Lecture Constructive Algorithms for Discrepancy Minimization

Nikhil Bansal *Scribe: Daniel Dadush*

1 Combinatorial Discrepancy

Given a universe $U = \{1, \ldots, n\}$ and a set system $\mathcal{S} = \{S_1, \ldots, S_m\}$ on U, we define the *discrepancy* of the set system S as

$$\operatorname{disc}(\mathcal{S}) = \min_{\chi: U \to \{-1,1\}} \max_{S \in \mathcal{S}} |\chi(S)|$$

where $\chi(S) = \sum_{i \in S} \chi(i)$. Here we want to color the elements of U such that each set $S_1, \ldots, S_m$ is colored as evenly as possible, i.e. we want nearly as many 1s and -1s in each set.

The *Minimum Discrepancy Problem* for the set system $\mathcal{S}$ is to find a coloring $\chi : U \to \{-1, 1\}$ which minimizes discrepancy, i.e. the function

$$\max_{S \in \mathcal{S}} \left| \sum_{i \in S} \chi(i) \right| .$$

We will be interested in polytime algorithms for computing these low discrepancy colorings.

In terms of applications, the min discrepancy problem appears in many varied areas of both Computer Science (Computational Geometry, Comb. Optimization, Monte-Carlo simulation, Machine learning, Complexity, Pseudo-Randomness) and Mathematics (Dynamical Systems, Combinatorics, Mathematical Finance, Number Theory, Ramsey Theory, Algebra, Measure Theory,...). One may consult any of the following books [Cha01, AS00, Mat10] for an in depth view of the subject.

1.1 Discrepancy Bounds for General Set Systems

In this section, we will be interested in computing upper bounds for the discrepancy of an arbitrary set system $\mathcal{S}$ based solely on n and m ($U = [n]$ and $|S| = m$). For simplicity, we will mostly examine the case where $m = n$.

The following lemma gives us a simple bound on $\operatorname{disc}(S)$.

Lemma 1 (Random Coloring)**.** $\operatorname{disc}(S) = O(\sqrt{n \log n})$

Proof. Let $\chi(i)$ denote a uniform $\{-1, 1\}$ random variable for each $i \in [n]$. Now note that for each $S \in \mathcal{S}$, the variance of $\chi(S) = \sum_{i \in S} \chi(i)$ is exactly $|S| \leq n$. Therefore by the Chernoff bounds, we get that

$$\Pr(|\chi(S)| \geq c\sqrt{n}) < 2e^{-\frac{1}{4}c^2}$$

for $c \leq \sqrt{n}$. Let $c = \sqrt{4 \log(4n)}$, we get that

$$\Pr\left(|\chi(S)| \geq \left(\sqrt{4 \log(4n)}\right) \sqrt{n}\right) < \frac{1}{2n}$$

Applying the union bound on the n sets in S, we get that a random coloring achieves discrepancy $\sqrt{4n\log(4n)}$ with probability at least $\frac{1}{2}$. Hence $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n\log n})$ as claimed. $\square$

We note that the above lemma immediately implies an algorithm for the discrepancy problem, i.e. that of choosing a random coloring on U. Until recently, this was indeed the best known algorithm for coloring a general set system. Furthermore, the above analysis of the random coloring technique is tight, i.e. there are examples where a random coloring yields discrepancy $\Theta(\sqrt{n\log n})$ with very high probability.

Even though random colorings gave the best algorithmic technique, a beautiful theorem of Spencer [Spe85] shows that the discrepancy of general set systems is in fact always better than $O(\sqrt{n\log n})$.

Theorem 2 (Spencer)**.** $\mathrm{disc}(\mathcal{S}) \leq 6\sqrt{n}$

For general set systems, i.e. $|\mathcal{S}| = m$, Spencer shows that $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n\log\frac{m}{n}})$. Furthermore, for $m = n$, the theorem is tight, since one can show that Hadamard set systems have discrepancy $.5\sqrt{n}$ [Cha01].

Unfortunately, Spencer's proof is inherently non-constructive (uses pigeon hole principle on exponentially many elements) and hence cannot directly be made into an algorithm.

Challenge: Can we find a Spencer type coloring *algorithmically*?

