Progress and Problems in Discrepancy

Nikhil Bansal
(CWI and Eindhoven)
Discrepancy

Universe: $U = [1, \ldots, n]$
Subsets: $S_1, S_2, \ldots, S_m$

Color elements red/blue so each set is colored as evenly as possible.

Given $\chi: [n] \rightarrow \{ -1, +1 \}$
Disc $\left( \chi \right) = \max_S \left| \sum_{i \in S} \chi(i) \right| = \max_S |\chi(S)|$

Disc (set system) $= \min_{\chi} \max_S |\chi(S)|$

Capture various properties of the set system.
Lots of questions/applications in various areas.
Given an $m \times n$ matrix $A$, find $x \in \{-1,1\}^n$, to minimize 
\[ \text{disc}(A) = \|Ax\|_\infty \]

Vector balancing view: Given vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ find $x \in \{-1,1\}^n$ to minimize 
\[ \|\sum_i x_i v_i\|_\infty \]

Can also consider more general norms $K$: symmetric convex body
Find $x \in \{-1,1\}^n$ to minimize 
\[ \|\sum_i x_i v_i\|_K \]
Discrepancy: All about beyond the probabilistic method

Two problems: Spencer’s setting, Komlos’ problem
Open questions (could geometry of polynomials help?)

Classical methods from discrepancy:
Partial Coloring Method, Banaszczyk’s method
(Non-constructive, argue about all colorings simultaneously)

Recent algorithmic approaches.
New algorithmic ways to go beyond the probabilistic method

Can they help for Kadison Singer, applications of interlacing poly?
Two examples

**Spencer Setting:** Discrepancy of any set system on n elements and m sets?

**[Spencer’85]:** (independently by Gluskin’87)
For \( m = n \) discrepancy \( \leq 6n^{1/2} \)

Tight: Cannot beat \( 0.5 \, n^{1/2} \) (Hadamard Matrix).

**Random coloring** gives \( O(n \log n)^{1/2} \)
Proof: For set \( S \), \( \Pr \{ \text{disc}(S) \approx c |S|^{1/2} \} \approx \exp(-c^2) \)
Set \( c = O(\log n)^{1/2} \) and apply union bound.

**Tight.** Random gives \( \Omega(n \log n)^{1/2} \) with very high prob.
Spencer setting

More generally: $O\left(n^{1/2} \log^{1/2} \left(\frac{m}{n}\right)\right)$ for $m$ sets, $n$ elements

Random: $(n \log m)^{1/2}$

Nothing special about 0/1. e.g. $|a_{ij}| \leq 1$ also fine.

Invented the Partial Coloring Method (a key tool in discrepancy)

Open Problem: Matrix Spencer

Given (symmetric) $A_1, \ldots, A_n$ with spectral norm $\leq 1$

Is there a signing s.t. $\left|\sum_i x_i A_i\right| = O\left(n^{1/2}\right)$

For simultaneously diagonalizable matrices, follows from Spencer.
Komlos Problem

Given vectors $v_1, \ldots, v_n \in \mathbb{R}^m$ with $|v_i|_2 \leq 1$
find a signing to minimize $|\sum_i x_i v_i|_{\infty}$

Random coloring gives $\Omega(n^{1/2})$
E.g. If $m=1$, have $v_i \in [-1,1]$

Partial Coloring: $O(\log n)$
Banaszczyk: $O((\log n)^{1/2})$

Conjectured bound: $O(1)$
Beck Fiala Conjecture

Discrepancy of low degree set systems, where each element lies in at most $t$ sets? (i.e. 0-1 matrix where each column has $\leq t$ 1’s).

Scaling by $\frac{1}{t^{1/2}}$ gives unit columns.

Random: $\Omega(n^{1/2})$ (a row could have $n$ 1’s).

Beck-Fiala’81: $2t - 1$, $2t - \log^\ast t$ [Bukh’16]

Banaszczyk’97: $t^{1/2} (\log n)^{1/2}$

Conjecture: $O(t^{1/2})$
Non-constructive methods

1) Partial Coloring Method:
   Beck/Spencer early 80’s: Probabilistic Method + Pigeonhole
   Gluskin’87: Convex Geometric Approach

   Very versatile
   Loss adds over $O(\log n)$ iterations

2) Banaszczyk’98: Based on a deep convex geometric result
   Produces full coloring directly
Spencer’s $O(n^{1/2})$ result

Partial Coloring suffices: For any set system with $m$ sets, there exists a coloring on $\geq n/2$ elements with discrepancy

