Gilmore and R.E. Gomory. A linear programming approach to the cutting-stock problem. *Oper. Res.*, 9:849–859, 1961.

[69] Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 2004. Stochastic Modelling and Applied Probability.

[70] Efim Davydovich Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of banach spaces. *Mathematics of the USSR-Sbornik*, 64(1):85, 1989.

[71] Oded Goldreich, Shari Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *J. ACM*, 45(4):653–750, July 1998.

[72] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[73] Anupam Gupta, Moritz Hardt, Aaron Roth, and Jonathan Ullman. Privately releasing conjunctions and the statistical query barrier. In *STOC*, pages 803–812, 2011.

[74] Anupam Gupta, Aaron Roth, and Jonathan Ullman. Iterative constructions and private data release. In *TCC*, pages 339–356, 2012.

[75] V. Guruswami. Inapproximability results for set splitting and satisfiability problems with no mixed clauses. *Approximation Algorithms for Combinatorial Optimization*, pages 155–166, 2000.

[76] M. Hardt and G. Rothblum. A multiplicative weights mechanism for privacy-preserving data analysis. *Proc. 51st Foundations of Computer Science (FOCS). IEEE*, 2010.

[77] Moritz Hardt, Katrina Ligett, and Frank McSherry. A simple and practical algorithm for differentially private data release. In *NIPS*, 2012. To appear.

[78] Moritz Hardt and Kunal Talwar. On the geometry of differential privacy. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 705–714, New York, NY, USA, 2010. ACM.

[79] Shlomo Hoory, Nathan Linial, and Avi Widgerson. Expander graphs and their applications. *Bull. Am. Math. Soc., New Ser.*, 43(4):439–561, 2006.

[80] Zengfeng Huang and Ke Yi. The communication complexity of distributed ε-approximations. 2014. To appear in FOCS 2014.

[81] Russell Impagliazzo and David Zuckerman. How to recycle random bits. In *30th Annual Symposium on Foundations of Computer Science, Research Triangle Park, North Carolina, USA, 30 October - 1 November 1989*, pages 248–253, 1989.

[82] T. S. Jayram and David P. Woodruff. Optimal bounds for johnson-lindenstrauss transforms and streaming problems with subconstant error. *ACM Transactions on Algorithms*, 9(3):26, 2013.

[83] Gil Kalai. Erdős discrepancy problem 22. `http://gowers.wordpress.com/2012/08/22/edp22-first-guest-post-from-gil-kalai/`, 09 2012.

[84] Daniel M. Kane, Jelani Nelson, and David P. Woodruff. An optimal algorithm for the distinct elements problem. In *Proceedings of the Twenty-Ninth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, PODS 2010, June 6-11, 2010, Indianapolis, Indiana, USA*, pages 41–52, 2010.

[85] Narendra Karmarkar and Richard M. Karp. An efficient approximation scheme for the one-dimensional bin-packing problem. In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 312–320, 1982.

[86] B. Klartag. An isomorphic version of the slicing problem. *J. Funct. Anal.*, 218(2):372–394, 2005.

[87] Boris Konev and Alexei Lisitsa. A sat attack on the erdős discrepancy conjecture. *CoRR*, abs/1402.2184, 2014.

[88] Eyal Kushilevitz and Noam Nisan. *Communication complexity*. Cambridge University Press, Cambridge, 1997.

[89] Kasper Green Larsen. On range searching in the group model and combinatorial discrepancy. *SIAM J. Comput.*, 43(2):673–686, 2014.

[90] A. S. Lewis. The convex analysis of unitarily invariant matrix functions. *J. Convex Anal.*, 2(1-2):173–183, 1995.

[91] Chao Li, Michael Hay, Vibhor Rastogi, Gerome Miklau, and Andrew McGregor. Optimizing linear counting queries under differential privacy. 2010.

[92] L. Lovász. Coverings and coloring of hypergraphs. In *Proceedings of the Fourth Southeastern Conference on Combinatorics, Graph Theory, and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1973)*, pages 3–12. Utilitas Math., Winnipeg, Man., 1973.

[93] L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *European Journal of Combinatorics*, 7(2):151–160, 1986.

[94] László Lovász. Integer sequences and semidefinite programming. *Publ. Math. Debrecen*, 56(3-4):475–479, 2000. Dedicated to Professor Kálmán Győry on the occasion of his 60th birthday.

[95] S. Lovett and R. Meka. Constructive discrepancy minimization by walking on the edges. *Arxiv preprint arXiv:1203.5747*, 2012.

[96] Adam Marcus, Daniel A. Spielman, and Nikhil Srivastava. Interlacing families I: bipartite ramanujan graphs of all degrees. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013, 26-29 October, 2013, Berkeley, CA, USA*, pages 529–537, 2013.

