Algorithmic Aspects of Combinatorial Discrepancy

Nikhil Bansal

Abstract This chapter describes some recent results in combinatorial discrepancy theory motivated by designing efficient polynomial time algorithms for finding low discrepancy colorings. Until recently, the best known results for several combinatorial discrepancy problems were based on counting arguments, most notably the entropy method, and were widely believed to be non-algorithmic. We describe some algorithms based on semidefinite programming that can achieve many of these bounds. Interestingly, the new connections between semidefinite optimization and discrepancy have lead to several new structural results in discrepancy itself, such as tightness of the so-called determinant lower bound and improved bounds on the discrepancy of union of set systems. We will also see a surprising new algorithmic proof of Spencer's celebrated six standard deviations result due to Lovett and Meka, that does not rely on any semidefinite programming or counting argument.

1 Introduction

In this chapter we consider the algorithmic aspects of several problems arising in discrepancy theory. In particular, we will focus on combinatorial discrepancy, which deals with the following type of question. There is a set-system (V,C) specified by the elements $V = \{1,\ldots,n\}$ and a collection of subsets $C = \{S_1,\ldots,S_m\}$ of V. Find a red-blue coloring of V such that each set in C is colored as evenly as possible.

Formally, let us use -1 and $+1$ to denote the colors red and blue. Given a coloring $\chi : V \to \{-1,1\}$, let us define the discrepancy of χ for a set S as $\mathrm{disc}(\chi,S) := |\chi(S)|$ where $\chi(S) := \sum_{i\in S}\chi(i)$. Note that $\chi(S)$ measures the imbalance between the number of elements in S that are colored -1 and 1. The discrepancy of the coloring χ for the system (V,C) is defined as the maximum discrepancy over all sets $S \in C$,

Nikhil Bansal
Eindhoven University of Technology, Eindhoven, the Netherlands e-mail: n.bansal@tue.nl

$$\mathrm{disc}(\chi, C) := \max_{S \in C} |\chi(S)|.$$

The discrepancy of the system (V,C) is the minimum discrepancy over all possible colorings, i.e.

$$\mathrm{disc}(C) = \min_{\chi} \max_{j \in [m]} |\chi(S_j)|.$$

More generally, one can define the discrepancy of a $m \times n$ matrix A as $\mathrm{disc}(A) := \min_{x \in \{-1,1\}^n} \|Ax\|_\infty$. Clearly, if A is the incidence matrix of a set system, then this is precisely the discrepancy of the set system as defined above. Almost all of the results that we consider in this chapter generalize to arbitrary matrices in a straightforward way. However we will focus on the case of set systems to keep the notation simple, and also for historical reasons, as most results in combinatorial discrepancy are stated for set systems.

Roughly speaking, the discrepancy of a set system is a useful measure of its inherent complexity, and hence its understanding it is related to several areas in mathematics and theoretical computer science. In computer science for example, discrepancy is useful in topics such as probabilistic and approximation algorithms, computational geometry, numerical integration, derandomization, complexity theory, data structures and so on. We discuss one such application in section 1.1, but for more details we refer the reader to [13, 10, 21].

Algorithmic aspects: Motivated by these applications, several interesting and non-trivial techniques have been developed for both upper bounding and lower bounding the discrepancy of various classes of set systems. As we shall see, many of these techniques are non-constructive in the sense that they prove the existence of a low discrepancy coloring, but give no clue about how to find one *efficiently*. As usual, an algorithm is called efficient if its running time on every input instance is polynomial in the size of the description of the instance. For us, this means that the algorithm must run in time polynomial in n and m. In particular, the algorithm that enumerates over all possible 2^n colorings and picks the best one is not efficient.

Designing efficient algorithms for finding low discrepancy colorings has several motivations. First, in many applications (see section 1.1) one actually needs to find a such a coloring efficiently. Another motivation is from theoretical computer science where one wishes to understand which problems admit efficient algorithms and which ones do not. Until recently, many techniques for upper-bounding discrepancy were believed to be inherently non-algorithmic.

Approximating Discrepancy: Perhaps the most natural question is the following. Given a set system (V,C) can we determine its discrepancy exactly, or perhaps even approximately, in polynomial time? Unfortunately, it turns out that this general question is essentially hopeless in a very strong sense. In particular, recently Charikar, Newman and Nikolov [9] showed the following.

Theorem 1. *Given a set system on n elements and $m = O(n)$ sets, it is NP-hard to distinguish whether the system has discrepancy* 0 *or* $\Omega(\sqrt{n})$.

In particular, theorem 1 says that assuming P≠NP, no polynomial time algorithm can distinguish even among the following two extreme cases (i) whether a given set system with $O(n)$ sets has discrepancy 0 or (ii) has discrepancy at least $c\sqrt{n}$ for some universal constant c.

We assume that the reader is familiar with basic notions of computational complexity such as P≠ NP. For more details, an excellent reference is [26].

Theorem 1 is surprisingly tight, as the celebrated "six standard deviations suffice" result of Spencer [27] shows that

Theorem 2. *Every set system with $m = n$ sets has discrepancy at most $6\sqrt{n}$. More generally, for $m > n$, the discrepancy is $O((n\log(m/n))^{1/2})$.*

We shall prove theorems 1 and 2 in sections 6 and 2.2.

Hereditary Discrepancy: One reason why discrepancy is so hard to estimate (in the sense of theorem 1) is that it is a very *fragile* quantity. The following example is instructive.

Let (V,C) be a set system with high discrepancy, where $V = \{1,\ldots,n\}$ and $C = \{S_1,\ldots,S_m\}$. Let (V',C') be a copy of (V,C) on a different ground set. That is, let $V' = \{n+1,\ldots,2n\}$ and $C' = \{S'_1,\ldots,S'_m\}$ where each $S'_i = \{n+j : j \in S_i\}$ is a copy of S_i on V'. Consider the system $(W,D) := ((V \cup V', \{S_1 \cup S'_1,\ldots,S_m \cup S'_m\})$, that is, with elements $V \cup V'$ and with i-th set $S_i \cup S'_i$. As the system (W,D) contains the system (V,C) (restricting (W,D) to V gives (V,C)), it is at least as complex as C, and hence one would expect its discrepancy to be no less than that of (V,C). However, (W,D) has discrepancy zero for trivial reasons, as one can color all the elements in V by 1 and those in V' by -1.

To get around these kinds of anomalies, it is very useful to define the *hereditary* discrepancy of a set system (V,C). Specifically, given $V' \subseteq V$, let $C|_{V'}$ denote the collection $\{S \cap V' : S \in C\}$. Then, the hereditary discrepancy of (V,C) is defined as

$$\mathrm{herdisc}(C) = \max_{V' \subseteq V} \mathrm{disc}(C|_{V'}).$$

Most upper bounds stated in terms of discrepancy also imply the same bound for hereditary discrepancy (for example if a class of systems is closed under taking restrictions of the ground set, then bounding the discrepancy of systems in this class, also implies bounds on hereditary discrepancy).

Let us consider an application to see how the concepts of discrepancy and hereditary discrepancy can be useful.

1.1 An Application

Suppose we are given a fractional solution $x \in \mathbb{R}^n$, to a linear system $Ax = b$ on n variables and m constraints. We would like to *round* x to an integral solution $\tilde{x}$ such

that the error in each of the m equations is as low as possible, i.e. find $\tilde{x} \in \mathbb{Z}^n$ that minimizes $\|A(x-\tilde{x})\|_\infty$.

The answer to this question turns out to be closely related to the hereditary discrepancy of the matrix A.

Theorem 3 (Lovász, Spencer, and Vesztergombi [18]). *For any $x \in \mathbb{R}^n$ satisfying $Ax = b$, there is a $\tilde{x} \in \mathbb{Z}^n$ with $\|\tilde{x}-x\|_\infty < 1$, such that $\|A(x-\tilde{x})\|_\infty \leq \text{herdisc}(A)$.*

Proof. Let $x = (x_1, \ldots, x_n)$, and consider the binary expansion of each x_i. That is, we write x_i as $\lfloor x_i \rfloor + \sum_{j \geq 1} q_{ij} 2^{-j}$, where $q_{ij} \in \{0,1\}$ denotes the j-th bit of x_i after the decimal point. The idea of the proof is to round the bits q_{ij} to 0, while introducing low error in each of the m equations.

Let k be a fixed positive integer. Consider the k-th bit q_{ik} of each x_i. Let $A^{(k)}$ be the sub-matrix of A restricted to those columns i for which $q_{ik} = 1$. By the definition of hereditary discrepancy, there is a $\{-1,1\}$ coloring $\chi^{(k)}$ of the columns of $A^{(k)}$ such that the discrepancy of each row of $A^{(k)}$ is at most herdisc(A). Viewing $\chi^{(k)}$ as a vector in $\mathbb{R}^n$ with entries $\{-1,0,1\}$ (where the 0 entries correspond to the columns not in $A^{(k)}$), consider the vector $x' = x + 2^{-k}\chi^{(k)}$. Now, the k-th bit of each x'_i is 0, as $\chi^{(k)}(i) \in \{-1,1\}$ if $q_{ik} = 1$ and 0 if $q_{ik} = 0$. Moreover, $Ax' - Ax = A(x'-x) = 2^{-k}A\chi^{(k)}$ and hence $\|Ax - Ax'\|_\infty \leq 2^{-k} \cdot \text{herdisc}(A)$.

We now iterate this process (treating x' as x) on the bits at position $k-1, k-2, \ldots, 1$. This produces a vector $x' = x + 2^{-k}\chi^{(k)} + 2^{-k+1}\chi^{(k-1)} + \ldots + 2^{-1}\chi^{(1)}$, where the first k bits of each x'_i are 0, and $|Ax - Ax'| \leq \text{herdisc}(A)(2^{-k} + 2^{-k+1} + \ldots 2^{-1}) < \text{herdisc}(A)$. Making k arbitrarily large implies the final result. □

Let us note a few things about the proof, as these ideas will also be useful later. (i) For each bit position j, the coloring $\chi^{(j)}$ is used as a guide whether to the round the bit q_{ij} up or down. (ii) After the update is applied to the j-th bit, the bits in positions $j-1$ or earlier might change due to carry-overs, and hence the submatrix $A^{(j-1)}$ and the coloring $\chi^{(j-1)}$ at the next step depend on the previous choices $\chi^j, \chi^{(j+1)}, \ldots$. (iii) The proof does not give an efficient algorithm to find $\tilde{x}$, as the low hereditary discrepancy property only shows the existence of some good χ^j.

We remark that Doerr [12] gives an improved error bound of $(1 - 1/2m)\text{herdisc}(A)$ in theorem 3.

1.2 Chapter Overview

The chapter is organized as follows. We first describe some classical methods for both upper bounding and lower bounding the discrepancy of various types of set systems. We then discuss the more recent results, that will build upon these previous ideas. To keep our discussion manageable, in this chapter we mostly focus on two problems which will suffice to convey the main ideas.

Arbitrary set systems: Given n elements and an arbitrary collection of $m = n$ sets, find a minimum discrepancy coloring.

Bounded degree set systems: Given n elements and an arbitrary collection of sets such that each element lies in at most t sets, find a minimum discrepancy coloring.

We now give a brief overview of the topics and the results that we will consider.

1.2.1 Classical upper bound methods

Random Coloring: Given a set system (V,C), perhaps the simplest idea is to color each element randomly and independently ± 1 with probability $1/2$ each.

Lemma 1. *For any set system (V,C) on n elements and m sets, a random coloring has discrepancy $O(\sqrt{n\log m})$ with high probability.*

Proof. For any set $S \in C$, $\chi(S)$ is a random variable with mean 0 and variance $|S|$. Thus, by standard Chernoff bounds (see for e.g. [1]), $\Pr[|\chi(S)| \geq c\sqrt{|S|}] \leq 2\exp(-c^2/2)$. Choosing, say $c = 2\sqrt{\log m}$, this probability is at most $2/m^2$, and hence by a union bound over the m sets in C, $\max_S |\chi(S)| \leq 2\sqrt{n\log m}$ with probability at least $1-2/m$. □

In particular, for $m = n$ this gives a discrepancy of $O(\sqrt{n\log n})$. This is reasonably good as there exist set systems on n sets with discrepancy at least $\Omega(\sqrt{n})$ (we will see this in section 6). However, a random coloring is not very interesting as it is completely oblivious to the problem structure. For example, even for bounded degree set systems with $t = O(1)$, a random coloring only gives $\Omega(\sqrt{n})$ discrepancy, while the actual discrepancy is $O(1)$ (see theorem 4 below). Moreover, note that theorem 2 always beats random coloring for any set system.

