Some Core Topics in Combinatorial Discrepancy Theory and Directions of Progress

Thomas Zhang

Contents

1 Introduction

In my experience so far, the core problems in combinatorial discrepancy theory intuitively boil down to "balancing" objects, where "balancing" refers to assigning $\pm$ signs to the objects such that a quantity dependent on these objects and their signs is minimized−hence the name "discrepancy". For the first two sections, these objects are coordinate vectors of some sort; for the third section, these objects are matrices of some sort. At face value, this seems like an innocuous enough of a problem. However, what is remarkable about combinatorial discrepancy theory is the spread and depth of the techniques researchers have thrown at it over the past decades−and the pool of techniques continues to grow at a terrific rate to this day!

From my perspective, the work on combinatorial discrepancy theory can be partitioned into two distinct eras: pre-constructive algorithms (for Spencer's Theorem), and post-constructive algorithms. During the first era, we saw remarkable existence proofs, many of which we will go over in the following two sections, such as Spencer's Theorem [14] and Banaszczyk's bound [1] on the Komlós conjecture. During this time, combinatorial discrepancy theory was motivated by, well, combinatorial problems such as hypergraph coloring problems [15], but it was later discovered that some aspects of convex geometry were deeply connected to the combinatorial narrative [1, 8]. However, all the important results of this time were existence proofs, and the algorithms that were floating around, such as the one proposed by Beck-Fiala [3] for sparse vector systems were far from the conjectured best bounds. This was more or less the state of the subject until around 2010, when Nikhil Bansal proposed an efficient randomized algorithm for Spencer's Theorem, introducing a stroke of mathematical optimization in the form of semi-definite programming to discrepancy theory. Since then, activity surrounding combinatorial discrepancy has exploded once again. Tools from the depths of algorithmic theory and mathematical programming have been adapted in myriad creative ways to create constructive algorithms for all the main theorems in discrepancy−as well as for all the best upper bounds for the big conjectures in the subject [12, 3, 6, 11]. Very recently even, we are witnessing connections being made via the new technique of interlacing polynomials, forming a bridge between random matrix theory results and discrepancy [7, 16].

There are essentially two discrepancy problems that I will address in this article−one which I will call Spencer's Problem, and the other Komlós' Problem.

Definition 1.1 (Spencer's Problem ver. 1) *Given a set system $\mathbb{S}$ on $[n] := \{1, \ldots, n\}$ with $|\mathbb{S}| = m$ sets S, let us consider colorings $\chi \in \{\pm 1\}^n$ of $[n]$. Define the (set) discrepancy $\mathbb{S}$ to be:*

$$\begin{aligned} disc\,(\mathbb{S}) &= \min_{\chi \in \{\pm 1\}^n} \max_{S \in \mathbb{S}} \sum_{i=1}^{|S|} \chi(S(i)) \\ &:= \min_{\chi \in \{\pm 1\}^n} \max_{S \in \mathbb{S}} |\chi(S)|\,. \end{aligned}$$

I often find it easier to consider the equivalent matrix form of the problem:

Definition 1.2 (Spencer's Problem ver. 2) *Given a matrix $A \in M_{m,n}$, where its entries $A(i,j) \in \{0,1\}$, let us consider ± 1 vectors $x \in \{\pm 1\}^n$. Define the (set) discrepancy of A to be:*

$$\begin{aligned} disc\,(A) &= \min_{x \in \{\pm 1\}^n} \max_{i=1,\ldots,m} |(Ax)_i| \\ &= \min_{\chi \in \{\pm 1\}^n} \|Ax\|_\infty\,. \end{aligned}$$

To translate from ver 1. to ver 2, view the columns of A as characteristic vectors of sets $S \in \mathbb{S}$.

What is this problem intuitively asking? Consider the following: we have m shoppers, each holding a shopping catalog with n items. Each shopper can either check or skip each item, where their completed catalog corresponds to a set S. Now, we want to know what is the best way to assign each shopper to one of two groups (corresponding to ± 1) such that when we add up the number of times each item is checked on the catalogs of each group, no item is overrepresented in one group over the other.

The Komlós problem is almost identical to Spencer's problem, except where the values in A are 0/1 in Spencer's problem, we generalize this to real values now with a norm restriction.

Definition 1.3 (Komlós Problem) *Given a square matrix $A \in M_n$, where each column has Euclidean norm no more than 1: $\|A_{\cdot,j}\|_2 \leq 1$, let us consider ± 1 vectors $x \in \{\pm 1\}^n$. Define the (set) discrepancy of A to be:*

$$\begin{aligned} disc\,(A) &= \min_{x \in \{\pm 1\}^n} \max_{i=1,\ldots,m} |(Ax)_i| \\ &= \min_{x \in \{\pm 1\}^n} \|Ax\|_\infty . \end{aligned}$$

In the next section, we will discuss Spencer's Theorem, which establishes a uniform upper bound on the discrepancy given *any* 0/1 matrix A, with special attention given to the constructive proof by Lovett and Meka [6]. This section will serve to introduce many proof techniques and structures that will be useful to understand in later sections.

In section 3, we will encounter the Komlós *conjecture*, which posits that the discrepancy in the Komlós problem is constant regardless of the dimension. We will then see why the Komlós conjecture is a powerful claim. The latter subsections will be dedicated to the semi-definite relaxation of the Komlós conjecture introduced by Nikolov [11], where we introduce a new proof to Nikolov's result using no more than elementary matrix theory, and then propose a randomized rounding algorithm inspired by Goeman's and Williamson's approximation scheme [5].

In the last section, we will generalize Spencer's and Komlós' problems to the matrix case. We then discuss novel preliminary results, including a semi-definite relaxation of the "Matrix Komlós" conjecture, and a surprising generalized random coloring lemma using the non-commutative Khintchine's inequality. We then conclude this article with a few remarks on current proof directions as well as numerically supported conjectures.

2 Spencer's Theorem

2.1 Result, Some Recurring Lemmas, and a Lower Bound

In the celebrated paper "Six Standard Deviations Suffice" [14], Spencer answers the question of what one should expect for the worst-case discrepancy.

Theorem 2.1 (Spencer's Theorem, 1985) *Let $\mathbb{S}$ be a set system on $[n]$ with $|\mathbb{S}| = m \geq n$. Then,*

$$disc\,(\mathbb{S}) = O\left(\sqrt{n \log(2m/n)}\right).$$

In particular, if $m = O(n)$, then $disc\,(\mathbb{S}) = O(\sqrt{n})$.

We will first show that Spencer's bound is tight up to constants by introducing the Hadamard set system (defined by Hadamard matrices) [8]. In other words, we should not hope for any tighter bound than Spencer's Theorem.

Definition 2.2 (Hadamard matrix) *$H \in M_n(\mathbb{R})$ is a Hadamard matrix if all its entries satisfy $H(i,j) \in \{\pm 1\}$ and H is an orthogonal matrix (scaled by $\sqrt{n}$). We note that Hadamard matrices do not exist for all dimensions.*

Let us define a recursive family of Hadamard matrices $H_k \in M_{2^k}(\mathbb{R})$:

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_k = \begin{bmatrix} H_{k-1} & H_{k-1} \\ H_{k-1} & -H_{k-1} \end{bmatrix}.$$

It is easy to verify that H_k satisfy $H_k^\top H_k = 2^k I_{2^k}$ by using the fact that $H_1^\top H_1 = 2I_2$ and then using induction and block matrix multiplication.