On this front, Spencer proved that no non-adaptive or online algorithm can achieve better that $\sqrt{n\log m}$ (see for e.g. [AS00]). Given these bounds, and the non-constructive nature of Spencer's proof, the following has been conjectured [AS00]:

Conjecture: No efficient algorithm to find such a coloring exists.

1.2 Approximating Discrepancy

Question: If $\mathcal{S}$ has low discrepancy (say $\mathrm{disc}(\mathcal{S}) << \sqrt{n}$), can we find a good coloring?

Unfortunately, the answer to this question was shown to be strongly negative by Charikar, Newman, and Nikolov [CNN11]:

Theorem 3 (Hardness)**.** *It is NP-Hard to distinguish between* $\mathrm{disc}(\mathcal{S}) = 0$ *and* $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n})$*, even for systems with* $m = O(n)$*.*

The above result suggests that $\mathrm{disc}(\mathcal{S})$ maybe too strong a quantity to measure directly. One may ask, if there is not a more "robust" measure of the discrepancy of $\mathcal{S}$ which would be amenable to approximation. In particular, we may examine the Hereditary Discrepancy of $\mathcal{S}$, which is defined as

$$\mathrm{herdisc}(\mathcal{S}) = \max_{U' \subseteq U} \mathrm{disc}(\mathcal{S}_{|U'})$$

where $\mathcal{S}_{|U'} = \{S \cap U' : S \in \mathcal{S}, S \cap U' \neq \emptyset\}$ is the restriction of the $\mathcal{S}$ to the subuniverse $U' \subseteq U$. The notion of Hereditary Discrepancy indeed appears quite naturally in many settings, for example in the definition of totally unimodular matrices and certain geometric contexts.

Question: (Matousek) Can we find a low discrepancy coloring of $\mathcal{S}$ given that herdisc(S) is low?

2 Results

We present our algorithmic results here.

Theorem 4 (Hereditary Discrepancy)**.** *Given a set system $\mathcal{S}$, $|\mathcal{S}| = m$ on a universe $U = [n]$, a coloring $\chi : U \to \{-1, 1\}$ can be computed in polytime achieving discrepancy at most $O(\log(mn))\text{herdisc}(\mathcal{S})$, or more precisely $O(\sqrt{\log m \log n})\text{herdisc}(\mathcal{S})$.*

Building upon the techniques of this theorem, with some additional ideas, we can also show that

Theorem 5 (Constructive Spencer)**.** *Given a set system $\mathcal{S}$, $|\mathcal{S}| = n$ on a universe $U = [n]$, a coloring $\chi : U \to \{-1, 1\}$ can be computed polytime achieving discrepancy at most $O(\sqrt{n})$.*

This effectively makes Spencer's theorem constructive.

The general techniques developed for these algorithms can also be used to get constructive bounds for various other discrepancy problems such as those arising in geometric contexts, and to build low discrepancy colorings in both the k-permutation (set system is defined by all prefixes from a collection of k permutations of U) and beck-fiala setting (each element of $u \in U$ appears in at most t sets).

3 Algorithm

We first write down two convex relaxations for Spencer's problem. Given the set system $\mathcal{S}$, $|\mathcal{S}| = n$ on $U = [n]$, we first examine the following linear program

$$\begin{aligned} \sqrt{n} \leq \sum_{i \in S} x_i^+ - \sum_{i \in S} x_i^- \leq \sqrt{n} \quad \forall S \in \mathcal{S} \\ x_i^+ + x_i^- = 1 \quad \forall i \in [n] \\ x_i^+, x_i^- \geq 0 \quad \forall i \in [n] \end{aligned}$$

Here the intent is for the LP to set $x_i^+ = 1$ if it wishes to color i to 1 and set $x_i^- = -1$ if it wants to color i to -1. Unfortunately, the above LP is completely useless, since no matter the set system it will set $x_i^+ = x_i^- = \frac{1}{2}$, achieving a discrepancy of 0 for each set.

To achieve a more meaningful convex program, we examine the following SDP relaxation. We may interpret the following SDP as coloring each element of U with vectors as opposed to -1 or 1.

$$\begin{aligned} \|\sum_{i \in S} v_i\|^2 \leq n \quad \forall S \in \mathcal{S} \\ \|v_i\|^2 = 1 \quad \forall i \in [n] \\ v_i \in \mathbb{R}^n \end{aligned}$$

Here the intended solution is $v_i = (-1, 0, \dots, 0)$ or $v_i = (1, 0, \dots, 0)$ if element i is colored -1 or 1 respectively. Here again, unfortunately, irrespective of the set system $\mathcal{S}$, the solution $v_i = e_i$ (the i^{th} unit vector) satisfies each of the above constraints at equality. Even though, the above example may make this SDP seem useless, modifications of the above program will form the central component of our discrepancy algorithm.