Linear Algebraic Method: A simple but powerful approach for bounding discrepancy is based on basic linear algebra. An interesting application of this method is the following result for bounded degree set systems.

Theorem 4 ([6]). *If A is a set system where each element lies in at most t sets, then* $\mathrm{disc}(A) \leq 2t-1$.

We shall see the proof of this theorem is section 2.1. While the idea itself is quite simple, the underlying idea will play a key role in later results.

The Entropy Method: One of the most powerful and widely used tools in combinatorial discrepancy is the so-called partial coloring method due to [5] and its refinements [27] based on the so-called entropy method. For example, Spencer's original proof of theorem 2 was based on the entropy method. It can also be used to prove a $O(\sqrt{t\log n})$ bound [28] for bounded degree systems[1], which can sometimes be better than the bound in theorem 4.

[1] Interestingly, an improved $O(\sqrt{t\log n})$ bound is also known [2] using a different method based on convex geometry.

As we shall see in section 2.2, the entropy method is based on a clever application of the pigeonhole principle to the space of all 2^n possible colorings, and only proves the existence of a low discrepancy coloring, without giving any algorithmic insight on how to find it efficiently. For example, even though theorem 2 was long known, no polynomial time algorithm better than random coloring was known until recently for general set systems.

1.2.2 A lower bound method

A variety of techniques have also been developed for proving lower bounds on discrepancy. Many of these are based on deep results and connections to various areas of mathematics (see for e.g. [10, 21]). One of the strongest known results in this direction is the following *determinant lower bound* due to [18]. For a real matrix A, define

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

where the maximum is over all $k \times k$ submatrices B of A. For a set system (V,C), let $\mathrm{detlb}(C)$ denote $\mathrm{detlb}(A)$ where A is the incidence matrix of (V,C).

Theorem 5 ([18]). *For every set system C,* $\mathrm{herdisc}(C) \geq \frac{1}{2}\mathrm{detlb}(C)$.

We shall prove theorem 5 in section 7.1. The proof is based on an interesting geometric interpretation of hereditary discrepancy which will be useful later.

1.2.3 Some recent results

Recently, several advances have been made in combinatorial discrepancy theory. First, many previous (non-constructive) results based on the entropy method, can now be made algorithmic. In particular, Bansal [3] showed the following algorithmic version of Spencer's result.

Theorem 6. *[3] For any set system on $m = O(n)$ sets, there is a randomized polynomial time algorithm to find an $O(\sqrt{n})$ discrepancy coloring.*

This result is based on semidefinite programming (SDPs) and in particular a new method to round SDP solutions based on designing correlated gaussian random walks. This method has several other applications. For example, it gives an algorithm to find $O(\sqrt{t}\log n)$ discrepancy coloring for bounded degree systems, matching the bound of [28]. It also gives a good approximation algorithm to find a low discrepancy coloring for systems with low hereditary discrepancy.

Theorem 7 ([3]). *For any set system (V,C) on n elements and m sets, there is a randomized polynomial time algorithm to find a coloring with discrepancy at most $O((\log m \log n)^{1/2}) \cdot \mathrm{herdisc}(C)$.*

Theorem 7 directly implies the following algorithmic version of theorem 3.

Corollary 1. *Given any fractional solution $x \in \mathbb{R}^n$ to the system $Ax = b$ on m equations, there is a polynomial time algorithm to round x to $\tilde{x} \in \mathbb{Z}^n$ such that* $\|A(x-\tilde{x})\|_\infty = O((\log m \log n)^{1/2}) \cdot \text{herdisc}(A)$.

In section 3 we give the relevant background on semidefinite programming and prove theorem 7. Then we show how these ideas can be refined and combined with the entropy method to obtain theorem 6. However, instead of proving the $O(\sqrt{n})$ bound, we show a weaker $O((n \log\log\log n)^{1/2})$ bound (which is already much stronger than random coloring). This weaker bound illustrates all the ideas involved without the tedious calculations needed for the $O(\sqrt{n})$ bound.

Somewhat surprisingly, even though the algorithm in theorem 6 is polynomial time, it crucially relies on the non-constructive entropy method in its design, and in particular does not give a truly constructive proof of Spencer's result. This unsatisfying situation was resolved very recently by Lovett and Meka [19] who gave a simpler and completely constructive proof of Spencer's result based on gaussian random walks and linear algebraic ideas. In particular, their proof does not use entropy method. We will see their proof in section 5. Interestingly, their main result implies a variant of the partial coloring that is quantitatively stronger (in a certain regime) than the one obtained by the entropy method. This variant can also be viewed as a "robust" version of the so-called iterated rounding technique in approximation algorithms, and has found some other very interesting algorithmic uses recently [25].

The connections between discrepancy and semidefinite programming have also been useful in other ways besides algorithm design. In a very interesting result, Matoušek [20] gave the first non-trivial upper bound on the gap between the determinant lower bound and hereditary discrepancy.

Theorem 8 ([20]). *For any set system (V,C),* $\text{herdisc}(C) \leq O(\log n \sqrt{\log m}) \cdot \text{detlb}(C)$.

This result is remarkably tight, as there exist set systems for which $\text{herdisc}(C) = \Omega(\log n) \cdot \text{detlb}(C)$ [23]. The proof of theorem 8 is based on SDP duality, and relating the dual SDP solution to sub-determinants of A via eigenvalues. We will prove theorem 8 in section 7.1.

Among other things, theorem 8 also implies the following new structural result in discrepancy [20]: For any two set systems (V,C_1) and (V,C_2),

$$\text{herdisc}(C_1 \cup C_2) \leq \max(\text{herdisc}(C_1), \text{herdisc}(C_2)) \cdot O(\log n \sqrt{\log m}).$$

Previously such a result was known only for the special case when C_2 consists of a single set [16]! A further extension of this result to the union of t set-system can also be found in [20].

Some related results that we do not discuss: A very recent and surprising result that we do not discuss in this chapter, is the first algorithm for approximating hereditary discrepancy to within poly-logarithmic factors [22]. Note that apriori it is not

even clear whether the minimum hereditary discrepancy problem is in NP.[2] For this to hold, given a set system (V,C) and a target hereditary discrepancy λ, there must exist a (short) polynomial time verifiable witness that certifies that $\text{disc}(C|_J) \leq \lambda$ for each of the 2^n subsets $J \subseteq V$. The result of [22] is based on combining the geometric interpretation of hereditary discrepancy together with powerful results and ideas from convex geometry, most notably the restricted invertibility principle by Bourgain and Tzafriri [7] and its refinements due to Vershynin [30].

Finally, the recent algorithmic ideas developed for addressing discrepancy related questions have also led to some very interesting results in optimization and algorithms (e.g. [14, 24, 25, 8]. However, a discussion of these is beyond the scope of this chapter.

2 Some Classic Results

We describe some classical results and techniques which will be very useful later.

2.1 Linear Algebraic Method

Beck and Fiala proved the following bound on the discrepancy of bounded degree set systems, based on the linear algebraic method. Even though very elementary, this will be an important proof for us, as its high level idea will be used many times later.

Theorem 9 ([6]). *If (V,C) is a set system such that each element lies in at most d sets, then* $\text{disc}(S) \leq 2d-1$. *Moreover, such a coloring can be found in polynomial time.*

Proof. We will give an algorithm that starts with the all 0-coloring χ_0, i.e. $\chi_0(i) = 0$ for all i, and iteratively updates the colors over time until they all reach -1 or 1. During the intermediate steps, the elements may be assigned a fractional coloring, i.e. in $[-1,1]$ instead of $\{-1,1\}$. We now describe how the algorithm proceeds.

Let $\chi_t = (\chi_t(1),\ldots,\chi_t(n))$ denote the coloring after the t-th iteration. Initially, $\chi_0(i) = 0$ for all i. We say that i is *alive* at time t, if $|\chi_{t-1}(i)| < 1$. Otherwise, if $\chi_{t-1}(i) \in \{-1,1\}$, it is considered *fixed* and is never updated again. Call a set S *safe* at time t if at most d elements of S are alive. Otherwise, we call S *dangerous*.

The algorithm will ensure that at least one alive variable becomes fixed in each iteration, and hence it terminates in at most n steps. As an element's color is not updated once it is fixed, once a set becomes safe it stays safe henceforth. The coloring

[2] Observe that proving a c-approximation for a problem, implies that the (approximate) problem is both in NP and co-NP. Note that both theorem 5 and theorem 7 can be used to give a co-NP witness.

is updated in each iteration as follows. At iteration t, let C^t denote the collection of dangerous sets. The crucial observation is that the number of dangerous sets is strictly less than the number of alive elements. This follows as each (alive) element lies in at most d sets, and each dangerous set contains strictly more than d alive elements.

Let $v = (v_1, \ldots, v_n)$, and consider the linear system defined by $S_j \cdot v = 0$ (i.e. $\sum_{i \in S_j} v_i = 0$) for each dangerous set S_j and $v_i = 0$ for each element i that is fixed. By the observation above, there are strictly less than n constraints and hence by basic linear algebra, there must exist a non-zero update direction $v = (v(1), \ldots, v(n))$ such that $v(i) = 0$ if i is fixed, and $S_j \cdot v = 0$ if j is dangerous. We set $\chi_t = \chi_{t-1} + \delta v$ where $\delta > 0$ is the smallest real such that some alive variable reaches -1 or 1 and gets fixed.

To see that the discrepancy of any set S is at most $2d-1$, observe that as long as a set is dangerous, the update rule ensures that $S \cdot \chi_t = 0$. However, once S becomes safe, it has at most d alive variables and thus no matter how these variables will be subsequently updated, the discrepancy added with be strictly less than $2d$. As the discrepancy is an integer, it can be at most $2d-1$ □

Let us note some key points about the proof. (i) The algorithms works with fractional colorings at intermediate steps. (ii) At each step it makes progress towards becoming an integral coloring. (iii) During the algorithm a set is protected as long as it is dangerous.

2.2 Entropy Method

The following result is known as the partial coloring lemma and is one of the most widely used techniques in combinatorial discrepancy. Its proof is based on a refined counting approach called the entropy method, and a clever pigeonhole principle argument. We closely follow the exposition in [21].

Theorem 10 (Partial Coloring Lemma via the Entropy Method). *Let (V,C) be a set system on n elements, and let a number $\Delta_S > 0$ be given for each set $S \in C$. Suppose the Δ_S satisfy the condition*

$$\sum_{S \in C} g\left(\frac{\Delta_S}{\sqrt{|S|}}\right) \leq \frac{n}{5} \tag{1}$$

where

$$g(\lambda) = \begin{cases} Ke^{-\lambda^2/9} & \text{if } \lambda > 0.1 \\ K\ln(\lambda^{-1}) & \text{if } \lambda \leq 0.1 \end{cases}$$

and K is some absolute constant. Then there is a partial coloring χ that assigns ± 1 to at least $n/10$ variables (and 0 to the rest), satisfying $|\chi(S)| \leq \Delta_S$ for each $S \in C$.

We begin with some standard results that we will need to prove theorem 10.

Entropy: Let Y be a discrete random variable that takes value y with probability p_y. Then, its *entropy* is defined as $H(Y) := \sum_y p_y \log_2(1/p_y)$.

Entropy satisfies the following properties (see e.g. [11]).

1. If $H(Y) \leq k$, then $p_y \geq 2^{-k}$ for some value y.
2. If Y attains ℓ different values, then $H(Y) \leq \log_2 \ell$. The equality is attained iff $Y = U_\ell$, the uniform distribution on ℓ values.
3. Subadditivity: If $Y_1, \ldots, Y_m$ are arbitrary (correlated) random variables, and $X = (Y_1, \ldots, Y_m)$ is a random vector with components $Y_1, \ldots, Y_m$, then $H(Y) \leq \sum_i H(Y_i)$.

Roughly speaking, the entropy $H(X)$ measures the amount of randomness in X.

