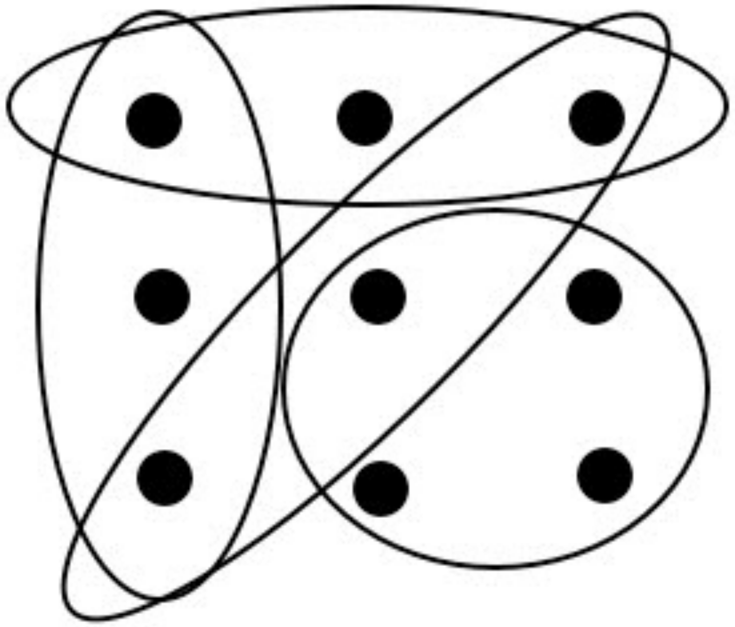
Algorithm for Komlós discrepancy problem matching Banaszczyk's bound

Nikhil Bansal (TU Eindhoven)
Daniel Dadush (CWI, Amsterdam)
Shashwat Garg (TU Eindhoven)

Combinatorial discrepancy

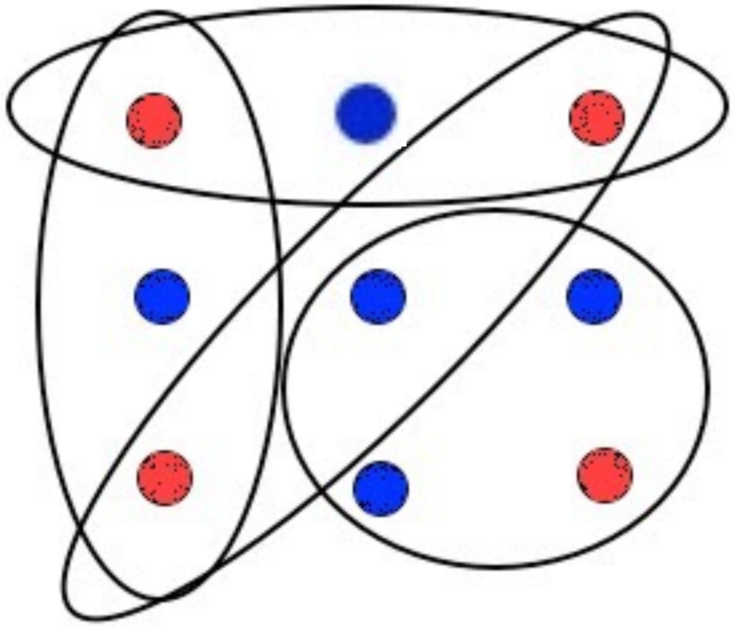
Set system $\mathcal{S} = \{S_1, \ldots, S_m\}$