[97] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families ii: Mixed characteristic polynomials and the kadison-singer problem. *arXiv preprint arXiv:1306.3969*, 2013.

[98] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: theory of majorization and its applications*. Springer Series in Statistics. Springer, New York, second edition, 2011.

[99] J. Matoušek. Tight Upper Bounds for the Discrepancy of Halfspaces. *Discrete and Computational Geometry*, 13(1):593–601, 1995.

[100] Jiří Matoušek. Derandomization in computational geometry. *J. Algorithms*, 20(3):545–580, 1996.

[101] Jiří Matoušek. An L_p version of the Beck-Fiala conjecture. *European J. Combin.*, 19(2):175–182, 1998.

[102] Jiří Matoušek. On the discrepancy for boxes and polytopes. *Monatsh. Math.*, 127(4):325–336, 1999.

[103] Jiri Matousek and Aleksandar Nikolov. Combinatorial discrepancy for boxes via the ellipsoid-infinity norm, 2014.

[104] Jiří Matoušek and Joel Spencer. Discrepancy in arithmetic progressions. *J. Amer. Math. Soc.*, 9(1):195–204, 1996.

[105] Jiří Matoušek. *Geometric Discrepancy (An Illustrated Guide)*. Springer, 1999.

[106] Jiří Matoušek. The determinant bound for discrepancy is almost tight. http://arxiv.org/abs/1101.0767, 2011.

[107] Jiří Matoušek and Aleksandar Nikolov. Combinatorial discrepancy for boxes via the ellipsoid-infinity norm. 2014.

[108] Darakhshan Mir, S. Muthukrishnan, Aleksandar Nikolov, and Rebecca N. Wright. Pan-private algorithms via statistics on sketches. In *PODS '11: Proceedings of the thirtieth ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pages 37–48, New York, NY, USA, 2011. ACM.

[109] S. Muthukrishnan. Data streams: Algorithms and applications. *Foundations and Trends in Theoretical Computer Science*, 1(2), 2005.

[110] S. Muthukrishnan and Aleksandar Nikolov. Optimal private halfspace counting via discrepancy. In *STOC '12: Proceedings of the 44th symposium on Theory of Computing*, pages 1285–1292, New York, NY, USA, 2012. ACM.

[111] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM J. Comput.*, 22(4):838–856, 1993.

[112] A. Narayanan and V. Shmatikov. De-anonymizing social networks. In *Security and Privacy, 2009 30th IEEE Symposium on*, pages 173–187. Ieee, 2009.

[113] Alantha Newman, Ofer Neiman, and Aleksandar Nikolov. Beck's three permutations conjecture: A counterexample and some consequences. In *FOCS '12: Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 253–262, Washington, DC, USA, 2012. IEEE Computer Society.

[114] Harald Niederreiter. *Random number generation and quasi-Monte Carlo methods*, volume 63 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[115] Aleksandar Nikolov. The komlos conjecture holds for vector colorings. *Submitted to Combinatorica*, 2013.

[116] Aleksandar Nikolov and Kunal Talwar. Approximating discrepancy via small width ellipsoids. 2013.

[117] Aleksandar Nikolov and Kunal Talwar. On the hereditary discrepancy of homogeneous arithmetic progressions. *Submitted to Proceedings of the AMS*, 2013.

[118] Aleksandar Nikolov, Kunal Talwar, and Li Zhang. The geometry of differential privacy: the sparse and approximate cases. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing*, STOC '13, pages 351–360, New York, NY, USA, 2013. ACM.

[119] A. Nilli. On the second eigenvalue of a graph. *Discrete Math.*, 91(2):207–210, 1991.

[120] M. L. Overton and R. S. Womersley. Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices. *Math. Programming*, 62(2, Ser. B):321–357, 1993.

[121] Yuval Rabani and Amir Shpilka. Explicit construction of a small epsilon-net for linear threshold functions. *SIAM J. Comput.*, 39(8):3501–3520, 2010.

[122] Prabhakar Raghavan and Clark D. Thompson. Randomized rounding: a technique for provably good algorithms and algorithmic proofs. *Combinatorica*, 7(4):365–374, 1987.

[123] G. Raskutti, M.J. Wainwright, and B. Yu. Minimax rates of estimation for high-dimensional linear regression over¡ formula formulatype=. *Information Theory, IEEE Transactions on*, 57(10):6976–6994, 2011.

[124] R. Tyrrell Rockafellar. *Convex analysis*. Princeton Mathematical Series, No. 28. Princeton University Press, Princeton, N.J., 1970.