Lemma 2. *Let $k = 2^{4n/5}$ and suppose $\chi_1, \ldots, \chi_k$ are k distinct ± 1 colorings of $[n]$, such that for every two colorings χ_i and χ_j, with $i, j \in [k]$, it holds that $|\chi_i(S) - \chi_j(S)| \leq 2\Delta_S$ for each set $S \in C$. Then, there exists a partial coloring χ that assigns ± 1 to at least $n/10$ elements (and 0 to the rest) satisfying $\chi(S) \leq \Delta_S$ for each $S \in C$.*

Proof. This follows from the standard isoperimetric inequality for the hamming cube [17], which states that any subset $C \subset \{-1,1\}^n$ of the cube with $|C| > \sum_{j=0}^{\ell} \binom{n}{j}$ contains two points in C with hamming distance at least 2ℓ.

As $2^{4n/5} > \sum_{h=0}^{n/20} \binom{n}{h}$ this implies that there must exist two colorings χ_i, χ_j in $\{\chi_1, \ldots, \chi_k\}$ with hamming distance at least $n/10$. Now, consider the vector $\chi = (\chi_i - \chi_j)/2$. Note that $\chi(\ell) = \pm 1$ whenever $\chi_i(\ell) \neq \chi_j(\ell)$, and is 0 otherwise. Moreover, for any set $S \in C$

$$|\chi(S)| = |\sum_{\ell \in S} \chi(\ell)| = |\frac{1}{2} \sum_{\ell \in S} (\chi_i(\ell) - \chi_j(\ell))| = |\frac{1}{2}(\chi_i(S) - \chi_j(S))| \leq \Delta_S.$$

Here the last inequality follows from the property of the colorings $\chi_1, \ldots, \chi_k$. □

We now prove theorem 10.

Proof. (Theorem 10) For a ± 1 coloring χ and a set $S \in C$, let $Y_\chi(S) := \text{round}\left(\frac{\chi(S)}{2\Delta_S}\right)$, where $\text{round}(x) = \lfloor x + 1/2 \rfloor$ is the rounding function to the nearest integer. Thus $Y_\chi(S)$ simply indicates which bucket of size $2\Delta_S$ the discrepancy $\chi(S)$ lies in.

Let $Y(S)$ be the random variable that takes value $Y_\chi(S)$ where the coloring χ is chosen uniformly at random among the 2^n possible colorings of $[n]$. We can determine $Y(S)$ exactly. In particular, if χ is a random coloring, then $\Pr[\chi(S) = k] = \binom{|S|}{(|S|-k)/2} 2^{-|S|}$ which is roughly $\exp(-k^2/(2|S|))$. In particular, $\chi(S)$ is distributed approximately uniformly in $[-\sqrt{S}, \sqrt{S}]$ and then decays super-exponentially in subsequent intervals of size $\sqrt{|S|}$. Let $\lambda_S := \Delta_S/\sqrt{|S|}$.

Claim. The entropy of Y_S satisfies $H(Y_S) \leq g(\lambda_S)$, for g as defined in theorem 10.

Proof. (Sketch) We refer the reader to [21] for the precise calculation, but the idea is the following. For $\lambda_S \leq 0.1$, Y_S is distributed essentially uniformly in

$[-1/\lambda_S, 1/\lambda_S]$ and the probability that Y_S takes values outside this range decreases super-exponentially. Thus $H(Y_S) \approx H(U_{2\lambda_S}) = O(\log(1/\lambda_S))$.

On the other hand if λ_S is large (say $\lambda_S = 10$), then $p_0 = \Pr[Y_S = 0] \approx 1 - \exp(-\lambda_S^2/2)$ (and hence very close to 1), and for $\ell > 1$, $\Pr[|Y_S| = \ell] \leq 2\exp(-\lambda_S^2\ell^2/2)$. Thus Y_S is essentially always 0 and has entropy roughly $p_0 \log(1/p_0) \approx \log(1/p_0) = O(\exp(-\lambda_S^2/2)$. Together this gives, $H(Y_S) \leq g(\lambda_S)$ for any set S. □

For $Y(S)$ as defined above, let Y denote the random vector $Y = (Y(S_1), \ldots, Y(S_m))$ where $S_1, \ldots, S_m$ are the sets in C. By sub-additivity of entropy and the claim above $H(Y) \leq \sum_j H(Y_j) \leq \sum_{S \in C} g(\lambda_S)$. As $\sum_S g(\lambda_S) \leq n/5$, this gives $H(Y) \leq n/5$ and hence Y attains some value $b = (b_1, \ldots, b_m)$ with probability at least $2^{-n/5}$. Equivalently, there exist $k \geq 2^{4n/5}$ different colorings χ_i, $i = 1, \ldots, k$ such that $Y(\chi_i(S_j)) = b_j$ for each $i = 1, \ldots, k$ and $j = 1, \ldots, m$, and thus for every $1 \leq i, i' \leq k$ and $j = 1, \ldots, m$ it holds that $|\chi_i(S_j) - \chi_{i'}(S_j)| \leq 2\Delta_S$. Applying lemma 2 now gives the result. □

Let us see how theorem 10 implies Spencer's result and the $O(\sqrt{t}\log n)$ bound for bounded-degree set systems.

Proof of theorem 2: The coloring is constructed in phases. Let $n_0 = n$ and let n_i denote number of uncolored elements left at the beginning of phase i, for $i = 0, 1, \ldots$. In phase i, we apply theorem 10 to these n_i elements with $\Delta_S^i = c(n_i \log(2n/n_i))^{1/2}$ and verify that (1) holds when c is a large enough constant. This gives a partial coloring on at least $n_i/10$ elements, with discrepancy for any set S at most Δ_S^i. This gives that $n_{i+1} \leq (0.9)n_i$ and hence $n_i \leq (0.9)^i n$. Summing up over the phases, the discrepancy for any set is at most

$$\sum_i \Delta_S^i \leq \sum_i c\left(n(0.9)^i \log\left(\frac{2n}{n(0.9)^i}\right)\right)^{1/2} = O(n^{1/2}).$$

Proof of $O(\sqrt{t}\log n)$ discrepancy for bounded degree systems [28]: The coloring is constructed in phases where at most $n_i \leq n(0.9)^i$ elements are left uncolored in phase i. In phase i, let $s_{i,j}$ denote the number of sets with the number of uncolored elements in the range $[2^j, 2^{j+1})$. As the degree of the set system is at most t, we have that $s_{i,j} \leq \min(m, n_i t/2^j)$. Using this fact, it can be verified that (1) holds if $\Delta_S = ct^{1/2}$ for some large enough constant c. Thus each set incurs $O(\sqrt{t})$ discrepancy in each phase and hence the total discrepancy incurred is $O(t^{1/2}\log n)$.

3 Systems with Low Hereditary Discrepancy

In this section we prove theorem 7.

Theorem 7. *For any set system (V, C) on n elements and m sets, there is a randomized polynomial time algorithm to find a coloring with discrepancy $O((\log m \log n)^{1/2} \cdot \text{herdisc}(C))$.*

The algorithm will be based on semidefinite programming, so we first give a brief overview of semidefinite programming.

Semidefinite Programming: Recall that a linear program (LP) consists of some collection of variables $x_1, \ldots, x_n$ where each x_i takes values in $\mathbb{R}$, and the goal is to optimize some linear objective function $\sum_i c_i x_i$, subject to some linear constraints $\sum_i a_{ji} x_i \leq b_j$ for $j = 1, \ldots, m$. Linear programs can be solved optimally in time that is polynomial in n, m, and the bit length required to describe the entries a_{ji}, b_j, c_i.

A semidefinite program (SDP) can be viewed as the following linear program. The variables are written in the form x_{ij} where $1 \leq i, j \leq n$ (the reason for writing variables in this form will become clear soon). There are m arbitrary linear constraints on x_{ij} of the type $\sum_{ij} a_{ij}^k x_{ij} \leq b^k$ for $k = 1, \ldots, m$ and there is some linear objective function $\sum c_{ij} x_{ij}$. Moreover, we require the symmetry condition $x_{ij} = x_{ji}$. Finally, one imposes the constraint that the $n \times n$ matrix $X = (x_{ij})$, consisting of entries x_{ij}, be positive semidefinite. That is, all its eigenvalues must be non-negative (this is well defined as X is symmetric and hence has only real eigenvalues), and we denote this by $X \succeq 0$. To summarize, a general SDP has the following form.

$$\begin{aligned} \min \quad & \sum c_{ij} x_{ij} & \\ s.t. \quad & \sum_{ij} a_{ij}^k x_{ij} \leq b^k, & 1 \leq k \leq m \\ & X \succeq 0 & \\ & x_{ij} = x_{ji}, & 1 \leq i, j \leq n \end{aligned}$$

where a_{ij}^k, b^k, c_{ij} are arbitrary real numbers.

While the condition $X \succeq 0$ may appear non-linear, note that it can be enforced by adding (infinitely many) linear constraints of the form $a^T X a \geq 0$, one for each vector $a \in \mathbb{R}^n$. Despite the infinitely many constraints, by standard optimization theory and in particular the Ellipsoid method, this program can be solved to any desired level of accuracy in polynomial time. In particular, given a candidate solution X there exists an efficient separation procedure as one can determine in polynomial time whether $a^T X a < 0$ for some a, by computing the least eigenvalue of X and checking if it is negative. For more details about solving semidefinite programs, we refer the reader to [29].

Vector Program view: Recall that a symmetric $n \times n$ matrix X is positive semidefinite if and only if it is the Gram matrix of some vectors $v_1, \ldots, v_n \in \mathbb{R}^n$. That is, each entry x_{ij} can be written as $x_{ij} = \langle v_i, v_j \rangle$ where $\langle , \rangle$ denotes the standard inner product. Moreover, given X, the vectors v_i can be computed in polynomial time using the Cholesky decomposition procedure.

This implies that an SDP can equivalently be viewed as an arbitrary linear program where the variables correspond to inner product of vectors. This is referred to as the vector program view of an SDP, and it will be extremely useful for our

purposes. Note that one can only impose constraints on the dot products of v_i's and not on the vectors v_i's themselves. Let us see how this is useful for discrepancy.

SDP relaxation for discrepancy: The natural SDP relaxation for the problem of finding a ± 1 coloring with discrepancy at most λ is the following.

$$\|\sum_{i \in S_j} v_i\|_2^2 \leq \lambda^2 \qquad \text{for each set } S_j \in C \tag{2}$$

$$\|v_i\|_2^2 = 1 \qquad \text{for each element } i \in V. \tag{3}$$

Here, as usual, $\|v\|_2 = (\langle v, v \rangle)^{1/2}$ denotes the length of v. The first constraint (2) says that the discrepancy of each set S_j must be at most λ. Observe that this is a linear constraint on the dot product of variables as the left hand side of (2) can be written as $\sum_{i \in S_j, i' \in S_j} \langle v_i, v_{i'} \rangle$. The second constraint says that each $\langle v_i, v_i \rangle = 1$, i.e. each v_i must be a unit vector.

This is a valid relaxation as any ± 1 coloring with discrepancy at most λ is a feasible solution to the above program (corresponding to the solution $v_i = (1, 0, \ldots, 0)$ or $v_i = (-1, 0, \ldots, 0)$ depending on whether i is colored 1 or -1). We will call any feasible solution to this SDP, a *vector-coloring* for C, and the smallest value λ for which this SDP is feasible as the *vector discrepancy* of C, denoted by $\text{vecdisc}(C)$. Clearly, $\text{vecdisc}(C) \leq \text{disc}(C)$.

Using this SDP: Before we describe our algorithm, we describe a natural approach for using this SDP that does not work, but it will give important insights.

Let A be a matrix with discrepancy λ (the reader should think of λ as being very small, say 0). First, we can assume that the algorithm knows λ, as it can try all values $0, 1, \ldots, n$ and pick the smallest λ for which the SDP given by (2)-(3) is feasible. Now let us consider some vector-coloring v_i obtained by solving this SDP. In this solution, the unit vectors v_i will be nicely correlated such that for every set S_j, the vector $\sum_{i \in S_j} v_i$ has length at most λ (say 0).

Our goal then is to convert these vectors v_i into the numbers ± 1 without increasing $\sum_{i \in S_j} v_j$ too much. A natural first step is to try to convert v_i into real numbers (hopefully close to ± 1) without substantially violating the sums $\sum_{i \in S_j} v_j$. So we project the vectors v_i on some vector $g \in \mathbb{R}^n$ to get real numbers $y_i = \langle g, v_i \rangle$. This seems reasonable as this maintains the correlations among v_i. In particular, these y_i satisfy

$$\sum_{i \in S_j} y_i = \sum_{i \in S_j} \langle g, v_i \rangle = \langle g, \sum_{i \in S_j} v_i \rangle \leq \|g\|_2 \cdot \|\sum_{i \in S_j} v_i\|_2,$$

implying that $\sum_{i \in S_j} y_i$ is also small if $\|\sum_{i \in S_j} v_i\|$ is small. To this end, it will be very convenient to project the v_i on to random gaussian vectors in $\mathbb{R}^n$.