Combinatorial discrepancy

Set system $\mathcal{S} = \{S_1, \ldots, S_m\}$

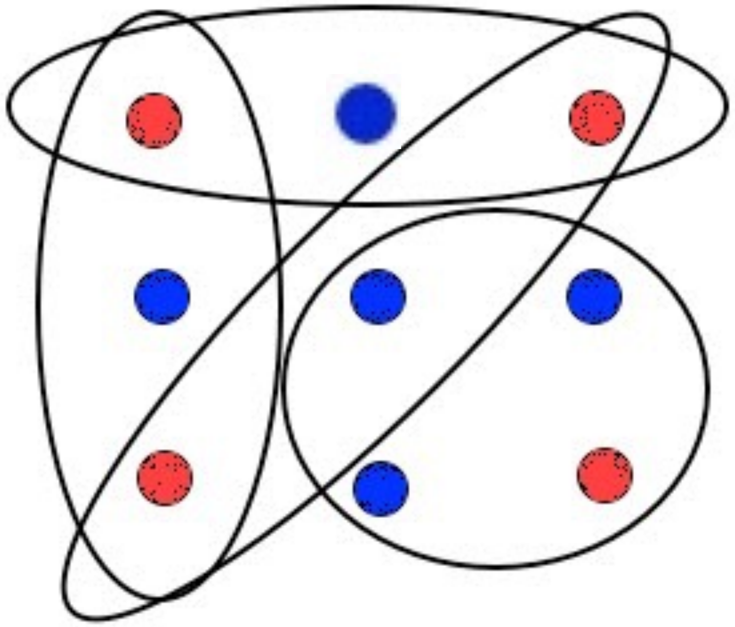
Color each point $\chi_i = -1$ or $+1$



Combinatorial discrepancy

Set system $\mathcal{S} = \{S_1, \ldots, S_m\}$

Color each point $\chi_i = -1$ or $+1$



$$disc(S_j, \chi) = |\sum_{i \in S_j} \chi_i|$$

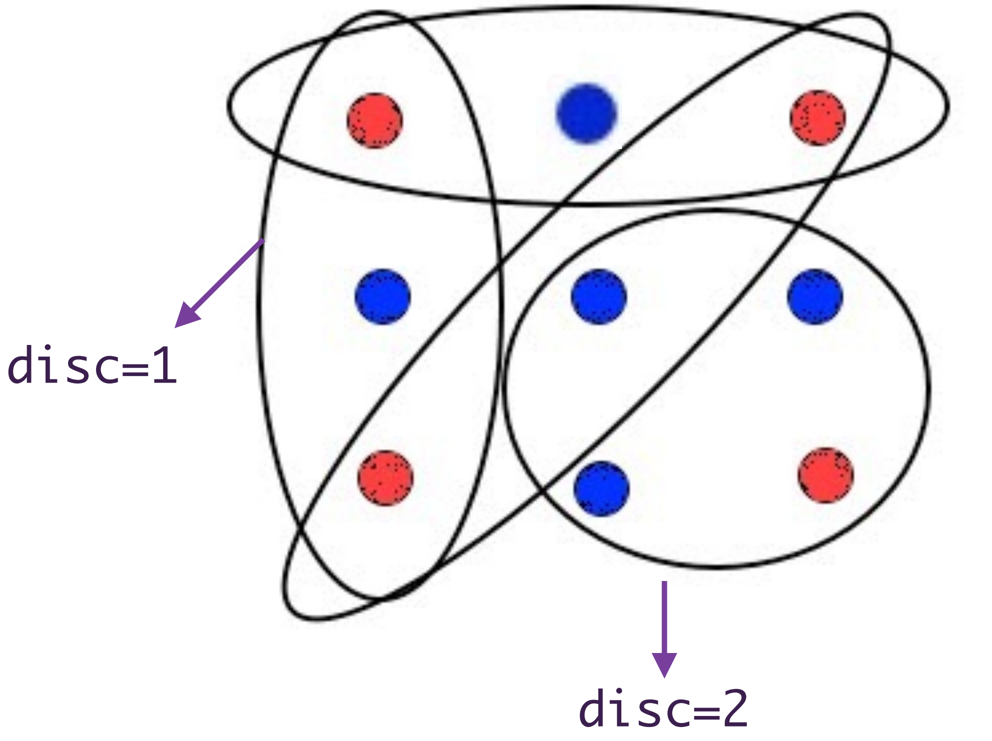
$$disc(\mathcal{S}, \chi) = \max_j disc(S_j, \chi)$$