Theorem 2.3 *Let $H \in M_n(\mathbb{R})$ be a Hadamard matrix from the family we constructed. Let $J = \mathbb{1}\mathbb{1}^\top$ be the all-ones matrix. Then $A := \frac{1}{2}(H + J)$ is a $0/1$ matrix and satisfies*

$$disc\,(A) \geq \frac{\sqrt{n}}{2}.$$

Proof of Theorem 2.3: We first observe the following equivalence of norms:

$$n\,\|x\|_\infty^2 \geq \|x\|_2^2\,.$$

We then observe that the first row of H must be all ones. Therefore, we know that $H^\top J = ne_1^\top \mathbb{1}$, i.e. the matrix that is 1 on the first row and 0 elsewhere, since the other rows of H are orthogonal to $\mathbb{1}$. Let $x \in \{\pm 1\}^n$. We have the following series of inequalities

$$\begin{aligned}
n\,\|Ax\|_\infty^2 &\geq \|Ax\|_2^2 \\
&= x^\top A^\top A x \\
&= \frac{1}{4} x^\top (H+J)^\top (H+J) x \\
&= \frac{1}{4} x^\top \left(H^\top H + J^\top H + H^\top J + J^\top J \right) x \\
&= \frac{1}{4} x^\top \left(nI_n + n\mathbb{1}^\top e_1 + ne_1^\top \mathbb{1} + nJ \right) x \\
&= \frac{1}{4} \left(n \sum_{i=1}^n x_i^2 + 2nx_1 \left(\sum_{i=1}^n x_i \right) + n \left(\sum_{i=1}^n x_i \right)^2 \right) \\
&\geq \frac{1}{4} \left(n^2 \right).
\end{aligned}$$

Thus, we have that

$$\|Ax\|_\infty \geq \frac{\sqrt{n}}{2},$$

regardless of our choice of $x \in \{\pm 1\}^n$. ■

Having established a lower bound for Spencer's Problem, let us now prove some simple, but key lemmas that get us somewhat close to Spencer's Theorem. We will first look at the Random Coloring Lemma, which says that assigning signs to sets by flipping a coin gets us the logarithmically close to Spencer's Theorem most of the time.

Theorem 2.4 (Random Coloring Lemma) *For any set system $\mathbb{S}$ on $[n]$, $|\mathbb{S}| = m$, a uniform random coloring $\chi \in \{\pm 1\}^n$ of $[n]$ satisfies:*

$$|\chi(S)| \leq \sqrt{2n \ln(4m)} = O(\sqrt{n \ln(m)}) \quad \forall S \in \mathbb{S} \textit{ with probability at least } 1/2.$$

Proof of Theorem 2.4: We observe that fixing a $S \in \mathbb{S}$, where $|S| \leq n$, $\chi(S)$ becomes a sum of Bernoulli (alternatively Radermacher) random variables, which asymptotically approaches a normal distribution with mean 0 and standard deviation at most $\sqrt{n}$. By an application of the Chernoff bound and union bound over the m sets in $\mathbb{S}$, we get

$$\Pr\left\{|\chi(S)| \geq t\sqrt{n}\right\} \leq 2me^{\frac{-t^2}{2}}$$

Plugging in $t = \sqrt{2\ln(4m)}$, we get the desired bound. ■

At first glance, it seems difficult to improve this bound, since there is little structure left in the problem that we haven't already used, but Spencer's Theorem shows that an improvement is in fact possible.

The next lemma is key to the standard proof of Spencer's Theorem [8, 3] and is in general a powerful structure used in many algorithms for discrepancy minimization. Where it may be difficult to find a good coloring for the entire set system at once, it might be easier to partially color the set system iteratively: indeed, this is the case in the constructive proof of Spencer's Theorem proposed by Lovett and Meka [6] that we will seen soon, as well as its descendants [12].

Theorem 2.5 (Partial Coloring Lemma) *For two set systems $\mathbb{M}$, $\mathbb{F}$ on $[n]$, where $\mathbb{F}$ satisfies*

$$\prod_{F \in \mathbb{F}} (|F| + 1) \leq 2^{\frac{n-1}{5}}$$

there exists a partial coloring $\chi \in \{-1, 0, +1\}^n$ such that χ has at least $n/10$ non-zero elements, and

$$\chi(F) = 0 \;\forall F \in \mathbb{F}$$
$$|\chi(M)| \leq \sqrt{2n \ln(4\,|\mathbb{M}|)} \;\forall M \in \mathbb{M}.$$

This basically tells us that for a small enough subcollection of a set system, we can find a "perfect" partial coloring on it while incurring discrepancy of no more than a good random coloring on the rest of the set system.

Proof of Theorem 2.5: Consider the collection of all full colorings $\mathcal{C}_0 := \{\pm 1\}^n$. By the Random Coloring Lemma, we know that at least one half of these colorings satisfy $|\chi(M)| \leq \sqrt{2n\ln(4\,|\mathbb{M}|)}\ \forall M \in \mathbb{M}$. Let us call those colorings $\mathcal{C}_1$, and $|\mathcal{C}_1| \geq \frac{1}{2}\,|\mathcal{C}_0| = 2^{n-1}$.

We will now set the stage for a pigeonhole principle argument. Let us define a mapping $g : \mathcal{C}_1 \to \mathbb{Z}^{|\mathbb{F}|}$. which sends a coloring $\chi \in \mathcal{C}_1$ to the $|\mathbb{F}|$-dimensional vector that records the discrepancy it incurs on each $F \in \mathbb{F}$. In other words, $g(\chi) = \big(\chi(F)\big)_{F\in\mathbb{F}}$. What is the cardinality of the range? We observe two things:

1. $|\chi(F)| \leq |F|$ for each $F \in \mathbb{F}$
2. $\chi(F) - F$ is always even.

By a simple counting argument, this gets us

$$|\mathrm{Ran}(g)| \leq \prod_{F \in \mathbb{F}} (|F| + 1) \leq 2^{\frac{n-1}{5}} \text{ (by assumption).}$$

Since $|\mathcal{C}_1| \geq 2^{n-1}$, by the pigeonhole principle, we know there is some vector $g(\chi_0)$ such that at least $2^{\frac{4}{5}(n-1)}$ colorings are sent to $g(\chi_0)$. Define this subcollection of colorings $\mathcal{C}_2 := \{\chi \in \mathcal{C}_1 : g(\chi) = g(\chi_0)\}$.

We now define another mapping $h : \mathcal{C}_2 \to \{-1, 0, +1\}^n$, where $h(\chi) = \frac{1}{2}(\chi - \chi_0)$, assuming that we've picked and fixed a representative of χ_0 from above. Now we want to argue that $\mathrm{Ran}(h)$ has some element that has at least $n/10$ non-zero elements. To do this, we can use a very crude estimate by counting all the ways to pick $k < n/10$ items out of $[n]$ to be non-zero and the 2^k possible partial colorings thereafter:

$$\sum_{0 \leq k < n/10} \binom{n}{k} 2^k.$$

We can use the estimate

$$\sum_{0 \leq k \leq z} \binom{n}{q} a^k \leq \left(\frac{ean}{z}\right)^z$$

to get

$$\sum_{0 \leq k < n/10} \binom{n}{k} 2^k \leq \left(\frac{2en}{n/10}\right)^{n/10} < 60^{n/10} < (2^6)^{n/10} < 2^{\frac{4}{5}(n-1)} \leq |\mathcal{C}_2|\,.$$

Therefore, by the above counting argument, we know that that $\mathcal{C}_2$ is large enough to contain some partial coloring that has at least $n/10$ non-zero elements. By our construction of $\mathcal{C}_2$, we know this partial coloring (call it χ') satisfies the following properties:

1. $\chi' = \frac{1}{2}(\chi_1 - \chi_0), \quad \chi_1 \in \mathcal{C}_2$
2. $\chi'(F) = 0 \quad \forall F \in \mathbb{F}$
3. $|\chi'(M)| \leq \sqrt{2n \ln(4\,|\mathbb{M}|)} \quad \forall M \in \mathbb{M}.$

This completes the proof of the partial coloring lemma. ■

Say that we have an algorithm that can get us a partial coloring that satisfies the conditions in the Partial Coloring lemma: we can get a full coloring by iteratively finding a partial coloring. The naivety of this algorithm is reflected in the bounds that we get from such a scheme.