Gaussian Random Variables: We recall the following standard facts about gaussian distributions.

1. The gaussian distribution $N(\mu,\sigma^2)$ with mean μ and variance σ^2 has probability distribution function
$$f(x) = \frac{1}{(2\pi)^{1/2}\sigma} e^{-(x-\mu)^2/2\sigma^2}.$$
2. If X is distributed as $N(0,\sigma^2)$, then $\Pr[|X| \geq t\sigma] \leq 2e^{-t^2/2}$ for any $t \geq 1$.
3. *Additivity:* If $g_1 \sim N(\mu_1,\sigma^2)$ and $g_2 \sim N(\mu_2,\sigma_2^2)$ are independent gaussian random variables, then for any $t_1,t_2 \in \mathbb{R}$, the random variable $t_1 g_1 + t_2 g_2$ is distributed as $N(t_1\mu_1 + t_2\mu_2, t_1^2\sigma_1^2 + t_2^2\sigma_2^2)$.

The additivity property of gaussians implies the following useful property.

Lemma 3. *Let $g \in \mathbb{R}^n$ be a random gaussian, i.e. each coordinate is chosen independently according to distribution $N(0,1)$. Then for any arbitrary vector $v \in \mathbb{R}^n$, the random variable $\langle g,v\rangle \sim N(0,\|v\|_2^2)$.*

Proof. As $\langle g,v\rangle = \sum_i g(i)v(i)$, where $g(i)$ and $v(i)$ denote the i-th coordinate of g and v, and as the $g(i)'s$ are independent, the additivity property implies that $\langle g,v\rangle$ is distributed as $N(0,\sum_i v(i)^2) = N(0,||v||_2^2)$. $\square$

If the vectors v_i satisfy (2)-(3), then if we choose a random gaussian g and let $y_i = \langle g, v_i\rangle$, lemma 3 implies that

1. Each y_i is distributed as $N(0,1)$. This follows as $\|v_i\|_2^2 = 1$.
2. For each j, the discrepancy $\sum_{i\in S_j} y_i$ is distributed as $N(0,\leq \lambda^2)$, i.e. as a gaussian with mean 0 and variance at most λ^2. This follows as $\|\sum_{i\in S_j} v_i\|^2 \leq \lambda^2$ by the SDP constraint.

This seems quite close to what we would like. As $y_i \sim N(0,1)$, we have that $y_i/(c(\log n)^{1/2}) \in [-1,1]$ with high probability (for some large enough constant c). Moreover, for any j, the discrepancy $|\sum_{i\in S_j} y_i| = O(\lambda(\log n)^{1/2})$ with high probability. Perhaps, one could now hope to round these y_i's to ± 1 without increasing the discrepancy substantially.

However, this possibility is ruled out by the hardness result in theorem 1. In particular, this result implies that there must exist systems with discrepancy $\Omega(\sqrt{n})$ but vector-discrepancy 0 (if such systems did not exist, then solving the discrepancy SDP with $\lambda = 0$ would give an algoirthm to distinguish between discrepancy 0 and $\Omega(\sqrt{n})$).

So, we adopt a different approach. Instead of trying to round the y_i's directly into ± 1, we will obtain a ± 1 solution gradually by combining several different collections of correlated $y_i's$ and solving several SDPs. This is also where we will really use that the guarantee is theorem 7 is with respect to the *hereditary* discrepancy.

Algorithm Overview: As mentioned above, instead of trying to obtain a coloring using a single SDP solution, we will gradually produce a solution by using several SDPs over time. At time 0, we start with the "empty" coloring $x^0 = (0,\ldots,0)$ where each element is colored 0. We slowly modify it over time as follows: Suppose x^{t-1}

denotes the coloring of elements at time $t-1$, we obtain the coloring x^t by adding a small perturbation vector u^t (how this is chosen will be described later) to x^{t-1}, i.e. $x^t(i) = x^{t-1}(i) + u^t(i)$ for each element i. As the perturbations are added, the color of the elements will evolve over time. Whenever an element's color reaches -1 or $+1$, we freeze that element's color and it is not longer updated.

It remains to specify how to generate the updates u^t. This is done using the gaussian rounding idea described above. In particular, at time t, we consider the SDP given by (2)-(3) with $\lambda = \text{herdisc}(C)$ (but only restricted to elements i that are still *alive*, i.e. not frozen yet). As $\lambda = \text{herdisc}(C)$, no matter which variables are alive at time t, the SDP is always feasible. We take the vectors v_i^t corresponding to some feasible SDP solution and set $u_i^t = \gamma\langle g, v_i^t\rangle$, where g is a random gaussian in $\mathbb{R}^n$ and γ is some polynomially small scaling factor ($\gamma = 1/n^c$ for any $c \geq 1$ suffices).

3.1 The Algorithm

We now state the algorithm formally.

1. Let x^t denote the coloring at time t. Let $\gamma = 1/n$ and $\ell = 8\log n/\gamma^2$. We initialize, $x^0(i) = 0$ for all $i \in [n]$. The F^t denote the set of frozen variables by time t, where we initialize $F^0 = \emptyset$.
2. For each time step $t = 1, 2, \ldots, \ell$ repeat the following steps:
   a. Find a feasible solution to the following semidefinite program:
   $$\begin{aligned} \|\sum_{i \in S_j} v_i\|_2^2 &\leq \lambda^2 \qquad \text{for each set } S_j \\ \|v_i\|_2^2 &= 1 \qquad \forall i \notin F^{t-1} \\ \|v_i\|_2^2 &= 0 \qquad \forall i \in F^{t-1} \end{aligned}$$
   b. Pick a random gaussian vector $g^t \in \mathbb{R}^n$.
   c. For each $i \in [n]$, update $x^t(i) = x^{t-1}(i) + \gamma\langle g^t, v_i^t\rangle$.
   d. Set $F^t = F^{t-1}$. For each i, freeze i if $|x^t(i)| > 1$, and update $F^t = F^t \cup \{i\}$.
3. After time $t = \ell$, if some $|x^\ell(i)| < 1$ (i.e. some element is alive), return fail.
4. For each i, set $x_i^\ell = -1$ if $x_i^\ell < -1$, and $x_i^\ell = 1$ if $x_i^\ell > 1$. Output the coloring x_ℓ.

Remark: The SDP in step 2(a) changes only when the set of frozen variables F changes. Moreover, the SDP is always feasible irrespective of the set F^t as $\text{herdisc}(C) \leq \lambda$.

3.2 Analysis

We will prove that the algorithm above produces a coloring with discrepancy $O((\log m \log n)^{1/2}\lambda)$ with probability at least $1/2$, where λ is the hereditary discrepancy of A. The proof relies on two simple ideas, that we first describe informally.

The proof sketch: First we show that all elements are frozen by time ℓ with high probability. Let us consider some element i. Observe that its color $x^t(i)$ starts at 0, and evolves over time until it crosses ± 1 and is frozen. At each step the update $u^t(i) = \gamma\langle v_i^t, g^t\rangle$ is added. As $\|v_i^t\| = 1$, by lemma 3, $u_i^t \sim N(0,\gamma^2)$, and as g^t is chosen independently at each time step, $u^t(i)$ is independent of the previous updates $u^{t-1}(i), u^{t-2}(i), \ldots$ for $x(i)$ (this is not strictly true, but let us ignore this technicality here). As the increments are $N(0,\gamma^2)$, $x^t(i)$ will reach ± 1 in $O(1/\gamma^2)$ steps with constant probability, and thus the probability that it does not reach ± 1 until $\ell = O(\log n/\gamma^2)$ steps is at most $1/n^2$. So, with probability at least $1 - 1/n$, all elements will be frozen by time ℓ.

We show now that the discrepancy is bounded with high probability. Let us consider how the discrepancy $x^t(S_j) = \sum_{i\in S_j} x^t(i)$ of a set S_j evolves over time. It is 0 initially at $t = 0$, at each step t, it is updated by $u^t(S_j) = \sum_{i\in S_j} \gamma\langle g^t, v_i^t\rangle$. Let $\lambda_t := \|\sum_{i\in S_j} v_j^t\|_2$. The SDP constraint ensures that $\lambda_t \leq \lambda$. Thus (roughly speaking) $x^t(S_j)$ evolves as a random walk with steps $N(0, \leq \gamma\lambda)$. By standard tail bounds, $\Pr[x^\ell(S_j) > c(\log m)^{1/2} \cdot \sqrt{\ell} \cdot \gamma\lambda] \leq 1/m^2$ for some suitable constant c. By union bound over the sets implies that

$$\mathrm{disc}(C) = O((\log m)^{1/2} \cdot \sqrt{\ell} \cdot \gamma\lambda) = O(\lambda \cdot (\log m \log n)^{1/2}).$$

Finally, note that truncating $x^\ell(i)$ to ± 1 in the last step introduces very low error. As $|x_i^t| < 1$ holds just before it is frozen and the next increment is a $N(0,\gamma^2)$ update, it must hold that $|x_i^t| < 1 + \gamma \cdot O((\log n)^{1/2})$ with high probability when it freezes. Thus, truncation adds at most $n \cdot \gamma \cdot O((\log n)^{1/2}) = O((\log n)^{1/2})$ error to any set.

The formal proof:

Recall that a sequence of random variables $X_0, \ldots, X_t$ forms a martingale, if $\mathbb{E}[X_t | X_{t-1}, \ldots, X_0] = X_{t-1}$. We first need the following tail bound on martingales with gaussian increments.

Lemma 4. *Let $0 = X_0 = X_1, \ldots, X_n$ be a martingale with increments $Y_i = X_i - X_{i-1}$. Suppose for $1 \leq i \leq n$, we have that $Y_i|(X_{i-1}, \ldots, X_0)$ is distributed as $\eta_i G$, where G is a standard gaussian $N(0,1)$ and η_i is a constant such that $|\eta_i| \leq 1$ (note that η_i may depend on $X_0, \ldots, X_{i-1}$). Then,*

$$\Pr[|X_n| \geq \lambda\sqrt{n}] \leq 2e^{-\lambda^2/2}.$$

Proof. Let α be a parameter to be optimized later. We have,

$$\begin{aligned}\mathbb{E}[e^{\alpha Y_i}|X_{i-1},\ldots,X_0] &\leq \int_{-\infty}^{\infty} e^{\alpha y}\cdot\left(\frac{1}{(2\pi)^{1/2}\eta_i}e^{-y^2/2\eta_i^2}\right)dy\\ &= e^{\alpha^2\eta_i^2/2}\cdot\int_{-\infty}^{\infty}\left(\frac{1}{(2\pi)^{1/2}\eta_i}e^{-(y-\alpha\eta_i^2)^2/2\eta_i^2}\right)dy\\ &= e^{\alpha^2\eta_i^2/2}\leq e^{\alpha^2/2}.\end{aligned}$$

Now,

$$\mathbb{E}[e^{\alpha X_n}] = \mathbb{E}[e^{\alpha X_{n-1}}e^{\alpha Y_n}] = \mathbb{E}[e^{\alpha X_{n-1}}\mathbb{E}[e^{\alpha Y_n}|X_{n-1},\ldots,X_0]] \leq e^{\alpha^2/2}\mathbb{E}[e^{\alpha X_{n-1}}].$$

Thus it follows by induction that $\mathbb{E}[e^{\alpha X_n}]\leq e^{\alpha^2 n/2}$. Finally by Markov's inequality,

$$\Pr[X_n\geq\lambda\sqrt{n}] = \Pr[e^{\alpha X_n}\geq e^{\alpha\lambda\sqrt{n}}]\leq e^{-\alpha\lambda\sqrt{n}}\mathbb{E}[e^{\alpha X_n}]\leq e^{-\alpha\lambda\sqrt{n}+\alpha^2 n/2}.$$

Setting $\alpha=\lambda/\sqrt{n}$ and noting that $\Pr[X_n\geq\lambda\sqrt{n}] = \Pr[X_n\leq-\lambda\sqrt{n}]$ implies the claim. □

Lemma 5. *Let $X_t = g_1+\ldots+g_t$ where g_i are iid $N(0,1)$ random variables. Then there is some universal constant $c>0$ such that for any integer $k\geq 1$,*

$$\Pr[|X_t|<\sqrt{n}\ \text{for}\ t=1,\ldots,k\sqrt{n}]<(1-c)^{-k+1}.$$

While more precise estimates known for the above quantity, the following simple proof suffices for our purposes.