$$disc(\mathcal{S}) = \min_\chi \max_j disc(S_j, \chi)$$

Combinatorial discrepancy

Set system $\mathcal{S} = \{S_1, \ldots, S_m\}$

Color each point $\chi_i = -1$ or $+1$



$disc(\mathcal{S}, \chi) = 2$

$$disc(S_j, \chi) = |\sum_{i \in S_j} \chi_i|$$

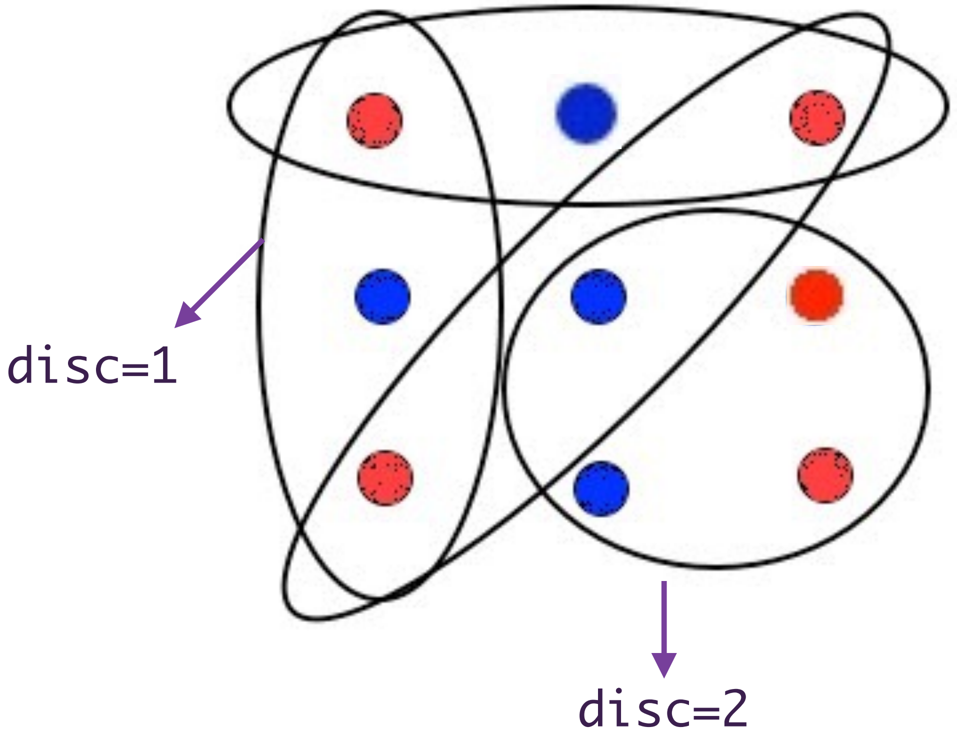
$$disc(\mathcal{S}, \chi) = \max_j disc(S_j, \chi)$$

$$disc(\mathcal{S}) = \min_\chi \max_j disc(S_j, \chi)$$

Combinatorial discrepancy

Set system $\mathcal{S} = \{S_1, \ldots, S_m\}$

Color each point $\chi_i = -1$ or $+1$



$$disc(S_j, \chi) = |\sum_{i \in S_j} \chi_i|$$

$$disc(\mathcal{S}, \chi) = \max_j disc(S_j, \chi)$$

$$disc(\mathcal{S}) = \min_\chi \max_j disc(S_j, \chi)$$

$$disc(\mathcal{S}, \chi) = 2$$

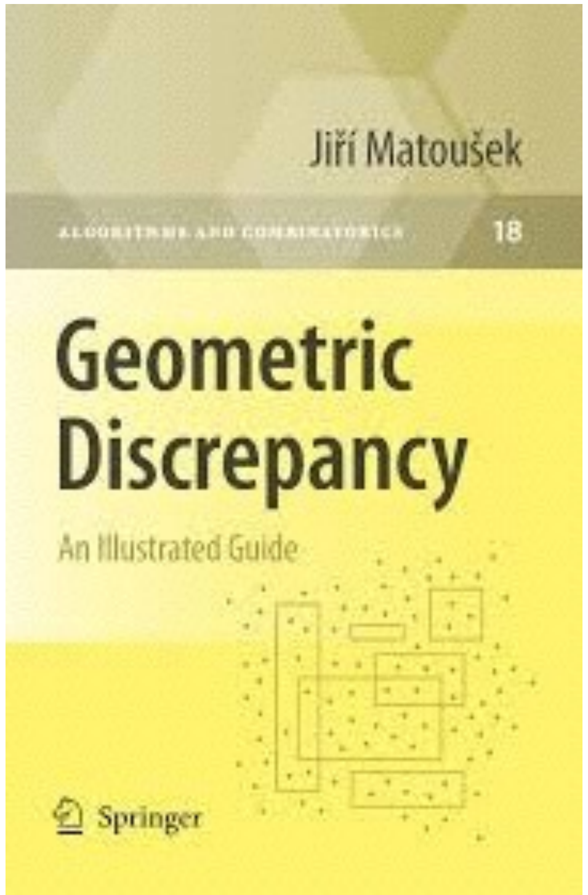
$$disc(\mathcal{S}) = 1$$

Matrix Notation

$$\text{Incidence matrix } A = \begin{pmatrix} 1 & 0 & \cdots & 1 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{pmatrix}$$