$O(n^{1/2} \log^{1/2} (2m/n))$  

[For $m=n$, disc $= O(n^{1/2})$]

Algorithm for total coloring:

Repeatedly apply partial coloring lemma

Total discrepancy

$O( n^{1/2} \log^{1/2} 2 )$  [Phase 1]
$+ O( (n/2)^{1/2} \log^{1/2} 4 )$  [Phase 2]
$+ O((n/4)^{1/2} \log^{1/2} 8 )$  [Phase 3]
$+ \ldots$  

$= O(n^{1/2})$
A geometric view

Spencer’85: Any 0-1 matrix \((n \times n)\) has disc \(\leq 6 \sqrt{n}\)

Gluskin’87: Convex geometric approach

Consider polytope \(P(t) = -t \mathbf{1} \leq A x \leq t \mathbf{1}\)

\(P(t)\) contains a point in \([-1,1]^n\) for \(t = 6 \sqrt{n}\)

Gluskin’87: If \(K\) symmetric, convex with large (Gaussian) volume (> \(2^{-n/100}\)) then \(K\) contains a point with many coordinates \([-1,+1]\)

d-dim Gaussian Measure: \(\gamma_d(x) = \exp(-|x|^2/2) (2\pi)^{-d/2}\)

\(\gamma_d(K)\): \(\Pr[(y_1, ..., y_m) \in K]\) each \(y_i\) iid \(N(0,1)\)

What is the Gaussian volume of \([-1,1]^n\) cube
A geometric view

Gluskin’87: If K symmetric, convex with large (Gaussian) volume (> $2^{-n/100}$) then K contains a point with many coordinates \{-1,+1\}

Similar to Minkowski’s theorem:
K symmetric has a non-zero point in $\mathbb{Z}^n$, if Vol(K) > $2^n$

Proof: Look at $K + x$ for all $x \in \{-1,1\}^n$
Total volume of shifts $= 2^{\Omega(n)}$ 
$\gamma_n(K + x) \geq \gamma_n(K) \exp(-|x|^2/2)$
Some point $z$ lies in $2^{\Omega(n)}$ copies

$z = k + x$ and $z = k' + x'$ where $x, x'$ have large hamming distance
Gives $(x - x')/2 = (k - k')/2 \in K$. 


Gluskin for Polytopes

Gluskin’87: If K symmetric, convex with large (Gaussian) volume (> $2^{-n/100}$) then K contains a point with many coordinates \{-1,+1\}

Spencer’s result proof:
Consider polytope $P(t) = -t \mathbf{1} \leq Ax \leq t \mathbf{1}$
Show Gaussian volume large enough for $t = c\sqrt{n}$

Sidak’s Thm: $\gamma_n(K) \geq \Pi_i \gamma_n(Slab_i)$
$Slab_i = -t \leq a_i x \leq t$

Thm: Given an $m \times n$ matrix $A$, there is a partial coloring satisfying
$|a_i x| \leq \lambda_i |a_i|_2$ for each row $i$, provided $\sum_i e^{-\lambda_i^2} \leq \frac{n}{5}$
Comparison w/ random coloring

Given an m x n matrix A, there is a partial coloring satisfying
\[ |a_i x| \leq \lambda_i |a_i|_2 \] for each row i, provided
\[ \sum_i e^{-\lambda_i^2} \leq \frac{n}{5} \]

Can view as extending Chernoff bounds
1) \( n/5 \) vs 1 (Chernoff)
2) Partial Coloring vs Full (Chernoff)

E.g. Can get 0 discrepancy on \( n/10 \) rows (very powerful)

Key tool in most discrepancy problems
Application: Komlos

Claim: Get partial coloring with $O(1)$ discrepancy.

Assume $n \leq m$ (linear algebraic argument)
For each column $j$, $\sum_i a_{ij}^2 \leq 1$
Sum of $a_{ij}^2$ over all matrix entries $\leq n$
Average sum per row $\leq n/m \leq 1$.

Call a row $i$ big if $\sum_j a_{ij}^2 > 10$. At most $n/10$ of these.
Set $\lambda_i = 0$ for big rows. Else $\lambda_i = O(1)$.

Gluskin: $|a_i x| \leq \lambda_i |a_i|_2$ for each row $i$, provided $\sum_i e^{-\lambda_i^2} \leq \frac{n}{5}$
Annoying loss of $O(\log n)$ to get full coloring
Ideal case

Beck-Fiala Setting: At most $n/10$ big (>10t) sets

Partial Coloring: 0 for big sets.