Proof. By additivity of gaussians, X_n is distributed as $N(0,n)$. Thus $\Pr[|X_n|>2\sqrt{n}]\geq c$ for some constant $c>0$. Now, if $|X_t|<\sqrt{n}$ holds for each $t=1,\ldots,k\sqrt{n}$, it must necessarily hold that $|X_n|<\sqrt{n}$ and that $|X_{jn}-X_{(j-1)n}|<2\sqrt{n}$ for each $j=2,\ldots,k$. As each of these events is independent and the probability of each latter event is at most $1-c$. This implies the claimed result. □

We can now prove theorem 7.

Lemma 6. *With probability at least $1-1/n$, all elements will be frozen by time $\ell = O(\log n/\gamma^2)$.*

Proof. Let us consider an element i. We bound the probability that its color $x^t(i)$ never crosses ±1 until time ℓ. Starting from 0, at each step t, the update $u^t(i) = \gamma\langle v_i^t, g^t\rangle$ is added to $x^{t-1}(i)$. Now, while the vector v_i^t may depend on previous choices of the gaussian vectors $g^{t'}$ for $t'<t$ (as they determine the SDP at time t), note that u_i^t is distributed as $N(0,\gamma^2)$ and is independent of $u_i^{t'}$ for $t'<t$, whenever $\|v_i^t\|=1$. By lemma 5 there is some constant c' such that the probability that x_i does not each ±1 by $c'\log n/\gamma^2$ steps is at most $1/n^2$. The result follows by a union bound over the n elements.

Lemma 7. *With probability at least $1-1/m$, for each set S, it holds that $disc(S) = O(\lambda\cdot(\log m\log n)^{1/2})$.*

Proof. The discrepancy of a set S_j at time t is $x^t(S_j) = \sum_{i \in S_j} x^t(i)$. It is 0 at $t = 0$, and at each step t it gets updated by $u^t(S_j) = \sum_{i \in S_j} \gamma \langle g^t, v_i^t \rangle$. Let $\lambda_t := \| \sum_{i \in S_j} v_j^t \|_2$. The SDP constraint ensures that $\lambda_t \leq \lambda$. Now, λ_t may depend on the previous choices $g^{t'}$ for $t' \leq t-1$ (as these choices affect the SDP at time t), but as g^t is chosen independently at time t, conditioned on the previous random choices $g^{t'}$, the update $u^t(S_j) \sim N(0, \gamma^2 \lambda_t^2)$. Thus, $x^t(S_j)$ forms a martingale with increments $N(0, \leq \gamma^2 \lambda^2)$. So by lemma 4,

$$\Pr[x^\ell(S_j) > c(\log m)^{1/2} \cdot \sqrt{\ell} \cdot \gamma\lambda] \leq 1/m^2$$

for some suitable constant c. By a union bound over the m sets, $\mathrm{disc}(S) = O((\log m)^{1/2} \cdot \sqrt{\ell} \cdot \gamma\lambda) = O(\lambda \cdot (\log m \log n)^{1/2})$ with probability at least $1 - 1/m$. □

Finally, as discussed previously, truncating a frozen variable $x^\ell(i)$ to ± 1 introduces only negligible error. Combining lemma 6 and 7 gives theorem 7.

3.3 Bounds based on Partial Hereditary Discrepancy

Let us define the *partial hereditary discrepancy* of a system as the smallest number λ such that for any sub-system, there exists a partial coloring (say that colors at least half the elements of that sub-system) with discrepancy at most λ. The result above can be refined to show that

Theorem 11. *There is a polynomial time algorithm that finds a coloring with discrepancy $O(\lambda (\log m \log n)^{1/2})$, where λ is the partial hereditary discrepancy of A.*

This is useful because for many problems better bounds are known on partial hereditary discrepancy than for hereditary discrepancy. For example, for bounded degree systems the best bound we know on hereditary discrepancy is $O((t \log n)^{1/2})$ [2], while the partial hereditary discrepancy is $O(\sqrt{t})$ (as we saw in section 2.2).

The algorithm is a direct modification of the one in section 3.1. We replace the SDP constraint 3 that $\|v_i\|^2 = 1$, by the conditions $\sum_i \|v_i\|^2 \geq n/2$ and $\|v_i\|^2 \leq 1$ for each i. In particular, we consider the following SDP.

$$\| \sum_{i \in S_j} v_i \|_2^2 \leq \lambda^2 \qquad \text{for each set } S_j \tag{4}$$

$$\sum_{i \notin F} \|v_i\|_2^2 \geq |F^c|/2 \tag{5}$$

$$\|v_i\|_2^2 \leq 1 \qquad \forall i \in F^c \tag{6}$$

$$\|v_i\|_2^2 = 0 \qquad \forall i \in F \tag{7}$$

Here F^c denotes the complement of F (i.e. alive variables) and λ is the partial hereditary discrepancy. Note that the constraints (5) and (6) only require that at least half of the alive variables must be colored.

Analysis: While the algorithm is same as before, the analysis needs some more care. The problem is that the alive variables do not necessarily satisfy $\|v_i\|_2 = 1$, but only the weaker condition (5). So, a priori it is possible that some variable always has $\|v_i\| \approx 0$ and hence never makes progress towards reaching ± 1. To get around this, one needs a more careful "energy increment" argument to show that after every $1/\gamma^2$ time steps, a constant fraction of the variables do reach ± 1 in expectation. One can then show that all elements are eventually colored ± 1 in $O(\log n/\gamma^2)$ time steps with probability at least $1/n^{O(1)}$.

The key result is the following.

Theorem 12. *Let $x \in [-1,+1]^n$ be an arbitrary fractional coloring with at most k alive variables. Starting from the coloring x, let z be the coloring obtained after applying the steps (2a)-(2d) of the algorithm for the SDP given by* (4)-(7)*, for $16/\gamma^2$ time units. Then the probability that z has more than $k/2$ alive variables is at most $1/4$.*

Proof. Let $u = 16/\gamma^2$ and for each $1 \le t \le u$, let x_t denote the coloring at the end of time t starting from the initial coloring $x_0 = x$, i.e. after t applications of steps (2a)-(2d). Let K denote the set of alive variables at time $t = 0$. Let k_t denote the number of variables alive at the end of time t, and let $k = k_0 = |K|$. We would like to show that $k_u \le k_0/2$ with probability at least $3/4$. To do this, we track how the "energy" $\sum_i x_t(i)^2$ of the coloring x^t evolves as the algorithm proceeds.

For each time $t = 1, \ldots, u$, let us define

$$r_t = \begin{cases} \sum_{i \in K} x_t(i)^2 & \text{if } k_{t-1} \ge k/2, \\ r_{t-1} + \gamma^2 k/4 & \text{otherwise.} \end{cases}$$

Lemma 8. *Conditioned on any coloring x_{t-1} at the end of time $t-1$, the expected increment $r_t - r_{t-1}$ at time step t is at least $\gamma^2 k/4$, where the expectation is over the choice of the random gaussian $g \in R^n$ at time t. That is, $\mathbb{E}[r_t - r_{t-1} | x_{t-1}] \ge \gamma^2 k/4$ for any x_{t-1}.*

Proof. If $k_{t-1} < k/2$, this follows trivially from the definition of r_t, as r_t is deterministically set to $r_{t-1} + \gamma^2 k/4$. So it suffices to consider the case when $k_{t-1} \ge k/2$. Let v^t the vector solution to the SDP (4)-(7). Then,

$$\begin{aligned} \mathbb{E}[r_t - r_{t-1} | x_{t-1}] &= \mathbb{E}[r_t | x_{t-1}] - r_{t-1} \\ &= \mathbb{E}_g \left[\sum_i (x_{t-1}(i) + \gamma \langle g, v_i^t \rangle)^2 \right] - \sum_i x_{t-1}(i)^2 \\ &= \sum_i \left(2\gamma x_{t-1} \mathbb{E}[\langle g, v_i^t \rangle] + \gamma^2 \mathbb{E}[(\langle g, v_i^t \rangle)^2] \right) \ge \gamma^2 k_{t-1} \ge \gamma^2 k/4. \end{aligned}$$

The first step follows as r_{t-1} is completely determined by x_{t-1}. The last step follows for the following reason. By lemma 3, $\langle g, v_i^t \rangle$ is distributed as $N(0, \|v_i^t\|^2)$ which implies that $\mathbb{E}_g[\langle g, v_i^t \rangle] = 0$ and $\mathbb{E}_g[(\langle g, v_i^t \rangle)^2] = \|v_i^t\|^2$. Now, the SDP constraint (5) ensures that $\sum_i \|v_i^t\|^2 \ge k_{t-1}/2$ which is at least $k/4$. $\square$

We now use this lemma together with Markov's inequality to finish off the proof.

The crucial observation is the following. Consider time $t = u$. If $k_u \geq k/2$, then by definition $r_u = \sum_{i \in K} x_u(i)^2 \leq k$. On the other hand, it always holds that $r_u \leq k + u\gamma^2 k/4$. This is because in any run of the algorithm, $r_t = \sum_{i \in K} x_t(i)^2 \leq k$ as long as $k_t \geq k/2$, and when k_t falls below $k/2$, then r_t increases deterministically by $\gamma^2 k/4$ at each subsequent time step.

This gives us that

$$\begin{aligned}\mathbb{E}[r_u] &= \mathbb{E}[r_u | k_u \geq k/2]\Pr[k_u \geq k/2] + \mathbb{E}[r_u | k_u < k/2]\Pr[k_u < k/2] \\ &\leq k \cdot \Pr[k_u \geq k/2] + (k + \gamma^2 uk/4) \cdot (1 - \Pr[k_u \geq k/2]) \\ &= k + (1 - \Pr[k_u \geq k/2]) \cdot (\gamma^2 uk/4).\end{aligned}$$

As $\mathbb{E}[r_u] \geq \gamma^2 uk/4$ by lemma 8, together this gives that $\gamma^2 uk/4 \leq k + (1 - \Pr[k_u \geq k/2]) \cdot (\gamma^2 uk/4)$, and hence that $\Pr[k_u \geq k/2] \leq k/(\gamma^2 uk/4) = 1/4$ as claimed. □

Lemma 9. *Let $\ell = 16 \log n/\gamma^2$. The probability that every element is colored ± 1 by time ℓ, is at least $1/n$.*

Proof. We apply theorem 12 repeatedly with the starting coloring $x = x_t$ at the time steps $t = 0, 16/s^2, 32/s^2, \ldots, (16 \log n)/s^2 = \ell$. With probability at least $(1 - 1/4)^{\log n} \geq 1/n$, the number of alive variables at least halves at each epoch and hence reaches 0. □

The proof of theorem 11 now follows directly by combining lemma 9 together with the argument in lemma 7.

4 Algorithmic version of Spencer's Result

In this section we consider theorem 6. This result turns out to be much more tricky that the algorithm in the previous section.

To keep the focus on the main ideas, we will first describe a weaker version of theorem 6, that gives an $O((n \log\log\log n)^{1/2})$ discrepancy coloring (note that is still much better than randomized rounding). Later, in section 5 we will describe the recent and much simpler algorithm due to Lovett and Meka [19] to find an $O(\sqrt{n})$ discrepancy coloring.

Some problematic issues: The reason that proving theorem 6 is much more tricky is the following. First, it is not at all clear whether semidefinite programming is useful. In particular, consider the SDP given by (2)-(3). The natural thing is to set $\lambda = O(\sqrt{n})$, and try to use this SDP solution.[3] However, if we set $\lambda \geq \sqrt{n}$, then this SDP can always return the trivially feasible solution $v_i = e_i$ for $i \in [n]$, where e_i

[3] Note that one cannot set $\lambda = o(\sqrt{n})$ in our setting, there are set systems on $m = O(n)$ sets (e.g. the Hadamard set system, that we will see in section 6) with vector discrepancy $\Omega(\sqrt{n})$.

denotes the unit vector in the ith direction. This SDP solution is feasible as the v_i are unit vectors $\|v_i\| = 1$ and their orthogonality implies that $\|\sum_{i\in S} v_i\|_2 = (|S|)^{1/2} \leq n^{1/2}$ for any S. Thus, the SDP does not seems to reveal any useful information.