The key point is that the trivial solution no longer becomes viable if we ask for slightly tighter discrepancy restrictions on some of the sets. I.e. if we ask that

$$\| \sum_{i \in S'} v_i \|^2 \leq \frac{n}{\log n}$$

for certain sets $S' \subseteq \mathcal{S}$, we force the SDP to give non-trivial information on the sets. Of course, the SDP may not be feasible if we set the discrepancy constraints too tight, so this will need to be handled carefully.

In the sections that follow, we will first discuss the high level approach for the SDP based algorithm. Then we will show how to apply this algorithm in the Hereditary Discrepancy setting, and subsequently the additional ideas needed to make it work in Spencer's setting.

3.1 High Level Description

To find a small discrepancy coloring χ of $[n]$ for $\mathcal{S}$, our algorithm will perform a random walk on the coordinates of χ, whose movement is seeded by a sequence of SDPs. Starting from an initial "fractional" coloring $\chi = (0, \dots, 0)$, we proceed as follows:

- At timestep t, use an SDP to generate a small random pertubation δ_t, and let $\chi \leftarrow \delta_t + \chi$.
- Freeze all the coordinates of χ which get pushed above 1 or below -1. Repeat.

Once all the coordinates of χ are frozen, we round each component of χ to the its nearest value in $-1, 1$ and return the associated vector. We will make sure the random walk increments are small enough so that the rounding cost (in terms of accumulated discrepancy) is negligible. The above process can therefore be thought of as a "sticky" random walk in $[-1, 1]^n$, which sticks to the faces of the cube it hits during its execution.

For the above approach to work, we will attempt to ensure the random perturbations induce a large variance random walk on the coordinates of χ, i.e. $\chi(i)\ \forall i \in [n]$, but a low variance random walk on the set discrepancies, i.e. $\sum_{i \in S} \chi(i)$. These two properties together will allow us to argue that many coordinates get frozen while keeping the set discrepancies low.

3.2 SDP / Random Walk for Hereditary Discrepancy

Examining the Hereditary Discrepancy setting, we are given a set system $\mathcal{S}$ on $[n]$ and a guarantee that $\text{herdisc}(\mathcal{S}) \leq \lambda$. We wish to construct a coloring χ whose discrepancy w.r.t. $\mathcal{S}$ is bounded by $O(\log(mn)\lambda)$.

Given the bound on the hereditary discrepancy, we have that for any choice of $U' \subseteq [n]$ the following SDP is feasible

$$\begin{aligned} \|\sum_{i \in S} v_i\|^2 &\leq \lambda^2 \quad \forall S \in \mathcal{S}_{|U'} \\ \|v_i\|^2 &= 1 \quad \forall i \in U' \\ v_i &\in \mathbb{R}^{|U'|} \end{aligned} \tag{1}$$

Therefore, no matter which coordinates we freeze in our random walk, the above SDP will always be able to give us useful information about the remaining free coordinates.

The random walk we will use works as follows. Let χ denote the current fractional coloring, and for notational simplicity let us assume that the set of free coordinates is $[n]$.

1. Let $v_1, \ldots, v_n$ denote a solution to the above SDP ($U' = [n]$).
2. Let $g = (g_1, \ldots, g_n)$ denote a random gaussian vector in $\mathbb{R}^n$, where each g_i are i.i.d. N(0,1).
3. Let $\eta_i = v_i \cdot g$, $\forall i \in [n]$.
4. Update χ by setting
$$\chi(i) = \chi(i) + \gamma \eta_i,$$
where $\gamma > 0$ is a positive parameter (we want small increments, but $\gamma = \frac{1}{n^2}$ will easily suffice).
5. Freeze all coordinates of χ where $|\chi(i)| \geq 1$.

To apply the above update when some of the coordinates of χ are frozen, we simply restrict both the SDP solution and the random coordinate updates to the set of free coordinates.

Properties of the Random Walk:

The following lemma recalls the fundamental property of the gaussian.