Corollary 2.6 (Full coloring from partial coloring) *Given a set system of size m on $[n]$, and we have a process to construct a partial coloring that satisfies the properties of the partial coloring lemma, then we can construct a full coloring that has discrepancy $\leq O(\sqrt{n \ln(m)})$.*

Proof of corollary: The algorithm is simple. We find the partial coloring, fix that coloring and recursively find a partial coloring on the remaining uncolored items. We prove that this process ends up with discrepancy $O(\sqrt{n \ln(m)})$. We ignore the distinction between the different set subsystems $\mathbb{F}$ and $\mathbb{M}$, because those can change between iterations. Instead, we pretend that a given set is always in $\mathbb{M}$ and in turn upper bound the resulting discrepancy. After iteration k, observe that we have at most $\left(\frac{9}{10}\right)^k$ remaining uncolored items, and hence iteration k will add discrepancy at most $\sqrt{2\left(\frac{9}{10}\right)^k n \ln(4m)}$ to each set. Therefore,

$$\begin{aligned}
\forall S \in \mathbb{S}, \quad \operatorname{disc}(S) &\leq \sqrt{2n\ln(4m)} + \sqrt{2\frac{9}{10} n \ln(4m)} + \cdots + \sqrt{2\left(\frac{9}{10}\right)^k n \ln(4m)} + \cdots \\
&= \sum_{k=1}^{\infty} \sqrt{2\left(\frac{9}{10}\right)^k n \ln(4m)} \\
&= \sqrt{2n\ln(4m)} \sum_{k=1}^{\infty} \left(\sqrt{\frac{9}{10}}\right)^k \\
&= C\sqrt{2n\ln(4m)} = O(\sqrt{n\ln(m)}).
\end{aligned}$$

■

2.2 A Leap: A Constructive Proof

In this section, we explore Lovett and Meka's constructive proof of Spencer's Theorem [6]. We should note that Bansal's original algorithm using semidefinite programming [2] is not actually a constructive proof, rather it requires Spencer's Theorem to be true in order to argue a certain semidefinite program is feasible.

Before we discuss the algorithm, it is important to note that Lovett and Meka's algorithm does not find the *minimum* discrepancy. In fact, we should not hope to find an efficient algorithm that finds the minimizing discrepancy due to the following complexity result:

Theorem 2.7 (Charikar-Newman-Nikolov [10]) *Given a set system $\mathbb{S}$ on $[n]$ such that $|\mathbb{S}| = O(n)$, it is NP-hard to determine whether*

$$disc(\mathbb{S}) = 0 \quad or \quad disc(\mathbb{S}) = \Omega(\sqrt{n}).$$

Therefore, we can only hope for Lovett and Meka's algorithm to return a coloring that incurs a discrepancy no more than the one in Spencer's Theorem. Before we look at the algorithm, let us first understand the high-level structure of the algorithm and the introduce some notation. The input of the algorithm are some parameters that are "small" and in initial point in the hypercube $[-1,+1]^n$, which we can take without loss of generality to be 0. The algorithm then iteratively finds a direction restricted to a subspace that it picks at random and walks along it for a certain step size γ. With high probability, this direction gets closer to a vertex. When a coordinate is within a sufficiently close margin δ to ± 1, we consider that direction to be "dead" and for the iterations that follow, we search orthogonal to that direction so that we maintain closeness to ± 1 in that coordinate. At the same time, when the location gets dangerously close to violating the discrepancy parameter with respect to some set (a.k.a. column vector in the matrix version of Spencer's Problem), we deem that set "dangerous" and for future iterations search orthogonal to it so that the discrepancy measured on that set never blows up. The claim that Lovett and Meka make is that such an algorithm finds with non-zero probability a half-coloring of $[n]$ that does not violate any of the discrepancy parameters.

Let us now propose the algorithm. We are given a 0/1 matrix $A \in M_{m,n}$, where we denote its rows $a_j \in \mathbb{R}^m$. Let us further denote $|a_j| = \|a_j\|_1$ to be the number of items in that set. Let $\lambda \in \mathbb{R}^m$ be a vector of parameters $\lambda_j \geq 0$ such that

$$\sum_{j=1}^{m} e^{-\lambda_j^2/32} \leq n/16.$$

Let the step size be $\gamma = \frac{1}{100n^2}$ and margin $\delta = 10\gamma \log(n)$. We denote subspaces by U_k.

Algorithm 1 LovettMeka$(A, \lambda, \gamma, \delta)$

Initialize $x_0 = 0$, Dead $= \emptyset$, Danger $= \emptyset$, $U_0 = \mathbb{R}^n$
for $k = 1, 2, \ldots, K = 8/\gamma^2$ **do**
    Sample random vector v_k: $v \sim N(0, I_n)$, $v_k = \mathrm{proj}_{U_k}(v)$
    Take a random step $x_k = x_{k-1} + \gamma v_k$
    For $i \in [n]$ such that $|x_k(i)| \geq 1 - \delta$, add $i \to$ Dead
    For $j \in [m]$ such that $|\langle x_k, a_j \rangle| \geq \lambda_j \sqrt{|a_j|} - \delta$, add $a_j \to$ Danger
    Set U_{k+1} such that $U_{k+1} \perp$ Dead and $U_{k+1} \perp$ Danger
end for
(Rounding) If $|x_K(i)| \geq 1 - \delta$, set $x'(i) = \mathrm{sign}(x_K(i))$.
return x'

The main result of this algorithm is the following

Theorem 2.8 *Given that we set the parameters λ, γ, δ as described above, then for any starting point $x \in [-1, +1]^n$, the algorithm finds a partial coloring $x' \in [-1, +1]^n$ where at least $n/2$ entries are ± 1 and*

$$\left| \left\langle x' - x, a_j \right\rangle \right| \leq \lambda_j \sqrt{|a_j|} \quad \forall j \in [m]$$

with probability at least $1/8$.

Proof of Spencer's Theorem using Theorem 2.8: Like we earlier mentioned for the Partial Coloring Lemma, getting Spencer's Theorem from Lovett and Meka's algorithm that returns a partial coloring simply involves running the algorithm $O(1/\log(n))$ times. Let's say we are on the t iteration of running the algorithm. Observe that the number of vertices left to color is at most $n_t \leq n/2^{t-1}$. We set

$$\lambda_j^t = \sqrt{32 \left(\log \left(\frac{m}{n_t} \right) + \log(16) \right)},$$

such that

$$\begin{aligned} \sum_{j=1}^{m} e^{-\lambda_j^2/32} &= \sum_{j=1}^{m} \exp \left(\log \left(\frac{n_t}{m} \right) - \log(16) \right) \\ &= n/16. \end{aligned}$$

By the algorithm, we know that x_t was found searching only over the dimension corresponding to the n_t uncolored items, and thus

$$\langle x_t, a_j \rangle \leq \lambda_j \sqrt{\left| S_j^t \right|} \leq \lambda_j \sqrt{n_t}.$$

Let's say at iteration T we have a full coloring x_T. We can bound the discrepancy incurred using the triangle inequality:

$$\begin{aligned}
\operatorname{disc}(x_T, A) &= |Ax_T|_\infty \\
&= \max_j |\langle x_T, a_j \rangle| \\
&= \max_j \left| \left\langle \sum_{t=0}^{T-1} x_{t+1} - x_t, a_j \right\rangle \right| \\
&\leq \max_j \sum_{t=0}^{T} |\langle x_{t+1} - x_t, a_j \rangle| \\
&\leq \max_j \sum_{t=0}^{T} 2\lambda_j \sqrt{n_t} \\
&\leq C \sum_{t=0}^{T} \sqrt{n_t \log\left(\frac{m}{n_t}\right)} \\
&\leq O\left(\sqrt{n \log\left(\frac{m}{n}\right)}\right) \quad \text{since } n_t \leq \frac{n}{2^{t-1}}.
\end{aligned}$$

Having demonstrated that this proves Spencer's Theorem, let us now show that the algorithm actually does what it claims to do.

Proof of Theorem 2.8: The proof of the theorem really comes down to understanding the properties of how Gaussian random vectors behave under orthogonal projections. Let us propose the following three related facts about Gaussian random vectors and projections:

Lemma 2.9 *If $u \in \mathbb{R}^n$ is an arbitrary vector, and $v \in \mathbb{R}^n$ is a Gaussian random vector with entries $v_i \sim N(0, \sigma^2)$, then*

$$\langle u, v \rangle \sim N\left(0, \sigma^2 \|u\|_2^2\right).$$

Corollary 2.10 *If $v \sim N(0, \sigma^2 I_n)$ is a Gaussian random vector, $S \subseteq \mathbb{R}^n$ is a subspace with $\dim(S) = k$, and $v' = proj_S(v)$, then*

$$\mathbb{E}\left[\|v'\|_2^2\right] = k\sigma^2.$$

Corollary 2.11 *If $u \in \mathbb{R}^n$ and $v \sim N(0, \sigma^2 I_n)$ is a Gaussian random vector, and $v' = proj_S(v)$ is the orthogonal projection of v onto a subspace $S \subseteq \mathbb{R}^n$, then*

$$\langle u, v' \rangle \sim N\left(0, \alpha \|u\|_2^2\right),$$

where $\alpha \leq \sigma^2$.