A second problem is that in the previous algorithm the discrepancy of a set performs a random walk over time. So, even if we the expected discrepancy of a set is $O(\sqrt{n})$, some of them are very likely to deviate from the expectation by factor $\Omega(\sqrt{\log n})$. So we do not seem to get anything better than the $O(\sqrt{n\log n})$ bound that random coloring would have given us anyway.

The additional idea: To get around these issues, the idea is to let the discrepancy bound λ_S for set S (in the SDP) vary over time depending on how the discrepancy of S evolves. If a set S gets dangerously closely to violating the target $c\sqrt{n}$ discrepancy bound, we set its λ_S (denoted by Δ_S in lemma 10) in the SDP to be much smaller than $\sqrt{n}$, thereby ensuring that its discrepancy is extremely unlikely to exceed in $c\sqrt{n}$ is subsequent iterations. The point is that the entropy method (lemma 10) guarantees a partial coloring provided the discrepancy bounds satisfy the condition (1). So if we can argue that not too many sets become dangerous, then the argument still goes through.

Before we describe the algorithm and its analysis, we state the following corollary of theorem 10 that we will need.

Corollary 2. *Let (V,C) be any set system on $m = O(n)$ sets, and $C' \subset C$ be any subcollection of $O(n/(\log\log n)^2)$ sets. Then there exists a partial coloring where the discrepancy of sets in C' is at most $\sqrt{n}/\log n$ and the ones in $C \setminus C'$ is $O(\sqrt{n})$.*

Proof. For each set $S \in C'$, setting $\lambda_S = \sqrt{n}/\log n$ contributes at most $g(1/\log n) = O(\log\log n)$ to the left hand side of (1). As $|C'| = O(n/(\log\log)^2 n)$, the overall contribution due to C' is $o(n)$. For the sets in $C \setminus C'$ we set $\lambda_S = c\sqrt{n}$ for c large enough, so that their total contribution to (1) is at most $n/10$. Thus the claimed partial coloring exists by lemma 10. □

4.1 Algorithmic Subroutine and Analysis

We now describe the algorithm. We only consider the first phase when the number of uncolored variables reduces from n to $n/2$. This is the hardest phase and contains all the main ideas.

Algorithm for the first phase: We start with the all 0 coloring, and consider the following partial coloring SDP.

$$\|\sum_{i\in S} v_i\|_2^2 \leq \lambda_S^2 \qquad \text{for each set } S = S_1,\ldots,S_m \tag{8}$$

$$\sum_{i\notin F} \|v_i\|_2^2 \geq |F^c|/2 \tag{9}$$

$$\|v_i\|_2^2 \leq 1 \qquad \forall i \notin F \tag{10}$$

$$\|v_i\|_2^2 = 0 \qquad \forall i \in F \tag{11}$$

Initially we set $\lambda_S = cn^{1/2}$ for each set S, where c is a large enough constant such that (1) is satisfied easily with some slack. As previously, for each time step t, we obtain the update $u_i^t = \gamma\langle g^t, v_i^t\rangle$ for $i = 1,\ldots,n$ and add it to the coloring thus far. We repeat this for $O(1/\gamma^2)$ steps, at which point we expect half the colors to reach ± 1.

During these steps, if the discrepancy $|x^t(S)|$ of set S ever exceeds $2(n\log\log\log n)^{1/2}$, we label S *dangerous* and set $\lambda_S = n^{1/2}/\log n$ in all the subsequent SDPs. This ensures that its discrepancy increment $u^t(S)$ will have standard deviation at most $O(\gamma\cdot(n^{1/2}/\log n))$ henceforth, making S extremely unlikely to incur an additional $\Omega(n^{1/2})$ discrepancy over the remaining $O(1/\gamma^2)$ steps.

Analysis: By design, the reduction of λ_S for dangerous sets ensures that after $O(1/\gamma^2)$ steps, the discrepancy of every set is $O((n\log\log\log n)^{1/2})$ with high probability. Moreover, lemma 9 implies that with probability at least $3/4$, at least $n/2$ elements are colored ± 1 by the end of the phase.

It suffices to show that with probability at least $3/4$, the SDP never becomes infeasible. Indeed, as the discrepancy of any set forms a martingale with gaussian increments with standard deviation $O(\sqrt{n})$, thus by lemma 4 the probability of a set ever becoming dangerous is $O(\exp(-2\log\log\log n)) = O(\log\log n)^{-2}$. Thus the total expected number of dangerous sets is $O(n/(\log\log n)^2))$, and by Markov's inequality, with probability at least $3/4$, this number does not exceed $O(n(\log\log n)^{-2})$. Corollary 2 now gives the claimed result.

Remarks:

1. The $O(n^{1/2})$ bound in theorem 6 follows by refining this idea by having multiple danger levels and by setting the bounds λ_S for a set S appropriately for each danger level.
2. Even though theorem 6 gives a polynomial time algorithm, it crucially uses the entropy method to argue about the feasibility of the SDP, and hence is not truly constructive.

Recently, Lovett and Meka [19] discovered a much simpler argument to obtain an algorithmic version of Spencer's result. Their idea was to combine the gaussian random walk approach above together with the linear algebraic ideas that we saw in section 2.1. We now describe their algorithm and analysis.

5 The result of Lovett and Meka

As usual, let (V,C) be a set system with n elements and m sets $S_1,\ldots,S_m$. Lovett and Meka [19] gave the following algorithmic version of the partial coloring lemma.

Theorem 13. *Let $x \in [-1,1]^n$ be some fractional coloring, with k alive elements (i.e. k elements i such that $x(i) \neq \pm 1$). For $j = 1,\ldots,m$, let λ_j be such that*

$$\sum_j \exp(-\lambda_j^2/16) \leq k/16. \tag{12}$$

Then there is a randomized polynomial time algorithm to find a coloring x' with at most $k/2$ alive variables such that the additional discrepancy added to each set S_j is at most $\Delta_{S_j} = \lambda_j\sqrt{|S_j|}$. That is, $|x'(S_j) - x(S_j)| \leq \Delta_{S_j}$ for each $j = 1,\ldots,m$.

Remark: The theorem is stated in the form above so that it can be applied repeatedly to obtain a complete coloring after $O(\log n)$ rounds. The reader may assume that the starting coloring x is $(0,\ldots,0)$.

An interesting aspect of theorem 13 is that it is quantitatively stronger than the entropy method (theorem 10). In particular, if we require $\lambda \ll 1$ for the sets, then using (1) this can only be done for $O(n/\log n)$ sets, while theorem 13 allows this for $\Omega(n)$ sets. For more discussion on the comparison with the entropy method we refer the reader to [19].

We now prove theorem 13. Let K denote the set of alive elements at $t = 0$. Without loss of generality let us assume that $K = \{1,\ldots,k\}$. As in the previous algorithms, the algorithm will start from the initial coloring $x_0 = x$ and update it over time as $x_t = x_{t-1} + u^t$ by adding a suitably chosen tiny update vector u^t at time t.

5.1 The Algorithm

Let $\gamma \leq 1/n^2$ be a tiny parameter as before, and let $\delta = (10\log n)\gamma$. Let us call a set S_j dangerous after time $t-1$, if $|x_{t-1}(S_j) - x_0(S_j)| \geq \lambda_j - \delta$. Also, we call a variable i frozen after time $t-1$ if $|x_{t-1}(i)| \geq 1-\delta$. The algorithm will ensure that the color $x_t(i)$ of any element i does not change once it freezes, and moreover the discrepancy of a set S also never changes once it becomes dangerous.

To achieve this the algorithm makes the update u^t at time t in an appropriate linear subspace defined as follows. Given the coloring x_{t-1} after time $t-1$, let $V^t \subseteq \mathbb{R}^k$ be the subspace of points $(v(1),\ldots,v(k))$ satisfying:

1. If the element i is frozen, then $v(i) = 0$.
2. If the set S_j is dangerous, then $\sum_{i \in S_j \cap [k]} v(i) = 0$.

Algorithm: Initialize $x_0 = x$ and $V^0 = \mathbb{R}^k$.
For $t = 1,\ldots,T$, where $T = 16/(3\gamma^2)$, repeat the following steps.

1. Pick $g \sim N(V^t)$. Update the coloring as $x_t = x_{t-1} + \gamma g$.
2. If some element i freezes or if some set S_j becomes dangerous at time t, update V^{t+1} accordingly.

Note that as the algorithm proceeds $\mathbb{R}^k = V^0 \supseteq V^1 \supseteq \cdots \supseteq V^t$. Moreover, as the update γg lies in V^t, it follows that for any element i, we have that $|x_t(i)|$ is at most $1 - \delta + O(\sqrt{\log(T)})\gamma \leq 1$ with high probability. Similarly, the additional discrepancy $|x_{t-1}(S) - x_0(S)|$ of any set does not exceed $\lambda_j - \delta + O(\gamma\sqrt{|S_j|}\sqrt{\log T}) \leq \lambda_j + 1/n$.

Thus one only needs to show that the algorithm does not get stuck (i.e. $V^t = \emptyset$, while more than half the variables are still alive).

5.2 Analysis

We begin with two simple properties of random gaussian vectors that play a key role in the analysis.

For a linear subspace $V \subseteq \mathbb{R}^n$, let $N(V)$ denote the standard multi-dimensional gaussian distribution supported on V. A random vector g is distributed according to $N(V)$ if $g = g(1)v_1 + \ldots + g(d)v_d$ where $\{v_1, \ldots, v_d\}$ is an orthonormal basis for V, and $g(1), \ldots, g(d)$ are independent $N(0,1)$ random variables. By the rotational invariance of the multi-dimensional gaussian distribution, it is easily seen that g is invariant of the choice of the basis $\{v_1, \ldots, v_d\}$.

Lemma 10. *If $g \sim N(V)$, then for all $u \in \mathbb{R}^n$, $\langle g, u\rangle \sim N(0, \sigma^2)$ where $\sigma^2 \leq \|u\|^2$.*

Proof. Let u' denote the projection of u onto V. Clearly, $\|u'\| \leq \|u\|$. As $g \in V$, $\langle g, u\rangle = \langle g, u'\rangle$, which by lemma 3 is distributed as $N(0, \|u'\|^2)$. □

Let e_i denote the unit vector in the i-th direction.

Lemma 11. *Let V be a d-dimensional subspace of $\mathbb{R}^n$ and $g \sim N(V)$. For $i = 1, \ldots, n$ let σ_i be such that $\langle g, e_i\rangle \sim N(0, \sigma_i^2)$. Then $\sum_{i=1}^n \sigma_i^2 = d$.*

Proof. If $v_1, \ldots, v_d$ is an orthogonal basis for V, then by lemma 3, $\sigma_i^2 = \sum_{j=1}^d \langle e_i, v_j\rangle^2$. Thus $\sum_{i=1}^n \sigma_i^2 = \sum_{i=1}^n \sum_{j=1}^d \langle e_i, v_j\rangle^2 = \sum_{j=1}^d \sum_{i=1}^n \langle v_j, e_i\rangle^2 = \sum_{j=1}^d \|v_j\|^2 = d$, where the second last equality follows as $\{e_1, \ldots, e_n\}$ is an orthogonal basis for $\mathbb{R}^n$ and hence $\|v\|^2 = \sum_{i=1}^n \langle v, e_i\rangle^2$ for every v. □

The proof of theorem now follows by arguments similar to the ones in section 3.3. We sketch the main idea here and refer the reader to [19] for the detailed computations.

First we claim that not many sets become dangerous in expectation. This follows as at each time step t, lemma 10 implies that irrespective of the choice of V^t, the discrepancy increment of each set S_j is distributed as $N(0, \leq \gamma^2|S_j|)$. Thus, by lemma 4,

$$\Pr[|x_T(S_j) - x_0(S_j)| \geq \Delta_{S_j}] \leq 2\exp(-\Delta_{S_j}^2/(T\gamma^2|S_j|)) = 2\exp(-3\lambda_j^2/16). \quad (13)$$

As λ_j satisfy (12), by a standard calculation that we skip, (13) implies that the probability that more than $k/8$ sets become dangerous is at most $1/8$.