About $s^{1/2}$ for small sets of size $s$.

“Ideal” life cycle of a set

Ideal case: Discrepancy $= t^{1/2} + (t/2)^{1/2} + (t/4)^{1/2} + …$
What can go wrong

Trouble: A set can get $t^{1/2}$ discrepancy, but very few elements colored.
Banaszczyk’s method
$O(\log^{1/2} n)$ for Komlos
Thm: Let $A$ have columns $v_1, \ldots, v_n \in \mathbb{R}^m$, $\|v_i\|_2 \leq 1/5$

$K = $ symmetric convex body with $\gamma_m(K) \geq \frac{1}{2}$

$\exists x \in \{-1,1\}^n$ s.t. $Ax \in K$

Constants somewhat arbitrary

For non-symmetric $K$, need $\gamma_m(K) > \frac{1}{2}$ to ensure $0 \in K$ (e.g. if halfspace)
Banaszczyk’s Theorem

Cube: \[ K = O(\log m)^{1/2} \quad [-1,1]^m \quad \gamma_m(K) \geq 1/2 \]

Komlos: Given unit vectors in \( R^m \),
\[ \exists \text{ signed sum w/ } \ell_\infty\text{-norm } O(\log m)^{1/2} \]

Surprising results for various bodies \( K \).
Proof idea

Given $v_1, \ldots, v_n$, each $|v_i| < 1/5$. $\gamma_m(K) \geq \frac{1}{2}$

Goal: Find signing $\sum_i x_i v_i \in K$

Key observation: Signing exists iff

Some signing of $v_2, \ldots, v_n$ with sum in $(K + v_1) \cup (K - v_1)$.

Convexify:

Remove regions of $K$ width $< 2|v_1|$ along $v_1$

Lose and gain volume.

(non-trivial) computation to show volume stays $\geq \frac{1}{2}$
Algorithmic history

Partial Coloring now constructive

Bansal’10: SDP + Random walk
Lovett Meka’12: Random walk + linear algebra
Rothvoss’14: Convex geometric
Many others by now [Harvey, Schwartz, Singh], [Eldan, Singh], [Lee], ...

Banaszczyk based approaches:

[B., Dadush, Garg’16]: $O(\log n)^{1/2}$ algorithm for Komlos problem
[B., Dadush, Garg, Lovett 18]: algorithm for general Banaszczyk.
Useful View

Independent rounding.

A (complicated) view
Brownian motion in cube.

Cube: \{ -1, +1 \}^n

dimension: element
vertex: coloring

Same as randomized rounding
Each coordinate rounded independently
(martingale property of the walk)
Useful View

If additional constraints.
Can tailor walk accordingly.

Pick covariance matrix for $\Delta x^t$
(slow down towards bad regions)

Design barrier functions

$$a_i x \leq \lambda_i |a_i|_2$$

Cube: $\{-1,+1\}^n$

dimension: element
vertex: coloring

$$(x_1, \ldots, x_n)$$

$$a_i x \geq -\lambda_i |a_i|_2$$
Lovett Meka Algorithm

Random walk, \( \gamma \) N(0,1) in each dimension

a) Fix \( j \) if \( x_j = \pm 1 \)

b) If row \( a_i \) gets tight (disc(\( a_i \)) = \( \lambda_i |a_i|_2 \))

Move in subspace \( a_i x = \lambda_i |a_i|_2 \)

(not violate discrepancy)

Thm: Given an m x n matrix A, finds a partial coloring satisfying

\[ |a_i x| \leq \lambda_i |a_i|_2 \text{ for each row } i, \text{ provided } \sum_i e^{-\lambda_i^2} \leq \frac{n}{5} \]
**Lovett Meka Algorithm**

Random walk, $\gamma \sim N(0,1)$ in each dimension

a) Fix $j$ if $x_j = \pm 1$

b) If row $a_i$ gets tight ($\text{disc}(a_i) = \lambda_i |a_i|_2$)

Move in subspace $a_i x = \lambda_i |a_i|_2$

(not violate discrepancy)

**Idea:** Walk makes progress as long as dimension = $\Omega(n)$

($E[\sum_i x_i^2]$ rises by $\Omega(n)\gamma^2$ per step)

After $\frac{10}{\gamma^2}$ steps: $\Pr[ \text{Row } a_i \text{ tight}] \approx \exp(-\lambda_i^2)$

As $\sum_i \exp(-\lambda_i^2) \leq \frac{n}{5}$ so $n/5$ tight rows in expectation