To prove the first lemma, we observe that $\langle u, v \rangle$ is a sum of i.i.d. Gaussian random variables, and then we use the additivity of Gaussian distributions to complete the proof.

To see how the first lemma implies the second, we find an orthonormal basis $\{s_1, \ldots, s_k\}$ of S and observe

$$\|v'\|_2^2 = \sum_{i=1}^{k} \langle v, s_i \rangle^2.$$

Similarly, the final lemma is proven using the first and an application of Cauchy-Schwarz.

Proof of Theorem 2.8: Our goal is to show that with non-zero probability, $|\mathrm{Dead}| \geq n/2$ after running the algorithm. We take a look at the final location x_K and bound its norm using the previously established facts about Gaussian random variables:

$$\begin{aligned}
\mathbb{E}\left[\|x_K\|_2^2\right] &= \mathbb{E}\left[\left\|\gamma \sum_{k=1}^{K} g_k\right\|_2^2\right] \\
&= \gamma^2 \sum_{k=1}^{K} \mathbb{E}\left[\|g_k\|_2^2\right] \quad \text{since } g_k \text{ are independent} \\
&= \gamma^2 \sum_{k=1}^{K} \mathbb{E}\left[\dim(V_k)\right] \quad \text{by Corollary 2.10} \\
&\geq \gamma^2 K \mathbb{E}\left[\dim(V_K)\right] \quad \text{since } V_k \subset V_{k-1} \\
&= \gamma^2 \frac{8}{\gamma^2} \mathbb{E}\left[\dim(V_K)\right] \\
&\geq 8(n - |\mathrm{Dead}| - |\mathrm{Danger}|).
\end{aligned}$$

Since we know that the coordinates of x_K freeze before exceeding 1, we know that

$$\mathbb{E}\left[\|x_K\|_2^2\right] \leq n,$$

which put together with the above lower bound gets us

$$\begin{aligned}
n &\geq 8(n - |\mathrm{Dead}| - |\mathrm{Danger}|) \\
\iff \quad |\mathrm{Dead}| &\geq \frac{7}{8} n - |\mathrm{Danger}|.
\end{aligned}$$

By our choice of γ, δ and λ, with a little effort we can show that

$$\frac{9}{16} \leq \mathbb{E}[|\mathrm{Dead}|] \leq \left(1 - \frac{p}{2}\right) n,$$

where p is the probability that $\mathbb{E}[|\mathrm{Dead}|] < n/2$. This gets us a chance of failure $p < 7/8$, i.e. the algorithm has at least a $1/8$ chance of succeeding. ■

3 The Komlós Conjecture

3.1 Statement and Importance

The statement of the Komlós conjecture is quite simple.

Conjecture 3.1 (Komlós) *Given square matrix $A \in M_n$ where the columns of A have Euclidean norm at most 1, we have*

$$disc\,(A) = \max_{x \in \{\pm 1\}^n} \|Ax\|_\infty \leq K$$

where K is a constant independent of dimension.

We observe that this conjecture is quite powerful. Clearly, it implies the $m = O(n)$ case of Spencer's theorem, where we normalize the columns of A by $\sqrt{n}$, and in the same vein it implies the related Beck-Fiala conjecture, which states that if A is a $0/1$ matrix and each column is t-sparse, then $\mathrm{disc}\,(A) \leq O(\sqrt{t})$.

Taking the conjecture head on has been difficult: so far no bounds have been able to shake off the logarithmic factor. The current best bound comes from convex geometry: Banaszczyk's used results related to Gaussian volume to get a bound of $O(\sqrt{\log(n)})$ [1]. Recently, the algorithmic world has caught up to the convex geometric result: Levy et al. [12] have built upon earlier work by Bansal to create a deterministic algorithm that attains Banaszczyk's bound. The algorithm relies on extremely careful subspace arguments akin to the ones seen in Lovett and Meka, except where Lovett and Meka narrow down a subspace where a random direction improves discrepancy with good probability, Levy et al. narrow down a subspace where *any* direction improves discrepancy, hence the need for far more careful subspace arguments. However, the inescapable partial coloring nature of the algorithm forces it to accrue a residual logarithmic factor−a recurring theme in algorithms for Komlós conjecture that rely on partial coloring.

Parallel to the recent algorithmic work, Nikolov [11] introduced a novel notion of discrepancy that is naturally connected to the Komlós conjecture. We discuss this in the following section.

3.2 Semidefinite Relaxation is True

Recall that the discrepancy of a matrix A as defined prior to this section is defined

$$\mathrm{disc}\,(A) = \min_{x \in \{\pm 1\}^n} \|Ax\|_\infty .$$

To avoid confusion, we will refer to this as the "set" discrepancy of A, as it describes the combinatorial coloring of objects in sets. Nikolov proposes a version of discrepancy problem that is relaxed in two ways. Firstly, instead of having A as a $0/1$ matrix, we can relax the conditions such that the columns of A are bounded in the Euclidean norm, w.l.o.g. at most norm 1, where the scaling simply makes it relevant to the Komlós conjecture. Secondly, instead of considering $x \in \{\pm 1\}^n$, which is analogous to assigning a sign to each column of A, we instead consider assigning a unit-length vector to each column of A. Formalizing this notion: we observe that $Ax = \sum_{j=1}^n A_{\cdot,j} x(j)$, where $A_{\cdot,j}$ is the j-th column of matrix A and $x(j)$ is the j-th element of x, i.e. the "sign" of $A_{\cdot,j}$. Observe that

$$\|Ax\|_\infty = \max_{i=1,\ldots,m} \left| A_{i,\cdot}^\top x \right| = \max_{i=1,\ldots,m} \left| \sum_{j=1}^n A_{ij} x(j) \right| .$$

In the relaxation, we replace $x(j)$ with a unit vector u_j and take $\left\|\sum_{j=1}^n A_{ij} u_j\right\|_2$ to get the value analogous to $\sum_{j=1}^n A_{ij} x(j)$. We then define the "vector" discrepancy:

$$\begin{aligned} \operatorname{disc}(A) &:= \min_{x(1),\ldots,x(n)\in\{\pm 1\}} \max_{i=1,\ldots,m} \left| \sum_{j=1}^n A_{ij} x(j) \right| \\ \sim \quad \operatorname{SDPdisc}(A) &:= \min_{u_1,\ldots,u_n,\, \|u_j\|_2=1} \max_{i=1,\ldots,m} \left\| \sum_{j=1}^n A_{ij} u_j \right\|_2 . \end{aligned}$$

The name "vector discrepancy" is unfortunate, as later we will introduce a notion of "matrix discrepancy" that is analogous to set discrepancy on matrices, and this notion of "matrix discrepancy" also admits its own "vector discrepancy" relaxation. Thus, from now on, we look forward in time and rename "vector discrepancy" in Nikolov's work to "SDP discrepancy", as it is a variant of discrepancy that is induced by the semidefinite relaxation.

To see why vector discrepancy is a relaxation, we will see every instance of $x \in \{\pm 1\}^n$ can be translated to a collection of unit vectors $\{u_j\}_{j=1}^n$: simply set $u_j = x(j)v$, where v is any ℓ^2 unit vector. Given $x \in \{\pm 1\}^n$ that attains $\operatorname{disc}(A)$,

$$\begin{aligned} \operatorname{disc}(A) = \|Ax\|_\infty &= \max_{i=1,\ldots,m} \left| \sum_{j=1}^n A_{ij} x(j) \right| \\ &= \max_{i=1,\ldots,m} \left| \sum_{j=1}^n A_{ij} x(j) \right| \|v\|_2 \\ &= \max_{i=1,\ldots,m} \left\| \sum_{j=1}^n A_{ij} x(j) v \right\|_2 \\ &\geq \min_{u_1,\ldots,u_n,\, \|u_j\|_2=1} \max_{i=1,\ldots,m} \left\| \sum_{j=1}^n A_{ij} u_j \right\|_2 \\ &= \operatorname{SDPdisc}(A) . \end{aligned}$$

Let us now recall the Komlós conjecture:

Conjecture 3.2 (Komlós) *Given square matrix $A \in M_n$ where the columns of A have Euclidean norm at most 1, we have*

$$\mathit{disc}(A) = \max_{x\in\{\pm 1\}^n} \|Ax\|_\infty \leq K$$

where K is a constant independent of dimension.