Lemma 6. *For a gaussian random vector $g = (g_1, \ldots, g_n) \in \mathbb{R}^n$, where each g_i is iid $N(0,1)$, and $v \in \mathbb{R}^n$ then $v \cdot g$ is distributed as $N(0, \|v\|^2)$.*

Proof. Simply note that $v \cdot g = \sum_i v_i g_i$, and that sum of independent one dimensional gaussians is another gaussian with both mean and variance equal to sum of the individual means and variances. □

We now analyze the evolution of χ over time. The relevant quantities at timestep t as follows:

- χ^t denotes the fractional coloring at time t, with free variables $U_t \subseteq [n]$.
- v_i^t, $i \in U_t$, denotes the SDP solution, and $g^t \in \mathbb{R}^{|U_t|}$ denotes the gaussian vector generated at time t (g^t is independent of gaussian generated at previous steps).
- For $i \in U_t$, the coordinate updates at time t are $\chi^t(i) = \chi^{t-1}(i) + \gamma \eta_i^t$ where $\eta_i^t = v_i^t \cdot g^t \sim N(0, \|v_i^t\|^2) = N(0,1)$, and $\eta_i^t = 0$ for $i \in [n] \setminus U_t$.

We first note that for $t \geq 1$ that

$$\chi^t(i) = \sum_{i=0}^{t-1} \gamma \eta_i^t \sim N(0, \gamma^2 t)$$

Hence for each i, the color $\chi^t(i)$ performs a random walk of step size $\gamma N(0,1)$ at each timestep (though note that the random walks for the various colors are correlated). Now for $S \in \mathcal{S}$, and for $t \geq 0$, we see that

$$\sum_{i \in S} \eta_i^t \sim N(0, \| \sum_{i \in S \cap U_t} v_i^t \|^2) \tag{2}$$

where $\| \sum_{i \in S \cap U_t} v_i^t \|^2 \leq \lambda^2$ by construction. Hence we get that $\chi^t(S)$ also performs a random walk of step size at most $\gamma N(0, \lambda^2)$ at each timestep (again the walks for various sets may be correlated arbitrarily).

Analysis: Let a round consist of $T = 12\frac{1}{\gamma^2}$ random walk steps. We first show that after one round, we expect a good fraction of the free coordinates of our fractional coloring to be frozen.

Lemma 7. *For a round $k \in \mathbb{N}$, with probability at least $\frac{1}{2}$ at least half the of the free coordinates of $\chi^{T(k-1)}$ are frozen in χ^{Tk}.*

Proof. We only show the statement for $k = 0$; for general k is analysis is basically identical. We examine the probability that $\chi^T(i)$ is not frozen. Let $Y_1, \dots, Y_T$ be iid $N(0, \gamma^2)$ variables, and let $S_t = \sum_{i=1}^t Y_i$. By construction, we can generate the distribution of $\chi^T(i)$ by sampling $Y_1, \dots, Y_T$ independently, returning either S_i if i the first index where S_i jumps out of $(-1, 1)$ or S_T if no such $i \in [T-1]$ exists.

Now note that

$$\Pr(\chi^T(i) \text{ is not frozen }) = \Pr(\cap_{t=0}^T \{|S_t| < 1\}) \leq \Pr(|S_T| < 1)$$

Now S_T is distributed as $N(0, T\gamma^2) = N(0, 12)$, and hence has density $\frac{1}{2\sqrt{6\pi}} e^{-\frac{1}{2}\| \frac{x}{2\sqrt{3}} \|^2}$. Therefore the $\Pr(|S_T| \leq 1) \leq \frac{1}{2\sqrt{6\pi}}(2) = \frac{1}{\sqrt{6\pi}} < \frac{1}{4}$.

Let E_i, $i \in [n]$, be the indicator denoting whether $\chi^T(i)$ is not frozen. The above computation shows that $\Pr(E_i) \leq \frac{1}{4}$. Therefore by Markov's inequality, we have that $\Pr(\sum_{i=1}^n E_i \geq \frac{1}{2}n) \leq \frac{1}{2}$. Hence with probability at least $\frac{1}{2}$ at least $\frac{1}{2}n$ of the coordinates of χ^T are frozen as claimed. $\square$

Remark: We note that the above argument does not quite work for Spencer's setting, since for each element we have the constraint $|v_i| \leq 1$ together with an aggregate constraint that $|\sum_i v_i|^2 \geq n/2$ instead of the constraints $|v_i = 1|$ above. This may cause the variance of some element walks to be small (though on average the walks have large variance at each step). To handle this, we use a more careful argument based on potential functions. This can be found in our paper.