As stays in cube, $\Omega(n)$ variables must have hit $\pm 1$,
Recall trouble with Partial Coloring

Beck Fiala Setting

Trouble: A set can get $t^{1/2}$ discrepancy, but very few elements colored.
Correlations in Lovett-Meka

Consider set $S = \{1, 2, \ldots, t\}$

Ideal case: If \textit{randomly} color each element

\[
\begin{align*}
\text{Progress} &= t \\
\text{discrepancy} &\approx t^{1/2}
\end{align*}
\]

Suppose move in subspace $x_1 = x_2 = \cdots = x_t$

E.g. if have constraints $x_1 - x_2 = 0, x_2 - x_3 = 0, \ldots$

Can only color all $+1$ or all $-1$.

\[
\begin{align*}
\text{Progress} &= t \\
\text{discrepancy} &= t
\end{align*}
\]

In Lovett-Meka, such sets hit subspace at $t^{1/2}$ discrepancy, but progress is \textit{only} $t^{1/2}$
Suggests a solution

Used to get an algorithmic $O(\log^{1/2} n)$ bound for Komlos [B., Dadush, Garg’16]

Can we design a walk that moves in some subspace, but still looks “random” enough?

E.g. If constrained to move in subspace $x_1 = x_2 = \cdots = x_t$

Just set $\Delta x_i = 0$ for $i=1,2,\ldots,t$
Can still do a random walk for $i = t+1,\ldots,n$. 
Better covariance matrices

$W$: arbitrary subspace $\dim(W) \leq (1 - \delta)n$

Need to walk in $W^\perp$

Property 1: $w^T(\Delta x) = 0 \quad \forall w \in W$

$E[w^T \Delta x \Delta x^T w] = 0$ or $w^T Y w = 0$

Property 2: Still looks almost independent.

For any direction $c = (c_1, \ldots, c_n)$

$E[(\sum_i c_i \Delta x_i)^2] \leq \frac{1}{\delta} \sum_i c_i^2 E[\Delta x_i^2]$

$c^T Y c \leq \left(\frac{1}{\delta}\right) c^T \text{diag}(Y) c \quad \forall c \in \mathbb{R}^n$

$Y \preceq \left(\frac{1}{\delta}\right) \text{diag}(Y)$
Can find such a good walk

**Key Thm:** If $\dim(W) \leq (1 - \delta)n$

There is a non-zero solution $Y$ to the SDP

$$w^T Y w = 0 \quad \forall w \in W$$

$$Y \preceq \left(\frac{1}{\delta}\right) \text{diag}(Y)$$

$$Y \succeq 0$$

Proof: Using SDP duality
Algorithm for Komlos

Time $t$: If $n_t$ variables alive, at most $n_t/10$ big rows
Pick $W =$ span of these constraints.

Run the SDP walk.
No phases, continue till all variables $-1/+1$ (i.e. $n_t = 0$).

If row big = discrepancy 0
When becomes small, just like a random walk.

“Freedman type” martingale analysis (avoid dependence on time steps), gives the result.
Making Banaszczyk Algorithmic

**Thm [Banaszczyk 97]:** Input $v_1, \ldots, v_n \in \mathbb{R}^d$, $|v_i|_2 \leq 1$

$\forall$ convex body $K$, with $\gamma_d(K) \geq \frac{1}{2}$

$\exists$ coloring $x \in \{-1,1\}^n$ s.t. $\sum_i x(i)v_i \in 5K$

Coloring depends on the convex body $K$.

How is $K$ specified? (input size could be exponential)

**Idea** [Dadush, Garg, Lovett, Nikolov’16]: Minimax Thm. (2-player game)

Universal distribution on colorings that works for all convex bodies
Equivalent formulation

Alternate formulation [Dadush, Garg, Lovett, Nikolov’16]:

\[ \exists \text{ distribution on colorings } x \in \{-1,1\}^n, \]
\[ \text{s.t. } Y = \sum_i x(i) v_i \text{ is } \approx N(0,1) \text{ in every direction} \]

\[ Y \in R^d \text{ is } \sigma\text{-subgaussian if in all directions } \theta \in R^d, |\theta|_2 = 1, \]
\[ \langle \theta, Y \rangle \text{ has same tails as } N(0, \sigma^2) \text{ i.e. } \Pr[|\langle \theta, Y \rangle| \geq \lambda] \leq 2 \exp(-\lambda^2/2\sigma^2) \]

Lemma: \[ Y \in K \text{ (for } K \text{ convex, } \gamma_d(K) \geq \frac{1}{2} \text{) with constant prob.} \]