Nikolov's paper refutes the following line of thought: can we *disprove* the Komlós conjecture by relaxing the constraints of the set discrepancy problem and then showing the discrepancy of the relaxed problem is lower bounded by some non-constant factor, say $K\sqrt{\log(n)}$? It can be shown that this line of thought does not work, at least with the aforementioned relaxations to the problem by the following theorem:

Theorem 3.3 (SDPdisc ≤ 1) *Given any square matrix $A \in M_n$ where the columns of A have Euclidean norm at most 1, we have*

$$
\begin{aligned}
SDPdisc\,(A) &= \min_{u_1,\ldots,u_n,\, \|u_j\|_2=1} \max_{i=1,\ldots,m} \left\| \sum_{j=1}^n A_{ij} u_j \right\|_2 \\
&\leq 1.
\end{aligned}
$$

Remark 3.4 *It is not difficult to construct examples of matrices whose SDPdisc is 0 and set discrepancy is non-zero, which implies that SDPdisc and set discrepancy cannot satisfy an integrality gap:*

$$c \cdot disc(A) \leq SDPdisc(A) \leq C \cdot disc(A)$$

for all A with unit length columns.

Observe that this defeats any hope of naively using this relaxed version of the set discrepancy problem to either disprove or prove the Komlós conjecture. However, numerical evidence points toward the following phenomenon that seems to suggest that SDP discrepancy is not necessarily a bad approximation of set discrepancy.

Remark 3.5 *For matrices A that induce high set discrepancy, they also induce high SDP discrepancy. Furthermore, the optimal colorings u_j that attain the SDP discrepancy are all identical up to sign, which if we recall the translation from Komlós to SDP Komlós corresponds to a set coloring $x \in \{\pm 1\}^n$.*

We will now prove Theorem 3.3. We will not use Nikolov's proof, but it is quite a neat argument that can roughly be summarized in the following points:

1. Formulate a semi-definite program that is equivalent to the vector discrepancy problem
2. Derive the dual program that is feasible only for D such that SDPdisc$(A) \geq D$ (strong duality holds here)
3. Use linear algebra to formulate an ellipsoid argument that contradicts the dual program being feasible for any $D = 1 + \epsilon \implies$ optimal value of primal program is upper bounded by 1, and it is tight by trivial examples.

We will show that an elaborate ellipsoid argument is unnecessary and one can appeal directly to the matrix theory behind the primal and dual programs to derive the bound. Let us first rephrase the SDPdisc problem into a minimization problem:

$$
\begin{aligned}
&\min \quad t \\
&\text{subject to} \\
&t \geq \left\| \sum_{j=1}^n A_{ij} u_j \right\|_2^2 = \sum_{j=1}^n \sum_{k=1}^n A_{ij} A_{ik} u_j^\top u_k, \ i = 1, \ldots, m \\
&\|u_j\|_2 = 1, \ j = 1, \ldots, n.
\end{aligned}
$$

such that t^* is $\mathrm{SDPdisc}(A)^2$. Conveniently, we can embed this problem into a semi-definite program:

$$
\begin{aligned}
(P) \quad & \min t \\
& \text{s.t. } t \geq A_{i,\cdot}^\top X \\
& \qquad A_{i,\cdot}, \; i = 1, \ldots, m \\
& \qquad X_{jj} = 1, \; j = 1, \ldots, n \\
& \qquad X \succcurlyeq 0.
\end{aligned}
$$

We appeal to Cholesky factorization to establish the equivalence between the former and latter programs: given a list of u_j such that $\|u_j\| = 1$, we look at the matrix $X = U^\top U = \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}^\top \begin{bmatrix} u_1 & \ldots & u_n \end{bmatrix}$ and immediately observe that X is positive semi-definite and $X_{jj} = \|u_j\|_2^2 = 1$. On the other hand, given a positive semi-definite matrix X whose diagonal entries are all 1, we can use Cholesky factorization to get $X = U^\top U$, where U has n columns, and each column is a ℓ^2 unit vector.

We explicitly write out the dual of the above semi-definite program through finding the Lagrangian dual:

$$
\begin{aligned}
(D) \quad & \max \sum_{j=1}^n \lambda_j \\
& \text{s.t. } \sum_{i=1}^m \mu_i A_{i,\cdot} A_{i,\cdot}^\top \succeq \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \\
& \qquad \sum_{i=1}^m \mu_i = 1, \mu_i \geq 0, \; i = 1, \ldots, m.
\end{aligned}
$$

As it stands, we might have to consider the duality gap in such a problem. However, we will show that primal-dual Slater's condition [4] for strong duality holds here. Geometrically, we want to show that the feasible region for *both* the primal and dual problems has a non-empty interior. What this boils down to in our case that we have to show that we can attain an instance where all inequality constraints are simultaneously *strictly* satisfies. A helpful image is the Slater condition in the linear programming case: simultaneous strict satisfaction of inequality constraints corresponds to non-empty interior of the feasible region. For our primal program, this is rather simple: given matrix A, we find n linearly independent u_j such that $X = U^\top U$ has diagonals all 1 and is strictly positive definite. Since A has finitely many rows, we can pick any t such that it exceeds all $A_{i,\cdot}^\top X A_{i,\cdot}$. For our dual program, given a list of $\mu_i \geq 0$ such that $\sum_{i=1}^m \mu_i = 1$, we observe that $\sum_{i=1}^m \mu_i A_{i,\cdot} A_{i,\cdot}^\top$ is at least a psd matrix. Since λ_j are unconstrained, we simply find "negative enough" λ_j such that

$$
\sum_{i=1}^m \mu_i A_{i,\cdot} A_{i,\cdot}^\top - \mathrm{diag}(\lambda_1, \ldots, \lambda_n) \succ 0.
$$

Therefore, by showing that we can find feasible solutions to both the primal and dual programs that strictly satisfy the inequality constraints, we have shown that the feasible regions of both have nonempty interiors. Slater's condition then tells us that strong duality holds for this primal-dual pair, i.e. $t^* = \sum_{i=1}^m \lambda_j^*$.

At this point we claim the main theorem of this section

Theorem 3.6 *Given $A \in M_{m,n}$, $A = \begin{bmatrix} a_1^\top \\ \vdots \\ a_m^\top \end{bmatrix} = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}$,*

$$SDPdisc(A) \leq \min\left\{ \max_{i=1,\ldots,m} \|a_i\|_2^2 \, , \, \max_{j=1,\ldots,n} \|v_j\|_2^2 \right\}.$$

This immediately shows as a corollary the main result of Nikolov's paper:

Corollary 3.7 *Given $A \in M_{m,n}$ such that $\|v_j\|_2 \leq 1$ for $j = 1, \ldots, n$, then*

$$SDPdisc(A) \leq 1.$$

To prove the theorem, we turn our focus to the dual program in particular. Let $(\vec{\lambda}^*, \vec{\mu}^*)$ be an optimal solution to the dual program. We will prove that $\text{SDPdisc}(A) \leq \max_{i=1,\ldots,m} \|a_i\|_2^2$. Just by plugging our optimal values into the constraints of the dual program, we get:

$$\begin{aligned} \sum_{i=1}^m \mu_i^* a_i a_i^\top &\succeq \text{diag}(\lambda_1^*, \ldots, \lambda_n^*) \\ \implies \text{tr}\left(\sum_{i=1}^m \mu_i^* a_i a_i^\top\right) &\geq \sum_{j=1}^n \lambda_j^* = \text{SDPdisc}(A) \\ \text{tr}\left(\sum_{i=1}^m \mu_i^* a_i a_i^\top\right) &= \sum_{i=1}^m \text{tr}\left(\mu_i^* a_i a_i^\top\right) \\ &= \sum_{i=1}^m \text{tr}\left(\mu_i^* a_i^\top a_i\right) \\ &= \sum_{i=1}^m \mu_i^* \max_{i=1,\ldots,m} \|a_i\|_2^2 \\ &= \|a_i\|_2^2 \\ \implies \text{SDPdisc}(A) &\leq \max_{i=1,\ldots,m} \|a_i\|_2^2 \, . \end{aligned}$$