The next lemma helps us bound how much discrepancy we have accumulated in k rounds.

Lemma 8. *Take $S \in \mathcal{S}$. For any round $k \in \mathbb{N}$, and set $S \in \mathcal{S}$, we have that*

$$\Pr(|\chi^{Tk}(S)| \geq 6\sqrt{m}\sqrt{k}\lambda) \leq \frac{1}{m^2}$$

To see why the above is true, we remember that $\chi^t(S)$ performs a random walk of step size roughly $N(0, \gamma^2\lambda^2)$. Since we perform $O(\frac{k}{\gamma^2})$ steps of the random walk, we expect to total length of the walk to look like $N(0, k\lambda^2)$, where the bound we want would now follow from the standard gaussian tail bound. The formal proof uses a standard martingale argument.

We sketch the remainder of the proof.

- *Claim 1:* After $6 \log n$ rounds, all the coordinates are frozen with high probability.

  To see this, let E_i denote the indicator of whether the number of free coordinates dropped by half in the ith round. We call a round "successful" if $E_i = 1$. By Lemma 7 we know that $E[E_i] = \Pr(E_i = 1) \geq \frac{1}{2}$. Since there are only n elements to start with, we can have at most $\log n$ successful rounds before all the coordinates are frozen. Since the rounds are essentially independent, we can use the Chernoff bounds to show that in $6 \log n$ rounds there are at least $\log n$ successful rounds with probability at least $1 - \frac{1}{n^2}$. Hence with probability $1 - \frac{1}{n^2}$, all the coordinates are frozen in at most $6 \log n$ rounds.

- *Claim 2:* After $6 \log n$ rounds, every set $S \in \mathcal{S}$ has discrepancy $O(\sqrt{\log n \log m})$ with high probability.

  By Lemma 8, we have that

  $$\Pr(|\chi^{6 \log n T}(S)| \geq 20\sqrt{\log n \log m}\lambda) \leq \frac{1}{m^2}$$

  for each $S \in \mathcal{S}$. Therefore by the union bound, the probability that any set has discrepancy bigger than $20\sqrt{\log n \log m}\lambda$ after $6 \log n$ rounds is at most $m\frac{1}{m^2} = \frac{1}{m}$. Hence with probability at least $1 - \frac{1}{m}$ all sets have discrepancy $O(\sqrt{\log m \log n}\lambda)$ after $6 \log n$ rounds.

- *Claim 3:* With high probability, we output a valid coloring of U of discrepancy $O(\sqrt{\log m \log n}\lambda)$ after $6 \log n$ steps.

  The claim follows by combining claim 1 and 2 and applying the union bound. The only thing we did not discuss is the error in discrepancy induced by rounding the fractional solution to $\{-1, 1\}$. Here, since we set the step size to $\gamma = \frac{1}{n^2}$, the probability that any step of the random walk is larger than $\frac{1}{n}$ is exponentially small in n. Hence with overwhelming probability, we never have to update any coordinate of χ by more than $\frac{1}{n}$, and hence the error induced by rounding is at most additive $O(1)$, i.e. negligible since $\lambda \geq 1$.

Recap: In this section, we used Hereditary Discrepancy to guarantee that a specific SDP was feasible at each step. With this SDP, we generated a random walk which satisfied the following two properties:

1. *High Variance* on Coordinates: guaranteed *Rapid Convergence* of the algorithm.
2. *Low Variance* on Sets: guaranteed *Low Discrepancy* of the final solution.

In the next section, we will discuss what is needed to generalize this approach to work in Spencer's setting.

4 Spencer's Setting

Given a set system $\mathcal{S}$, $|S| = n$, on $[n]$, our goal is to find a coloring χ of discrepancy $O(\sqrt{n})$.