Suffices to sample \[ x \text{ implicitly from such a distribution.} \]
Goal: \( \exists \) distribution on colorings \( x \in \{-1, 1\}^n \), s.t. random vector \( Y = \sum_i x(i)v_i \) is \( O(1) \) subgaussian

\[ \forall \theta \in S^{m-1}, \quad \langle Y, \theta \rangle = \sum_i x(i) \langle v_i, \theta \rangle \text{ decays like } N(0,1). \]

Special cases:

1) \( v_i \) are **Orthogonal**: Random \( \pm \) coloring \( x_i \) works
   
   As \( \sum_i c_i x_i \approx N(0, \sum_i c_i^2) \)

   \[ \text{Var}(\langle Y, \theta \rangle) = \sum_i \langle v_i, \theta \rangle^2 \leq |\theta|^2 \leq 1 \]

2) All equal vectors

\( v_1 = \cdots = v_n = v \) random coloring bad: \( \Omega(\sqrt{n}) \) in direction \( v \)

Need dependent coloring: \( n/2 \) +1’s and \( n/2 \) -1’s
Gram Schmidt Walk

Algorithm: Consider vectors $v_1, ..., v_n$
Write $v_n = c_1 v_1 + ... c_{n-1} v_{n-1} + w_n$
where $w_n \in \text{span} \ (v_1, ..., v_{n-1})^\perp$

Let direction $c = (c_1, ..., c_{n-1}, -1)$
Update coloring $x$ as $\delta c$ s.t. $E[\delta] = 0$
i.e. $\Delta x = +\delta_1 c$ or $-\delta_2 c$

Key Point: $\Delta Y = \sum_i \Delta x(i) v_i = \delta (\sum_{i=1}^{n-1} c_i v_i - v_n) = -\delta w_n$.

As $\delta \leq 2$ and $E[\delta] = 0$
$\Delta \langle Y, \theta \rangle$ evolves as a martingale with variance $O(\langle \theta, w_n \rangle^2)$
Proof Idea (ideal case)

Suppose pivot is the one to freeze every time

\[ \Delta Y: \delta_n w_n \]
\[ \Delta Y: \delta_{n-1} w_{n-1} \]

\[ w_1, \ldots, w_n \] obtained by Gram Schmidt process.

\[ w_1 = v_1 \]
\[ \hat{w}_1 = w_1 / |w_1| \]
\[ w_2 = v_2 - \langle v_2, \hat{w}_1 \rangle \hat{w}_1 \]
\[ \hat{w}_2 = w_2 / |w_2| \]
\[ w_3 = v_3 - \langle v_3, \hat{w}_1 \rangle \hat{w}_1 - \langle v_3, \hat{w}_2 \rangle \hat{w}_2 \]
\[ \hat{w}_3 = w_3 / |w_3| \]

\[ Y = \delta_n w_n + \delta_{n-1} w_{n-1} + \cdots + \delta_1 w_1 \]

\[ Var (\langle Y, \theta \rangle) = \sum_i \delta_i^2 \langle w_i, \theta \rangle^2 \leq \sum_i \delta_i^2 \langle \hat{w}_i, \theta \rangle^2 \leq 4|\theta|^2 = 4 \]
Some more details

\(v_1, \ldots, \bigtimes_5, \ldots, v_n\)  

No reason why pivot should get fixed.

Suppose \(v_5\) gets fixed.

\(w_n\) becomes \(w'_n\) which can be longer.

Proof idea: Can charge increase in \(|w_n|^2\) to \(v_5\) disappearing.

Track evolution of \(E[e^{\lambda(\theta,Y)}]\) by a suitable potential
and show \(E[e^{\lambda(\theta,Y)}] = e^{O(\lambda^2)}\) for each \(\theta, \lambda\)

(Recall \(Z\) is \(\sigma\)-subgaussian iff \(E[e^{\lambda Z}] = e^{O(\lambda^2\sigma^2)}\) for all \(\lambda\))
Concluding remarks

Besides Matrix Spencer and Komlos conjecture, many problems in discrepancy still open
(Steinitz problem, Tusnady’s problem, … (log $n)^{1/2}$ gap)

Lots of progress on lower bounds (SDP duality, convex geometry)

[Rothvoss’14] Algorithm for Gluskin (general convex bodies)
[Nikolov, Talwar’15] Approximating hereditary discrepancy

Various new uses in algorithm design, beating “union bound”
Bin-packing [Rothvoss’13], iterated + randomized rounding [B.’19]
Questions!