To prove that $\text{SDPdisc}(A) \leq \max_{j=1,\ldots,n} \|v_j\|_2^2$, we once again plug in $(\vec{\lambda}^*, \vec{\mu}^*)$ to the dual program:

$$\begin{aligned} \sum_{i=1}^m \mu_i^* a_i a_i^\top &\succeq \text{diag}(\lambda_1^*, \ldots, \lambda_n^*) \\ \sum_{i=1}^m \mu_i^* a_i a_i^\top &= A^\top \text{diag}(\mu_1^*, \ldots, \mu_m^*) \\ &= \begin{bmatrix} v_1^\top \\ \vdots \\ v_n^\top \end{bmatrix} \text{diag}(\vec{\mu}^*) \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix} \\ &= \begin{bmatrix} v_1^\top \text{diag}(\vec{\mu}^*) v_1 & & * \\ & \ddots & \\ * & & v_n^\top \text{diag}(\vec{\mu}^*) v_n \end{bmatrix}. \end{aligned}$$

Since we have that $\mathrm{tr}\left(\sum_{i=1}^m \mu_i^* a_i a_i^\top\right) \geq \sum_{j=1}^n \lambda_j^*$, we then have

$$\mathrm{tr}\left(\begin{bmatrix} v_1^\top \mathrm{diag}(\vec{\mu}^*) v_1 & & * \\ & \ddots & \\ * & & v_n^\top \mathrm{diag}(\vec{\mu}^*) v_n \end{bmatrix}\right) \geq \sum_{j=1}^n \lambda_j^*$$
$$\iff \sum_{j=1}^n v_j^\top \mathrm{diag}(\vec{\mu}^*) v_j \geq \sum_{j=1}^n \lambda_j^*.$$

To upper bound the value on the left hand side, we consider the following linear program:

$$\begin{aligned} &\max \quad \sum_{j=1}^n v_j^\top \mathrm{diag}(\mu_1, \ldots, \mu_m) v_j \\ &\text{subject to} \\ &\sum_{i=1}^m \mu_i = 1, \;\; \mu_i \geq 0, i = 1, \ldots, m. \end{aligned}$$

Observe that the optimal value to this program is an upper bound for $\sum_{j=1}^n v_j^\top \mathrm{diag}(\vec{\mu}^*) v_j$. Since this is a linear (hence convex) maximization over a bounded polytope, its optimal value has a vertex solution, i.e. some $\mu_k = 1$ and $\mu_j = 0$ for other indices. Quick observation tells us the μ_k we should grant the value of 1 is the index k such that $v_k^\top v_k = \|v_k\|_2^2 = \max_{j=1,\ldots,n} \|v_j\|_2^2$. Therefore,

$$\max_{j=1,\ldots,n} \|v_j\|_2^2 \geq \sum_{j=1}^n v_j^\top \mathrm{diag}(\vec{\mu}^*) v_j \geq \sum_{j=1}^n \lambda_j^* \geq \mathrm{SDPdisc}(A).$$

This completes the proof of the theorem. ■

3.3 An Excursion: Randomized Algorithm via Goemans-Williamson

The goal here is to come up with a randomized rounding algorithm to get $x \in \{\pm 1\}^n$ for the set discrepancy problem via the semidefinite relaxation, where the randomized algorithm is inspired by the scheme introduced by Goemans and Williamson in their seminal paper on approximation algorithms through semidefinite relaxation [5]. We will later see at multiple points that the naive implementation of machinery from Goemans-Williamson only admits non-negative matrices A. Fortunately, this is fine for the original set discrepancy problem, where A is a 0/1 matrix.

The algorithm is very simple: the core lies in the simple rounding scheme: say that in solving the primal semi-definite program from above to optimality we get matrix X^* as an optimal solution. We then treat it as a covariance matrix and sample

$$\xi \sim N(0, X^*).$$

We then round ξ in the following manner to extract a ± 1 vector $x(\xi)$:

$$x(\xi)_i = \begin{cases} +1 & \text{if } \xi(i) > 0 \\ -1 & \text{if } \xi(i) < 0. \end{cases}$$

It can be shown that

$$\mathbb{E}[x(\xi)_i x(\xi)_j] = \frac{2}{\pi} \arcsin(X^*_{ij}).$$

We now introduce three lemmas, two of which are elementary, and the third coming from Goemans and Williamson.

Lemma 3.8 (Hoeffding's Lemma) *Suppose $Z \in [a, b]$. Then*

$$\Pr[Z \geq \mathbb{E}[Z] + t] \leq e^{\frac{-2t^2}{(b-a)^2}}.$$

Lemma 3.9 (Union (intersection) Bound) *Given random variables $Z_1, \ldots, Z_n$,*

$$\Pr\left[\bigcap_{i=1}^{n} Z_i\right] \geq 1 - \sum_{i=1}^{n} \Pr[Z_i^c],$$

where we intentionally denote Z_i^c as the complement of Z_i.

Lemma 3.10 *Suppose $x \in [-1, +1)$, then*

$$\frac{2}{\pi} \arcsin(x) \leq 1 - \alpha + \alpha x,$$

where $\alpha = 0.878\ldots$

Let us now define $Z_i = \left(a_i^\top x(\xi)\right)^2$, where $a_i^\top$ is the i-th row of the incidence matrix A. Observe that since $x(\xi) \in \{-1, +1\}^n$, Z_i is bounded: $Z_i \in [0, \|a_i\|_1^2]$. Furthermore, if we compute the expectation of Z_i, we get something interesting

$$\begin{aligned}
\mathbb{E}[Z_i] &= \mathbb{E}[a_i^\top x(\xi) x(\xi)^\top a_i] \\
&= a_i^\top \mathbb{E}[x(\xi) x(\xi)^\top] a_i \\
&= a_i^\top \begin{bmatrix} \frac{2}{\pi} \arcsin(X^*_{11}) & \cdots & \frac{2}{\pi} \arcsin(X^*_{1n}) \\ \vdots & \ddots & \vdots \\ \frac{2}{\pi} \arcsin(X^*_{n1}) & \cdots & \frac{2}{\pi} \arcsin(X^*_{nn}) \end{bmatrix} a_i \\
&= \frac{2}{\pi} a_i^\top \arcsin(X^*) a_i,
\end{aligned}$$

where $\arcsin(X^*)$ denotes arcsin applied element-wise to the matrix X^*.

Now applying Hoeffding's Lemma, we get

$$\Pr[Z_i \geq \mathbb{E}[Z_i] + t] = \Pr[Z_i \geq \frac{2}{\pi} a_i^\top \arcsin(X^*) a_i + t] \leq e^{\frac{-2t^2}{\|a_i\|_1^4}}.$$

Using the upper bound on arcsin, we get

$$\frac{2}{\pi} \arcsin(X^*) \leq (1 - \alpha) \mathbb{1}\mathbb{1}^\top + \alpha X^*,$$

which implies

$$\frac{2}{\pi} a_i^\top \arcsin(X^*) a_i \leq (1 - \alpha)(a_i^\top \mathbb{1})^2 + \alpha a_i^\top X^* a_i.$$

Here is where we make the assumption that A is a *positive matrix*, in which case $(a_i^\top \mathbb{1})^2 = \|a_i\|_1^2$. Returning to the probability bound, we now have

$$\begin{aligned}
&\Pr[Z_i \geq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + t] \\
\leq\ &\Pr[Z_i \geq \frac{2}{\pi} a_i^\top \arcsin(X^*) a_i + t] \\
\leq\ &e^{\frac{-2t^2}{\|a_i\|_1^4}}.
\end{aligned}$$