If we try to adapt the previous approach in this setting, we have to first understand what SDP to use. Previously, we were able to use the bound on Hereditary discrepancy to guarantee the feasibility of our SDPs. Now the naive approach would be to simply plug in an upper bound of $\sqrt{n}$ on the Hereditary discrepancy. This clearly performs poorly, since it only achieves discrepancy $O(\sqrt{n}\sqrt{\log m \log n})$ which is even worse than a random coloring. The next idea would be to tighten the bounds of the SDP used in the previous section to those predicted by Spencer's theorem itself. I.e. if we have k free variables and n sets left at time t, we may use the SDP from (1) with a right hand side of $O(k \log \frac{n}{k})$, which is guaranteed to be feasible by Spencer's theorem. Unfortunately, even with this strengthening, we can show that the coloring produced only has discrepancy $O(\sqrt{n}\sqrt{\log m})$ (the main issue is the union bound over all sets) which again is no better than random.

4.1 Partial Colorings

To find the right SDP, we will have to examine the details of Spencer's proof. The main crux of his proof is the following partial coloring lemma.

Lemma 9. *For any set system $\mathcal{S}$, $|\mathcal{S}| = m$, on $[n]$, there exists a coloring $\chi : \{-1, 0, 1\} \to [n]$ on $\frac{n}{2}$ elements (i.e. $|\{i : \chi(i) \neq 0\}| \geq \frac{n}{2}|$) elements such that*

$$\max_{S \in \mathcal{S}} |\chi(S)| = O\left(\sqrt{n}\sqrt{\log \frac{m}{n}}\right)$$

Note that for $m = n$, the above gives discrepancy $O(\sqrt{n})$. With above lemma in hand, the full proof of Spencer's theorem is straightforward: we color the first $\frac{n}{2}$ elements, then $\frac{n}{4}$ of the remaining, and so forth. We finish in at most $\log n$ rounds, and hence the total discrepancy will be at most (case $m = n$)

$$O(1) \sum_{i=1}^{\log n} \sqrt{\frac{n}{2^{i-1}}} \sqrt{\log \left(\frac{2^{i-1} m}{n}\right)} = O(\sqrt{n})$$

by a simple computation (the terms decrease faster than a geometric series).

Sketch of the Proof: For simplicity, we examine the case $m = n$. Spencer's main idea here is to use the Pigeon Hole principle to find two proper coloring $\chi_1, \chi_2 \in \{-1, 1\}^n$ satisfying the following two properties:

1. χ_1, χ_2 have "similar" discrepancy profiles:

$$\max_{S \in \mathcal{S}} |\chi_1(S) - \chi_2(S)| = \sqrt{n}$$

2. χ_1, χ_2 have large hamming distance:

$$|\{i \in [n] : \chi_1(i) \neq \chi_2(i)\}| \geq \frac{n}{2}$$

With these two properties, we get that $\frac{1}{2}(\chi_1 - \chi_2) \in \{-1,0,1\}^n$ is a valid partial coloring on at least $\frac{n}{2}$ elements with discrepancy $O(\sqrt{n})$.

To find these colorings we proceed as follows. For every vector $z \in \mathbb{Z}^n$, we associate the bucket

$$B_z = \sqrt{n} \ \ [\frac{z_1 - 1}{2}, \frac{z_1 + 1}{2}] \times \cdots \times [\frac{z_n - 1}{2}, \frac{z_n + 1}{2}]$$

Now let χ denote a random $\{-1,1\}^n$ coloring. Let $S_1, \ldots, S_n$ denote the sets of $\mathcal{S}$. Using the concentration of $\chi(S_i)$ for $i \in [n]$ (Chernoff bounds suffice), one can show that there exists a bucket B_z, $z \in \mathbb{Z}^n$ such that

$$\Pr((\chi(S_1), \ldots, \chi(S_n)) \in B_z) \geq 2^{-\frac{1}{5}n}$$

This implies that there are at least $2^{\frac{4}{5}n}$ colorings $c \in \{-1,1\}^n$ such that $(c(S_1), \ldots, c(S_n)) \in B_z$. Note now that any two such colorings satisfy property (1) above, so we need only show that we can find 2 such colorings that lie "far" apart. For this, we use the classic bound from combinatorics, which states that any subset of the n dimensional Hamming cube having of diameter of most d has at most

$$\sum_{i=0}^{d} \binom{n}{i}$$

elements (size of the Hamming ball of radius d). From here, one can work out that since the subset of colorings we are interested in has size at least $2^{\frac{4}{5}n}$, it must have diameter at least $\Omega(n)$. This allows us to produce the colorings χ_1, χ_2 as needed.