Let us condition on the event that no more than $k/8$ sets become dangerous. Then we claim that with probability at least $3/4$, at least $k/2$ elements become frozen. This follows by the argument in lemma 12. In particular, if fewer than $k/2$ elements are frozen, then at any time during the algorithm the dimension of the subspace V^t is at least $k - k/2 - k/8 \geq 3k/8$. Now, lemma 11 implies that the expected energy increment $\sum_i x_t(i)^2 - x_{t-1}(i)^2 \geq (3k/8)\gamma^2$. So, if fewer than $k/2$ elements are frozen, the total energy increment will be $T(3k/8)\gamma^2 = 2k$. But as $x_t(i)^2$ can never exceed k, we can use the argument in the proof of lemma 12 (with constants modified) to upper bound the probability that fewer than $k/2$ elements are fixed. Together, these two claims imply the result.

6 Inapproximability of Discrepancy

In this section we prove theorem 1, which shows that discrepancy is essentially hopeless to approximate. The proof has two main ingredients. One starts with a very weak hardness result for discrepancy, which states that it is NP-hard to determine if a set system has discrepancy 0 or if a constant fraction of sets have discrepancy at least 1. The next step is to amplify this hardness by composing it with the Hadamard set system. The Hadamard system is a classic example that shows that theorem 2 is tight up to constant factors.

Hadamard Set System: We recall the construction of the $2^k \times 2^k$ Hadamard matrix $H^{(k)}$, which is defined as the k-fold tensor product $H^{(k)} := H^{(1)} \otimes \cdots \otimes H^{(1)}$, where $H^{(1)}$ is the 2×2 matrix with entries $h(1,1) = h(1,2) = h(2,1) = 1$ and $h(2,2) = -1$.

Let us denote $n = 2^k$ and $H = H^{(k)}$. The only property we need of the H is that (i) it is symmetric and consists of ± 1 entries, (ii) its first row consists entirely of 1's and (iii) its columns are mutually orthogonal, i.e. $H^T H = nI$.

Let J denote the $n \times n$ matrix with all 1's, and let Z be the 0-1 matrix $Z := \frac{1}{2}(H + J)$. Then Z satisfies the following.

Lemma 12. *Let $x \in \mathbb{R}^n$ be such that $\sum_{i>1} x_i^2 = \Omega(n)$. Then $\|Zx\|_2^2 = \Omega(n^2)$.*

Proof. Let Z_i (resp. J_i, H_i) denote the i-th column of Z (resp. J, H). We have

$$\|Zx\|_2^2 = (Zx)^T(Zx) = \sum_{i,j} x_i Z_i^T Z_j x_j = \frac{1}{4}\sum_{i,j} x_i (H_i + J_i)^T (H_j + J_j) x_j.$$

As the columns of H are orthogonal, $\sum_{i,j} x_i x_j H_i^T H_j = \sum_i x_i^2 H_i^T H_i = n\|x\|_2^2$. Moreover, as $J_j = H_1$ for all j, $\sum_{i,j} x_i x_j H_i^T J_j = \sum_{i,j} x_i x_j H_i^T H_1 = (\sum_j x_j)(nx_1)$. Similarly,

$\sum_{i,j} x_i x_j J_i^T H_j = (\sum_i x_i)(nx_1)$. Finally, the last term $\sum_{i,j} x_i x_j J_i^T J_j = n(\sum_i x_i)^2$. Combining these terms gives

$$\|Zx\|_2^2 = n(\|x\|_2^2 + 2x_1(\sum_i x_i) + (\sum_i x_i)^2) = n(\|x\|_2^2 + (x_1 + \sum_i x_i)^2 - x_1^2) \geq n(\sum_{i>1} x_i^2).$$

□

Max-2-2-Set-Splitting Problem: We will use the following max-2-2-set-splitting problem (henceforth referred to as MSS). An instance of MSS consists of m sets, $C = \{C_1, \ldots, C_m\}$ on n elements, each set consisting of exactly four distinct elements. The objective is to assign each element either $\{-1, +1\}$ such that number of sets for which the values sum to exactly 0 is maximized. Clearly, for any assignment, the sets can have values $\{0, \pm 2, \pm 4\}$. The sets that have value 0 are called *split*, and otherwise they are *unsplit*.

Theorem 14 ([15]). *There exists a universal constant $\phi > 0$ such that it is NP-hard to distinguish between instances of MSS whether (i) there is an assignment such that all the sets are split, or (ii) any assignment will result in at least ϕm unsplit sets.*

Moreover, one can assume that in these instances no element appears in more than b sets for some universal constant b.

If C denotes the $m \times n$ incidence matrix of an MSS instance, the result above implies that it is NP-hard to distinguish whether (i) the discrepancy of C is 0 (ii) for any $y \in \{-1, 1\}^n$, at least $\Omega(m)$ sets have non-zero discrepancy. We amplify this gap by composing the MSS instance with the Hadamard system. At first, we obtain a matrix whose entries are not necessarily $\{0, 1\}$, but lie in the range $[0, b]$ for some constant b. Later, we show how to modify the argument to obtain a $\{0, 1\}$ matrix.

Theorem 15. *Given a $m \times n$ matrix B with $m = O(n)$ and integer entries in the range $[0, b]$, where b is a constant, it is NP-hard to distinguish between the cases: (i) the discrepancy of B is 0 (ii) for any $y \in \{-1, 1\}^n$, $\|By\|_2^2 = \Omega(n^2)$.*

Proof. Let us take an instance of MSS on n elements and $m = O(n)$ sets, and let C be its incidence matrix. By duplicating some sets if necessary, we can assume without loss of generality that m is an integer power of 2. Consider the matrix $B = ZC$, where Z is the $m \times m$ set system in lemma 12. As each column of C contains at most b 1's and Z is a $0-1$ matrix, each entry of B is an integer in $[0, b]$.

Now, if the MSS instance corresponding to C has an assignment y that splits each set, then $Cy = \mathbf{0}$ and hence $By = Z(Cy) = \mathbf{0}$. On the other hand if $\|Cy\|_2^2 = \Omega(m)$, then by lemma 12 and noting that the first entry of $Cy \in \{0, \pm 2, \pm 4\}$, it follows that $\|By\|_2^2 = \|Z(CY)\|_2^2 = \Omega(m^2) = \Omega(n^2)$. □

We now modify the argument so that B only has $\{0, 1\}$ entries. Let us partition the sets in the MSS instance into $h \leq 4b + 1$ parts $T^1, \ldots, T^h$ such that an element j appears in at most one set from any collection T^i. Such a partition exists as each set contains four elements and each elements lies in at most b sets, and hence each

set shares an element with at most $4b$ other sets. Duplicating sets if necessary, we assume that the number of sets m_i in T^i is an integer power of 2. For each $i = 1, \ldots, h$, let H^i denote the $m_i \times m_i$ Hadamard set system, and let $B^i = H^i T^i$. Note that B^i has $\{0,1\}$ entries as each element appears in at most one set in T^i. Let B be the $(\sum_i m_i) \times n$ matrix obtained by placing the rows of $B^1, \ldots, B^h$ one after the other.

Now if $Cy = 0$ then $T^i y = \mathbf{0}$ for each i, and hence each $B^i y = H^i T^i y = \mathbf{0}$ which implies that $By = \mathbf{0}$. On the other hand, if $\|Cy\|_2^2 = \Omega(m)$, then $\|T^i y\|_2^2 = \Omega(m/b) = \Omega(m)$ for some i, and hence $\|By\|_2^2 \geq \|B^i y\|_2^2 = \|H^i(T^i y)\|_2^2 = \Omega(m^2)$, where the last step follows from lemma 12.

7 Tightness of the Determinant Lower Bound

In this section, we see how the connection between discrepancy and semidefinite programming can be used to show that the determinant lower bound characterizes hereditary discrepancy up to poly-logarithmic factors. We begin by describing the determinant lower bound and proving theorem 5. Then in section 7.2, we prove theorem 8.

7.1 Determinant Lower Bound

Recall that for a real matrix A, we denote $\mathrm{detlb}(A) := \max_k \max_B |\det B|^{1/k}$, where the maximum is over all $k \times k$ submatrices B of A. We will show that

Theorem 5. *For any real matrix A,* $\mathrm{herdisc}(A) \geq \frac{1}{2}\mathrm{detlb}(A)$.

Let us define the linear discrepancy of a matrix A as

$$\mathrm{lindisc}(A) := \max_{x \in [-1,1]^n} \min_{y \in \{-1,1\}^n} \|Ax - Ay\|_\infty.$$

That is, it is the worst case error over all points $x \in [-1,1]^n$, when x is rounded to the "best" integral $y \in \{-1,1\}^n$. Theorem 3 directly implies that

$$\mathrm{lindisc}(A) \leq 2\,\mathrm{herdisc}(A) \tag{14}$$

(we lose the extra factor 2 here as $y \in \{-1,1\}^n$, instead of $y \in \{0,1\}^n$).

Lemma 13. *To prove theorem 5, it suffices to show that for any $k \times k$ matrix B,* $\mathrm{lindisc}(B) \geq \det(B)^{1/k}$.

Proof. Let A be any $m \times n$ matrix. By the definition of hereditary discrepancy and (14)

$$\mathrm{herdisc}(A) \geq \max_k \max_B \mathrm{herdisc}(B) \geq \frac{1}{2} \max_k \max_B \mathrm{lindisc}(B),$$

where B ranges over the $k \times k$ submatrices of A. By the claim, this is at least $\frac{1}{2}\max_k \max_B \det(B)^{1/k}$ which is exactly $\frac{1}{2}\text{detlb}(A)$. $\square$

Thus we focus on proving lemma 13 for square matrices. The proof is based on a geometric interpretation of lindisc.

Geometric Interpretation of Linear Discrepancy: Let B be an invertible $k \times k$ matrix (if B is non-invertible, then $\det(B) = 0$ and lemma 13 holds trivially). Let $P \in \mathbb{R}^k$ be the set of points x satisfying $-\mathbf{1} \leq Bx \leq \mathbf{1}$, where $\mathbf{1}$ is the all 1's vector in $\mathbb{R}^k$. Clearly, if $x \in P$ then $-x$ also lies in P and hence P is symmetric about the origin. Note that P is the inverse image of the unit cube $[-1,1]^k$ under B^{-1}, i.e. $P = \{B^{-1}z : z \in [-1,1]^k\}$. We claim the following.

Lemma 14. *Linear discrepancy is the smallest real number t such that placing a copy of tP at every point $\{-1,1\}^k$ covers the cube $[-1,1]^k$ completely.*

Proof. By definition of linear discrepancy and P, $\text{lindisc}(B)$ is smallest t such that for each point $x \in [-1,1]^k$, the polytope $tP + x$ (i.e. P scaled t times and shifted by x) contains some point y in $\{-1,1\}^k$. Let x be any point in $[-1,1]^k$. Now if $y \in \{-1,1\}^k$ is such that $y \in tP + x$, then $y - x \in tP$. But then, $x - y \in tP$ by the symmetry of P which implies that $x \in tP + y$. $\square$

We now finish the proof of theorem 5. For each $y \in \{-1,1\}^k$, let $R_y = [-1,1]^k \cap (tP + y)$ denote the intersection of the unit cube with the copy of tP at y. Observe that R_y is identical to $tP \cap I(-y)$ where $I(y)$ is the orthant $\Pi_i[0, y_i]$ formed by the origin and y. As $tP = \cup_{y=\{-1,1\}^k}(tP \cap I(y))$, this implies that R_y's give a partitioning of tP and hence

$$\text{Vol}(tP) = t^k \text{Vol}(P) = \sum_{y \in \{-1,1\}^k} \text{Vol}(R_y) \geq \text{Vol}([-1,1]^k) = 2^k \tag{15}$$

where the inequality above follows by lemma 14 and the definition of R_y.

As $P = \{B^{-1}z : z \in [-1,1]^k\}$ we have that

$$\text{Vol}(P) = \det(B^{-1})\text{Vol}([-1,1]^k) = 2^k \det(B^{-1}) = 2^k / \det(B). \tag{16}$$

Taking logarithms in (15) and using (16) this gives $t \geq 2/\text{Vol}(P)^{1/k} = \det(B)^{1/k}$, which implies lemma 13 and hence theorem 8.

7.2 Matoušek's Upper Bound

A natural question is whether the determinant lower bound essentially determines the hereditary discrepancy of a system. Hoffman gave an elegant example (see [21]) of a set system with $\text{detlb}(C) = O(1)$ and $\text{disc}(C) \approx (\log n)/(\log\log n)$, and hence implies that the gap between the two can be $\Omega(\log n/\log\log n)$. Recently, Pálvölgyi

[23] gave an improved example with an $\Omega(\log n)$ gap. However, there was no non-trivial result upper bound known on this gap until the following recent result [20] of Matoušek.