We now appeal to Lemma 3.2 to bound the intersection of probabilities across rows $a_i^\top$:

$$\begin{aligned}
&\Pr\left[\bigcap_{i=1}^m \left(Z_i \leq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + t\right)\right] \\
\geq\ &1 - \sum_{i=1}^m \Pr[Z_i \geq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + t] \\
\geq\ &1 - \sum_{i=1}^m e^{\frac{-2t^2}{\|a_i\|_1^4}}
\end{aligned}$$

Let us set $c = \max_{i=1,\ldots,m} \|a_i\|_1$ and $t = \frac{\sqrt{2}}{2} c^2 \sqrt{d}$:

$$\begin{aligned}
\Pr\left[\bigcap_{i=1}^m \left(Z_i \leq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + \frac{\sqrt{2}}{2} c^2 \sqrt{d}\right)\right] &\geq 1 - \sum_{i=1}^m e^{\frac{-c^4 d}{c^4}} \\
&= 1 - m e^{-d}.
\end{aligned}$$

For the final step, we set $d = \log(2m)$ to get

$$\Pr\left[\bigcap_{i=1}^m \left(Z_i \leq (1-\alpha)\|a_i\|_1^2 + \alpha a_i^\top X^* a_i + \frac{\sqrt{2}}{2} c^2 \sqrt{d}\right)\right] \geq 1/2.$$

Recall that $Z_i := (a_i^\top x(\xi))^2$, i.e. the squared net "charge" of the i-th row. Furthermore, recall the primal semi-definite program

$$\begin{aligned}
&\min \quad t \\
&\text{subject to} \\
&t \geq a_i^\top X a_i,\ i = 1, \ldots, m \\
&X_{jj} = 1,\ j = 1, \ldots, n \\
&X \succeq 0.
\end{aligned}$$

Observe that by the constraints of the program, each $a_i^\top X^* a_i \leq \mathrm{SDPdisc}(A)$. Therefore, for positive matrix A, the randomized algorithm will return $x(\xi)$ that attains a discrepancy no more than

$$\begin{aligned}
&\sqrt{(1-\alpha) \max_{i=1,\ldots,m} \|a_i\|_1^2 + \alpha\, \mathrm{SDPdisc}\,(A)^2 + \frac{\sqrt{2}}{2} \max_{i=1,\ldots,m} \|a_i\|_1^2 \sqrt{\log(2m)}} \\
= &\sqrt{\left(1 - \alpha + \frac{\sqrt{2}}{2}\sqrt{\log(2m)}\right) \max_{i=1,\ldots,m} \|a_i\|_1^2 + \alpha\, \mathrm{SDPdisc}\,(A)^2}
\end{aligned}$$

with probability at least 1/2.

4 Generalizations to Sums of Matrices

4.1 Matrix Komlós & Spencer and SDP Relaxation

Recall in the Komlós problem, we consider a square matrix $A \in M_n$ where its columns are at most unit length. Let us consider the following generalization of the Komlós problem. Consider a collection of matrices $\mathcal{A} = \{A_i\}_{i=1}^n$, $A_k \in M_n$. We define the matrix discrepancy (Mdisc) of $\mathcal{A}$ to be

$$\text{Mdisc}(\mathcal{A}) = \min_{x \in \{\pm 1\}^n} \left\| \sum_{i=1}^n x_i A_i \right\|_{op},$$

where $\|\cdot\|_{op}$ is the spectral norm, i.e. largest singular value. We now pose the matrix equivalent of the Komlós conjecture:

Conjecture 4.1 (Matrix Komlós) *Let $\mathcal{A} = \{A_i\}_{i=1}^n$ be a collection of symmetric matrices $A_i \in M_n$ where each A_i has Frobenius norm at most 1: $\|A_i\|_F \leq 1$. Then,*

$$Mdisc(\mathcal{A}) \leq K,$$

where K is a constant independent of n.

This is a strict generalization of the Komlós conjecture by the following embedding: given $A \in M_n$ where its columns have Euclidean norm at most 1, put each column into a diagonal matrix. The Frobenius norm of each diagonal matrix is clearly bounded by 1, and the operator norm of a diagonal matrix is the largest diagonal entry in absolute value, i.e. $\|\cdot\|_\infty$ on the diagonal.

In a similar vein, Spencer's Problem also admits a matrix generalization:

Conjecture 4.2 (Matrix Spencer) *Let $\mathcal{A} = \{A_i\}_{i=1}^n$ be a collection of symmetric matrices $A_i \in M_n$ where each A_i satisfies $\|A_i\|_{op} \leq 1$ (bounded in spectral norm). Then,*

$$Mdisc(\mathcal{A}) = O(\sqrt{n}).$$

The embedding of Spencer's Theorem into Matrix Spencer is identical to that of Komlós into Matrix Komlós. We remark that while Matrix Komlós seems to be a novel problem, Matrix Spencer has surfaced recently in discrepancy theory literature [9, 12]. Specifically in [12], a deterministic algorithm is proposed that gets discrepancy with the following bound

$$\left\| \sum_{i=1}^n x_i A_i \right\|_{op} \leq O\left(\sqrt{n \log(2q)}\right),$$

where q is the dimension of the largest block in any A_k. At the highest level of generality, $q = n$, and the bound recovers the random coloring bound [9].

What is specifically nice about the Matrix Komlós problem is that, similar to what we saw in Section 3.2, we can find its semi-definite relaxation. Let us reformulate the problem into a mathematical program

$$\begin{aligned} \min \quad & \left\| \sum_{i=1}^n x_i A_i \right\|_{op} \\ \text{s.t. } & x_i \in \{\pm 1\}, \quad i = 1, \ldots, n. \end{aligned}$$

which can be equivalently written

$$\begin{aligned}
&\min t \\
&\text{s.t. } tI_n \succcurlyeq \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}\left(A_i A_j^\top + A_j A_i^\top\right) \\
&\qquad x_i^2 = 1, \quad i = 1, \ldots, n.
\end{aligned}$$

We propose the following (primal) semidefinite relaxation to the above:

$$\begin{aligned}
(P) \quad &\min t \\
&\text{s.t. } tI_n \succcurlyeq \sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j\left(A_i A_j^\top + A_j A_i^\top\right) \\
&\qquad X = [X_{ij} = x_i x_j]_{n\times n} \succcurlyeq 0 \\
&\qquad X_{ii} = 1, \quad i = 1, \ldots, m.
\end{aligned}$$

whose dual program is

$$\begin{aligned}
(D) \quad &\max \sum_{i=1}^{n} y_i \\
&\text{s.t. } \operatorname{diag}(y) \preccurlyeq \left[\operatorname{tr}\left(A_i^\top Z A_j\right) + \operatorname{tr}\left(A_j^\top Z A_i\right)\right]_{n\times n} \\
&\qquad Z \succcurlyeq 0 \\
&\qquad \operatorname{tr}(Z) = 1.
\end{aligned}$$

It is similarly easy to verify that (P) and (D) both satisfy Slater's condition and thus the optimal values of both programs are bounded and coincide. Observe that when A_i are diagonal matrices corresponding to inputs of the original Komlós conjecture, then the above two programs reduce to the semidefinite relaxations introduced in Section 2.2.

4.2 An Asymmetric Random Coloring Lemma

It is known that the equivalent of the random coloring lemma (Theorem 2.4) is true for the Matrix Spencer problem [9]:

Theorem 4.3 ((Symmetric) Random Coloring Lemma) *For any symmetric matrices $A_1, \ldots, A_n \in M_n$, then*

$$\mathbb{P}_{x\in\{\pm 1\}^n}\left[\left\|\sum_{i=1}^{n} x_i A_i\right\|_{op} \geq t \left\|\sum_{i=1}^{n} A_i^2\right\|_{op}^{1/2}\right] \leq 2ne^{-t^2/2}.$$

In particular, if $\|A_i\|_{op} \leq 1$, then we have that for random $x \in \{\pm 1\}^n$, then with probability at least $1/2$ we get

$$\left\|\sum_{i=1}^{n} x_i A_i\right\|_{op} \leq O\left(\sqrt{n\log(n)}\right).$$

The above is a standard random matrix theory result and the proof comes from a Matrix Chernoff bound, analogous to the Chernoff bound used to prove the original random coloring lemma [17, 9].