[125] Aaron Roth and Tim Roughgarden. Interactive privacy via the median mechanism. In *Proceedings of the 42nd ACM symposium on Theory of computing*, STOC '10, pages 765–774, New York, NY, USA, 2010. ACM.

[126] K. F. Roth. On irregularities of distribution. *Mathematika*, 1:73–79, 1954.

[127] Klaus F Roth. Remark concerning integer sequences. *Acta Arithmetica*, 9:257–260, 1964.

[128] Thomas Rothvoß. Approximating bin packing within o(log OPT * log log OPT) bins. pages 20–29, 2013.

[129] Thomas Rothvoß. Constructive discrepancy minimization for convex sets. *CoRR*, abs/1404.0339, 2014.

[130] T.J. Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, pages 216–226, 1978.

[131] Wolfgang M. Schmidt. Irregularities of distribution. VII. *Acta Arith.*, 21:45–50, 1972.

[132] Alexander Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. B*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003. Matroids, trees, stable sets, Chapters 39–69.

[133] P. D. Seymour. Decomposition of regular matroids. *J. Combin. Theory Ser. B*, 28(3):305–359, 1980.

[134] Peter Shirley. Discrepancy as a quality measure for sample distributions. In *In Eurographics '91*, pages 183–194. Elsevier Science Publishers, 1991.

[135] Joel Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985.

[136] Joel Spencer. *Ten lectures on the probabilistic method*, volume 64 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1994.

[137] D.A. Spielman and N. Srivastava. An elementary proof of the restricted invertibility theorem. *Israel Journal of Mathematics*, pages 1–9, 2010.

[138] Aravind Srinivasan. Improving the discrepancy bound for sparse matrices: better approximations for sparse lattice approximation problems. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA, 1997)*, pages 692–701. ACM, New York, 1997.

[139] Nikhil Srivastava. Erdős discrepancy problem 22. `http://windowsontheory.org/2013/07/11/discrepancy-graphs-and-the-kadison-singer-conjecture-2/`, 2013.

[140] Salil P. Vadhan. Pseudorandomness. *Foundations and Trends in Theoretical Computer Science*, 7(1-3):1–336, 2012.

[141] T. van Aardenne-Ehrenfest. Proof of the impossibility of a just distribution of an infinite sequence of points over an interval. *Nederl. Akad. Wetensch., Proc.*, 48:266–271 = Indagationes Math. 7, 71–76 (1945), 1945.

[142] T. van Aardenne-Ehrenfest. On the impossibility of a just distribution. *Nederl. Akad. Wetensch., Proc.*, 52:734–739 = Indagationes Math. 11, 264–269 (1949), 1949.

[143] J.G. van der Corput. Verteilungsfunktionen. I. Mitt. *Proc. Akad. Wet. Amsterdam*, 38:813–821, 1935.

[144] J.G. van der Corput. Verteilungsfunktionen. II. *Proc. Akad. Wet. Amsterdam*, 38:1058–1066, 1935.

[145] Vijay V. Vazirani. *Approximation algorithms*. Springer, 2001.

[146] R. Vershynin. John's decompositions: Selecting a large part. *Israel Journal of Mathematics*, 122(1):253–277, 2001.

[147] Zhewei Wei and Ke Yi. The space complexity of 2-dimensional approximate range counting. In Sanjeev Khanna, editor, *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2013, New Orleans, Louisiana, USA, January 6-8, 2013*, pages 252–264. SIAM, 2013.

[148] Hermann Weyl. Über die gleichverteilung von zahlen mod. eins. *Mathematische Annalen*, 77(3):313–352, 1916.

[149] David P. Williamson and David B. Shmoys. *The Design of Approximation Algorithms*. Cambridge University Press, 2011.

[150] David P. Woodruff. Optimal space lower bounds for all frequency moments. In *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004, New Orleans, Louisiana, USA, January 11-14, 2004*, pages 167–175, 2004.

[151] David Paul Woodruff. *Efficient and private distance approximation in the communication and streaming models*. PhD thesis, Massachusetts Institute of Technology, 2007.

[152] Xiaokui Xiao, Guozhang Wang, and Johannes Gehrke. Differential privacy via wavelet transforms. In *ICDE*, pages 225–236, 2010.

[153] Andrew Chi-Chih Yao. Probabilistic computations: Toward a unified measure of complexity (extended abstract). In *18th Annual Symposium on Foundations of Computer Science, Providence, Rhode Island, USA, 31 October - 1 November 1977*, pages 222–227, 1977.

[154] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 209–213, 1979.

[155] Li Zhang. Nearly optimal minimax estimator for high dimensional sparse linear regression. *Annals of Statistics*, 2013. To appear.