Theorem 8.[20] *For any set system* (V,C), $\text{herdisc}(C) \leq O(\log n\sqrt{\log m}) \cdot \text{detlb}(C)$.

The proof uses several interesting ideas, and in particular the duality for semidefinite programming.

SDP Dual for Discrepancy: Let us recall the strong duality for convex programs which states the following. Let P be a convex programming problem (say, with minimization objective). Then any feasible solution to the dual program of P is a lower bound on the optimum solution for P. Moreover under some mild technical conditions (see [29] for details), the value of the optimum feasible dual solution is equal to the value of the optimum feasible solution for P. Let us apply this to the following discrepancy minimization SDP for the system (V,C) (having the optimum value $\text{vecdisc}(C)$).

$$\min\ \lambda^2,\ \text{s.t.}\ (i) \quad \|\sum_{i\in S_j} v_i\|_2^2 \leq \lambda^2 \quad \forall S_j \in C,\ \text{and}\ (ii) \quad \|v_i\|_2^2 = 1, \quad \forall i \in V.$$

Lemma 15. *For any set system* (V,C) *with n elements and m sets, we have that* $\text{vecdisc}(C) \geq D$ *if and only if there are nonnegative reals* $w_1,\ldots,w_m$ *with* $\sum_{j=1}^m w_j \leq 1$ *and reals* $z_1,\ldots,z_n$ *with* $\sum_{i=1}^n z_j \geq D^2$ *such that for all* $\mathbf{x} \in \mathbb{R}^n$,

$$\sum_{j=1}^m w_j \left(\sum_{i\in S_j} x_i\right)^2 \geq \sum_{i=1}^n z_i x_i^2. \tag{17}$$

While computing the dual formally needs some work (and we refer the reader to [20] for details), this dual has a rather intuitive interpretation. Suppose there exists a convex combination with coefficients w_i, of the set discrepancy constraints $(\sum_{i\in S_j} x_i)^2$ such that this sum always exceeds $\sum_{i=1}^n z_i x_i^2$ no matter what real values are assigned to x_i's. Or in other words, $\sum_{j=1}^m w_j(\sum_{i\in S_j} x_i)^2 - \sum_{i=1}^n z_i x_i^2$ is a positive definite quadratic form. Then, indeed $\sum_{i=1}^n z_i x_i^2 = \sum_i z_i$ for any ± 1 assignment to the x_i's and hence this certifies that $\sum_{i=1}^n z_i$ is a lower bound on the discrepancy. The duality says that if the vector discrepancy is D, then there always exists a choice of witnesses w_j's and z_i's of this form.

We now show the following result, which directly implies theorem 8.

Theorem 16. *Let* $C = \{S_1,\ldots,S_m\}$ *be a set system on* $[n]$ *with* $\text{vecdisc}(C) = D$. *Then* $\text{detlb}(C) = \Omega(D/\sqrt{\log n})$.

Before proving theorem 16, we show how it implies theorem 8.

Proof. (theorem 8) Let $J \subset [n]$ be such that $\text{vecdisc}(C|_J)$ is maximized. Then, it must hold that

$$\text{vecdisc}(C|_J) = \Omega(\text{herdisc}(C)/(\log m \log n)^{1/2}) \tag{18}$$

otherwise, one could use the algorithm in theorem 7 to color any subsystem $(J', C|_{J'})$ of (V,C) with discrepancy strictly less then herdisc(C), contradicting the definition of hereditary discrepancy.

Applying theorem 16 to $C|_J$ we obtain $\text{detlb}(C|_J) = \Omega(\text{vecdisc}(C|_J)/\sqrt{\log n})$. Together with (18), this gives that

$$\text{detlb}(C) \geq \text{detlb}(C|_J) \geq \Omega(\text{herdisc}(C)/(\sqrt{\log n}(\log m \log n)^{1/2})),$$

where the first inequality follows as any sub-matrix of $C|_J$ is also a sub-matrix of C. □

Proof. (Theorem 16) Let us consider the dual formulation of vector discrepancy from Lemma 15. For more convenient notation, let us write the nonnegative weight w_j as β_j^2. Moreover, let $L \subseteq [n]$ consist of the indices i with $z_i > 0$. By lemma 15 applied with $x_i = 0$ for $i \notin L$, and writing $z_i = \gamma_i^2$ for $i \in L$, we obtain

$$\sum_{i=j}^{m} \beta_j^2 \left(\sum_{i \in S_j \cap L} x_i \right)^2 \geq \sum_{i \in L} \gamma_i^2 x_i^2 \tag{19}$$

for all $x \in \mathbb{R}^L$, where $\|\beta\|^2 \leq 1$ and $\|\gamma\|^2 \geq D$.

Next, we select $K \subseteq L$ with $\|\gamma[K]\|^2 = \Omega(D/\sqrt{\log n})$ and such that all entries of $\gamma[K]$ are within a factor of 2 of each other (here $\gamma[K]$ denotes the vector γ restricted to coordinates in K). Such a subset K exists for the following reason: Let $\gamma_{\max} = \max_i |\gamma_i|$. For $\ell = 0, 1, 2, \ldots, 2\log n - 1$, let $K_\ell = \{i : |\gamma_i| \in (2^{-\ell-1}\gamma_{\max}, 2^{-\ell}\gamma_{\max}]\}$. The contribution to $\|\gamma\|$ of the components of γ with $\gamma_i \leq \gamma_{\max}/n^2$ is negligible, and so there exists some ℓ_0 for which $\sum_{i \in K_{\ell_0}} \gamma_i^2 = \Omega(\|\gamma\|^2/\log n)$.

Let us denote $k = |K|$ and $\tilde{D} = \frac{1}{2}\|\gamma[K]\|$. As $\sum_{i \in K} \gamma_i^2 = 4\tilde{D}^2$, and all these γ_i are within twice of each other, we have that $\gamma_i \geq \tilde{D}/\sqrt{k}$ for all $i \in K$. So, restricting (19) to vectors x with $x_i = 0$ for $i \notin K$, we have that

$$\sum_{j=1}^{m} \beta_j^2 \left(\sum_{i \in S_j \cap K} x_i \right)^2 \geq \frac{\tilde{D}^2}{k} \sum_{i \in K} x_i^2. \tag{20}$$

Let $C = A[*, K]$ be the $m \times k$ incidence matrix of the system $A|_K$ and let $\check{C}$ be the $m \times k$ matrix obtained from C by multiplying the jth row by β_j. Then (20) can be rewritten as

$$x^T \check{C}^T \check{C} x = \|\check{C}x\|^2 \geq \frac{\tilde{D}^2}{k}\|x\| \qquad \text{for all } x \in R^k.$$

This, by the usual variational characterization of eigenvalues, tells us that the smallest eigenvalue of the $k \times k$ matrix $\check{C}^T\check{C}$ is at least $\tilde{D}^2/k$. Now, as the determinant is the product of eigenvalues, this implies that $\det(\check{C}^T\check{C}) \geq (\tilde{D}^2/k)^k$. By the Binet–Cauchy formula we obtain

$$\det(\check{C}^T\check{C}) = \sum_I \det(\check{C}[I,*])^2, \tag{21}$$

where the summation is over all k-element subsets $I \subseteq [m]$ and $\check{C}[I,*]$ consists of the rows of $\check{C}$ whose indices lie in I. Setting $M := \max_I |\det(C[I,*])|$ and noting that $\det(\check{C}[I,*]) = \det(C[I,*])\prod_{j\in I}\beta_j$, we can bound the right-hand side of (21) as

$$\begin{aligned}\sum_I \det(\check{C}[I,*])^2 &= \sum_I \det(C[I,*])^2 \prod_{j\in I}\beta_j^2 \le M^2 \sum_I \prod_{j\in I}\beta_j^2 \\ &\le M^2 \frac{\left(\sum_{j=1}^m \beta_j^2\right)^k}{k!} \le \frac{M^2}{k!},\end{aligned}$$

where the second inequality follows as every term $\prod_{j\in I}\beta_j^2$ occurs $k!$ times in the multinomial expansion of $(\beta_1^2+\cdots+\beta_m^2)^k$. Letting $B := C[I,*]$ for an I maximizing $|\det C[I,*]|$, we have

$$\det(B)^2 \ge k!\det(\check{C}^T\check{C}) \ge k!(\tilde{D}^2/k)^k \ge (k/e)^k(\tilde{D}^2/k)^k = \Omega(\tilde{D})^{2k} = \Omega(D/\sqrt{\log n})^{2k}.$$

So the $k\times k$ matrix B witnesses that $\mathrm{detlb}(C) = \Omega(D/\sqrt{\log n})$, which implies the claimed result. □

References

1. N. Alon and J. H. Spencer. *The probabilistic method*. Wiley, New York, 2 edition, 2000.
2. W. Banaszczyk. Balancing vectors and gaussian measures of n-dimensional convex bodies. *Random Struct. Algorithms*, 12(4):351–360, 1998.
3. N. Bansal. Constructive algorithms for discrepancy minimization. In *Foundations of Computer Science (FOCS)*, pages 3–10, 2010.
4. N. Bansal and J. Spencer. Deterministic discrepancy minimization. In *ESA*, pages 408–420, 2011.
5. J. Beck. Roth's estimate on the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1:319–325, 1981.
6. J. Beck and T. Fiala. Integer-making theorems. *Discrete Applied Mathematics*, 3:1–8, 1981.
7. 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.
8. K. Chandrasekaran and S. Vempala. A discrepancy based approach to integer programming. *Arxiv*, 1111.4649, 2011.
9. M. Charikar, A. Newman and A. Nikolov. Tight hardness results for minimizing discrepancy. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2011.
10. B. Chazelle. *The discrepancy method: randomness and complexity*. Cambridge University Press, 2001.
11. T. M. Cover and J. A. Thomas. *Elements of information theory*, Wiley, 2006.
12. B. Doerr. Linear and hereditary discrepancy. *Combinatorics, Probability & Computing*, 9(4):349–354, 2000.
13. M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*. Springer-Verlag, 1997.

14. F. Eisenbrand, Dömötör Pálvölgyi, and Thomas Rothvoss. Bin packing via discrepancy of permutations. In *Symposium on Discrete Algorithms (SODA)*, pages 476–481, 2011.
15. V. Guruswami. Inapproximability results for set splitting and satisfiability problems with no mixed clauses. *Algorithmica*, 38(3):451469, 2003.
16. J. H. Kim, J. Matoušek, and V. Vu. Discrepancy after adding a single set. *Combinatorica*, 25(4):499–501, 2005.
17. D. Kleitman. On a combinatorial problem of Erdos. *J. Combinatorial Theory*, 1:209214, 1966.
18. L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *Europ. J. Combin.*, 7:151–160, 1986.
19. S. Lovett and R. Meka. Constructive Discrepancy Minimization by Walking on the Edges. *Foundation of Computer Science (FOCS)*, pages 61–67, 2012.
20. J. Matoušek. The determinant bound for discrepancy is almost tight. Manuscript, Arxiv 1101:0767.
21. J. Matoušek. *Geometric Discrepancy: An Illustrated Guide*. Springer, 2010.
22. A. Nikolov, K. Talwar and L. Zhang. The Geometry of Differential Privacy: the Approximate and Sparse Cases. *ACM Symposium on Theory of Computing (STOC)*, 2013, to appear.
23. D. Pálvölgyi. Indecomposable coverings with concave polygons. *Discrete Comput. Geom.*, 44:577–588, 2010.
24. T. Rothvoss. The entropy rounding method in approximation algorithms. *Symposium on Discrete Algorithms (SODA)*, pages 356–372, 2012.
25. T. Rothvoss. Approximating Bin Packing within $O(\log OPT \log\log OPT)$ bins. *Arxiv*, 1301.4010, 2013.
26. M. Sipser. *Introduction to the theory of computation*. PWS Publishing Company, 1997.
27. J. Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.
28. A. Srinivasan. Improving the discrepancy bound for sparse matrices: Better approximations for sparse lattice approximation problems. *Symposium on Discrete Algorithms (SODA)*, pages 692–701, 1997.
29. L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.
30. R. Vershynin. Johns decompositions: Selecting a large part. *Israel Journal of Mathematics*, 122:253-277, 2001.