Key to the proof is the fact that A_i are symmetric−in fact, that is how we are able to go from the first assertion to second. The surprising fact that we prove in this section is the following: symmetricity is not crucial to get the $O\left(\sqrt{n\log(n)}\right)$ bound. We use some less well-known moment inequalities [13], most notably the non-commutative Khintchine inequality, to establish this fact.

Theorem 4.4 (Asymmetric Random Coloring Lemma) *For matrices $A_1, \ldots, A_k$, where $A_i \in M_{m,n}$ and $k \leq \max(m,n)$, given a random coloring $x \in \{\pm 1\}^k$ and any $\beta \geq 1/2$, we have*

$$\mathbb{P}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{op} \geq \sqrt{2e(1+\beta)\max(m,n)\ln(\max(m,n))}\right] \leq (\max(m,n))^{-\beta}.$$

Note that setting $m = n$ and $\beta = 1/2$ gets us an $O\left(\sqrt{n\log(n)}\right)$ bound with good probability.

Proof of Theorem 4.4: Observe that since $\|A_i\|_{op} \leq 1$ and $k \leq \max(m,n)$, we have that

$$\sum_{i=1}^{k} A_i A_i^\top \preccurlyeq k I_m \preccurlyeq \max(m,n) I_m$$
$$\sum_{i=1}^{k} A_i^\top A_i \preccurlyeq k I_n \preccurlyeq \max(m,n) I_n.$$

We use the following variant of the non-commutative Khintchine inequality

Theorem 4.5 (NCKI, Lust-Piquard) *Let $x_1, \ldots, x_k$ be independent Bernoulli random variables, and let $A_1, \ldots, A_k$ be a sequence of arbitrary $m \times n$ matrices. Then given $p \geq 2$, there exists an absolute constant $\gamma_p > 0$ such that*

$$\mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{S_p}^{p}\right] \leq \gamma_p \max\left\{\left\|\left(\sum_{i=1}^{k} A_i A_i^\top\right)^{1/2}\right\|_{S_p}^{p}, \left\|\left(\sum_{i=1}^{k} A_i^\top A_i\right)^{1/2}\right\|_{S_p}^{p}\right\},$$

where S_p denotes the Schatten p-norm, i.e. the vector p-norm evaluated on the singular values, and $\gamma_p < p^{p/2}$.

We observe that

$$\left\|\left(\sum_{i=1}^{k} A_i A_i^\top\right)^{1/2}\right\|_{S_p}^{p} \leq m\max(m,n)$$
$$\left\|\left(\sum_{i=1}^{k} A_i^\top A_i\right)^{1/2}\right\|_{S_p}^{p} \leq n\max(m,n).$$

Therefore, using the fact $\|X\|_{op} \leq \|X\|_{S_p}$, Theorem 4.5 tells us

$$\mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{op}^{p}\right] \leq \mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{S_p}^{p}\right] \leq \gamma_p \max(m,n)^2 < p^{p/2}\max(m,n)^2.$$

Applying Markov's Inequality, for any $t > 0$ and $p \geq 2$, we have

$$\mathbb{P}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{op} \geq t\right] \leq \frac{1}{t^p}\mathbb{E}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{S_p}^p\right]$$
$$\leq \frac{p^{p/2}\max(m,n)^2}{t^p}.$$

Setting $t = \sqrt{2e(1+\beta)\max(m,n)\ln(\max(m,n))}$ and $p = \frac{t^2}{e\max(m,n)} > 2$ (since $\beta \geq 1/2$), we get

$$\mathbb{P}\left[\left\|\sum_{i=1}^{k} x_i A_i\right\|_{op} \geq \sqrt{2e(1+\beta)\max(m,n)\ln(\max(m,n))}\right] \leq (\max(m,n))^{-\beta},$$

which concludes the proof of Theorem 4.4. ■

4.3 Remarks and Conjectures

Remark 4.6 *One of the powerful motivations for working on the Matrix Komlós conjecture is the recent burst of matrix concentration inequalities originating from Marcus et al.'s proof of the Kadison-Singer problem [16, 7]. These concentration inequalities can often be translated into a Matrix Komlós setting except with additional rank assumptions. In the Marcus et al.'s paper, rank* 1 *matrices are considered.*

Conjecture 4.7 *The following conjecture is supported by numerical optimization. Given a dimension n, say we have a matrix A with columns a_j having size at most* 1 *that attains the worst possible discrepancy for that dimension. Then, we embed the columns of A into diagonal matrices. The claim here is that the discrepancy of that system of diagonal matrices cannot be improved by orthogonal transformations. In other words, given $\{diag(a_j)\}$ and orthogonal matrices $Q_1, Q_2 \ldots,$ then*

$$\left\|\sum_j diag(a_j)\right\|_{op} \geq \left\|\sum_j Q_j\, diag(a_j) Q_j^\top\right\|_{op}.$$

This would imply that worst-case systems in vector Komlós cannot be worsened in the matrix setting through orthogonal perturbations. Numerically, this seems to hold true, and is emphatically not true even for "bad" case Komlós inputs such as a scaled Hadamard matrix.

Conjecture 4.8 *Beyond the counterexamples inherited from vector Komlós, it is unclear to me when else SDPMdisc can be* 0 *while Mdisc is greater than* 0*. This leads me to believe that SDPMdisc is usually (almost everywhere) a decent approximation of Mdisc apart from when the collection contains only diagonal matrices (which is a null set with respect to the Lebesgue measure on M_n). Numerical optimization supports this conjecture.*

Conjecture 4.9 *With the generalized asymmetric random coloring lemma, it may be possible to refine the classic proof of Spencer's Theorem by entropy [8] to generalize to Matrix Spencer. In the entropy proof, the random coloring lemma is used as a way to quantify a standard deviation in the distribution of all possible colorings and the discrepancy they incur. The proof then works by showing that there must be non-zero mass at $O(\sqrt{n})$. If Matrix Spencer enjoys the same "standard deviations", it might suggest that similarly there is mass at $O(\sqrt{n})$. I unfortunately got stuck trying to go down this route. It would also be very interesting if one could in fact prove that in general there is zero mass at $O(\sqrt{n})$.*

References

[1] Wojciech Banaszczyk. Balancing vectors and gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

[2] Nikhil Bansal. Constructive algorithms for discrepancy minimization. *In Foundations of Computer Science, 51st Annual IEEE Symposium*, pages 3–10, 2010.

[3] Nikhil Bansal. Algorithmic aspects of combinatorial discrepancy. In Anand Srivastav William Chen and Giancarlo Travaglini, editors, *A Panorama of Discrepancy Theory*, chapter 6, pages 425–457. Springer-Verlag Berlin Heidelberg, 2014.

[4] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[5] M.X. Goemans and D.P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the Association for Computing Machinery*, 42(6):1115–1145, 1995.

[6] Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. *SIAM Journal on Computing*, 44(5):1573–1582, 2015.

[7] A. W. Marcus and N. Srivastava. The solution of the kadison-singer problem, 2017.

[8] Jiřì Matoušek. *Geometric Discrepancy: An Illustrated Guide*. Springer-Verlag Berlin Heidelberg, 1 edition, 1999.

[9] Raghu Meka. Discrepancy and beating the union bound, 2014.

[10] M. Charikar, A. Newman and A. Nikolov. Tight hardness results for minimizing discrepancy. In *SODA*, 2011.

[11] Aleksandar Nikolov. The komlós conjecture holds for vector colorings, 2013.

[12] A. Levy, H. Ramadas and T. Rothvoss. Deterministic discrepancy minimization via the multiplicative weight update method. In *IPCO*, 2017.

[13] Anthony Man-Cho So. Moment inequalities for sums of random matrices and their applications in optimization. *Mathematical Programming*, 130(1):125–151, 2011.

[14] Joel Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985.

[15] Joel Spencer. Balancing vectors in the max norm. *Combinatorica*, 6:55–65, 1986.

[16] A. W. Marcus, D.A. Spielman and N. Srivastava. Interlacing families ii: Mixed characteristic polynomials and the kadison-singer problem. *Annals of Mathematics*, pages 327–350, 2015.

[17] Joel A. Tropp. An introduction to matrix concetration inequalities, 2014.