The Entropy Method: To get at the right type of SDP for Spencer's problem, we will need a generalization of the partial coloring Lemma. Spencer's Lemma only deals with finding partial colorings having *uniform* discrepancy bounds. We will need a criterion to guarantee the existence of partial colorings having more general discrepancy requirements.

The entropy method, developed by Beck in [Bec81], provides such a criterion. We do not state the theorem here, but generally speaking, it says that for a set system $\mathcal{S} = (S_1, \ldots, S_m)$ on a universe $[n]$, there exists a partial coloring χ on $\frac{n}{2}$ elements satisfying $|\chi(S_i)| \leq \Delta_i$, $i \in [m]$, provided that the Δ_i's are not too "tight on average".

Therefore from the SDP standpoint, as long as the sequence Δ_i satisfies Beck's criterion, we get that the following SDP is feasible:

$$\begin{aligned} \|\sum_{i \in S_j} v_i\|^2 &\leq \Delta_j^2 \ \ \forall j \in [m] && \text{Low Discrepancy} \\ \|\sum_{i=1}^{n} v_i\|^2 &\geq \frac{n}{2} && \text{High Coordinate Variance} \\ \|v_i\|^2 &\leq 1 \ \ \forall i \in [n] \end{aligned} \tag{3}$$

As before, we generate a random n dimensional gaussian $g = (g_1, \ldots, g_n)$, where our coordinate updates will be $\eta_i = v_i \cdot g_i$. The vector $(\eta_1, \ldots, \eta_n)$ will "mimic" a true partial coloring.

New Approach: We will leverage the above type SDP to generate the random walk steps. We analyze the case where $\mathcal{S} = \{S_1, \ldots, S_n\}$ with each $S_j \subseteq [n]$. The main idea is as follows:

1. Start with random walk with SDP right hand sides, $\Delta_j = c\sqrt{n}$ for $j \in [n]$. Let χ denote our current fractional coloring.
2. Each $\chi(S_j)$ will do a random walk with expected discrepancy $O(\sqrt{n})$. However some of the $\chi(S_j)$ will become problematic, so we adjust their Δ's on the fly.
3. To do this we set thresholds $T_1 \le T_2 \le T_3 \ldots$ where say $T_i = (2-1/i)c'\sqrt{n}$ for some constant c', with associated delta values $d_1 \ge d_2 \ge d_3 \ge \ldots$. Whenever an $|\chi(S_j)|$ passes a threshold T_r, we set its $\Delta_j = d_r$ and continue.
4. The analysis will show that there are only a few problematic sets at any step, and hence the entropy penalty will be low enough for us continue (i.e. the SDP will always remain feasible).

5 Conclusion

We end with some remarks and open problems.

Problem: Find polytime algorithm which outputs a coloring achieving discrepancy at most $\text{herdisc}(\mathcal{S})\sqrt{\log m}$.

Upcoming Work: [Bansal - Spencer] All the algorithms above can be derandomized (add new constraint to the SDP). This work gives derandomizations for probabilistic inequalities that are stronger than Chernoff bounds (Exponential moment technique for derandomizing Chernoff loses $\sqrt{\log n}$).

One may wish to examine other non-constructive proofs:

1. Lattices (Minkowski's Theorem).
2. Fixed-Point Based (Nash, Sperner's Lemma).
3. Topological (Hypergraph Matching).

There are also other discrepancy problems where progress can be made:

1. Beck-Fiala Conjecture (More generally Komlos conjecture).
2. Erdos Discrepancy Problem.

References

[AS00] N. Alon and J.H. Spencer. *The probabilistic method*. Wiley-Interscience series in discrete mathematics and optimization. Wiley, 2000.

[Bec81] J. Beck. Roths estimate on the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1:319–325, 1981.

[Cha01] B. Chazelle. *The discrepancy method: randomness and complexity.* Cambridge University Press, 2001.

[CNN11] Moses Charikar, Alantha Newman, and Aleksander Nikolov. Tight hardness results for minimizing discrepancy. In *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2011.

[Mat10] J. Matousek. *Geometric Discrepancy: An Illustrated Guide.* Algorithms and Combinatorics. Springer, 2010.

[Spe85] Joel Spencer. Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706, 1985.