Coloring $\chi \in \{-1, +1\}^n$ to minimize $\|A\chi\|_\infty$

Applications



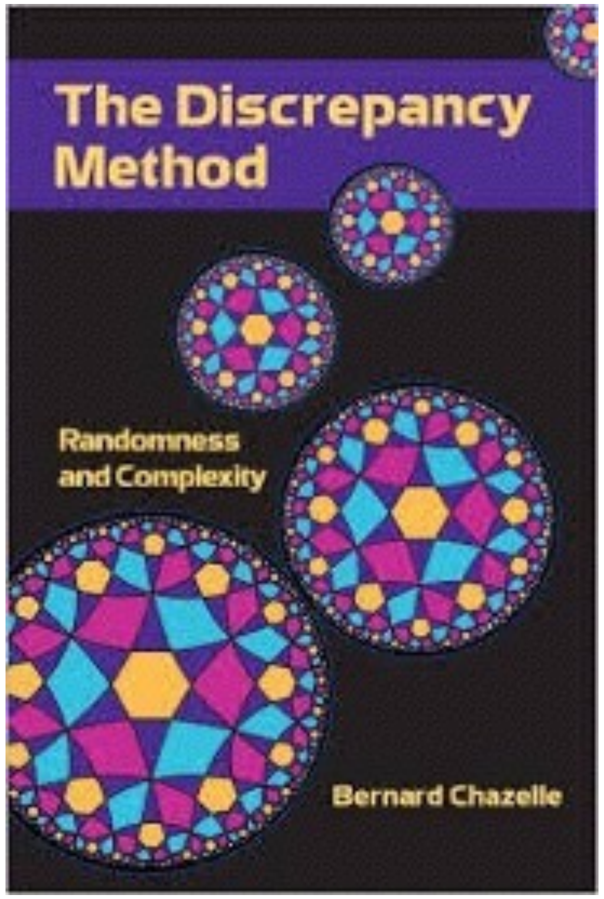
Jiří Matoušek

ALGORITHMS AND COMBINATORICS 18

Geometric Discrepancy

An Illustrated Guide

Springer



The Discrepancy Method

Randomness and Complexity

Bernard Chazelle

Approximation algorithms: Bin Packing [Rothvoss'13]

Data structures [Larsen'11]

many more …

Beck-Fiala Setting

- Every point lies in at most *t* sets
- Beck-Fiala['81] showed *2t-1* discrepancy
- Conjecture $O(\sqrt{t})$

Beck-Fiala Setting

- Every point lies in at most t sets
- Beck-Fiala['81] showed $2t-1$ discrepancy
- Conjecture $O(\sqrt{t})$

Komlós:

- Vectors of A of length at most 1
- conjecture $O(1)$

Beck-Fiala Setting

- Every point lies in at most *t* sets
- Beck-Fiala['81] showed *2t-1* discrepancy
- Conjecture $O(\sqrt{t})$

Partial coloring gives $O(\sqrt{t}\log n)$

Beck-Fiala Setting

- Every point lies in at most *t* sets
- Beck-Fiala['81] showed *2t-1* discrepancy
- Conjecture $O(\sqrt{t})$

Partial coloring gives $O(\sqrt{t}\log n)$

Banaszczyk['98]:

- Komlós : $O(\sqrt{\log n})$
- Beck-Fiala: $O(\sqrt{t \log n})$

Earlier work

- $2t - 1$ (constructive [Beck-Fiala'81])
- $O(\sqrt{t}\log n)$ (non-constructive)

  [Srinivasan'97]: entropy method
- $O(\sqrt{t}\log n)$ (constructive) : based on partial colorings

  [Bansal'10]: SDP based random walk

  [Lovett-Meka'12]: simplified random walk without SDP

  [Rothvoss'14]: random sample and project

  [Harvey, Schwartz, Singh'14], [Eldan, Singh'14]
- $O(\sqrt{t \log n})$ (non-constructive [Banaszczyk'98])

Our main result

There is an algorithm which returns a coloring of discrepancy $O(\sqrt{t \log n})$ in Beck-Fiala setting w.h.p.

Our main result

There is an algorithm which returns a coloring of discrepancy $O(\sqrt{t \log n})$ in Beck-Fiala setting w.h.p.

Also extends to Komlós with discrepancy $O(\sqrt{\log n})$

Earlier approaches

- Based on partial coloring
    - Color half the points globally
    - Discrepancy = $\sqrt{t}$
    - Repeat log n times
    - Total discrepancy = $\sqrt{t}\log n$

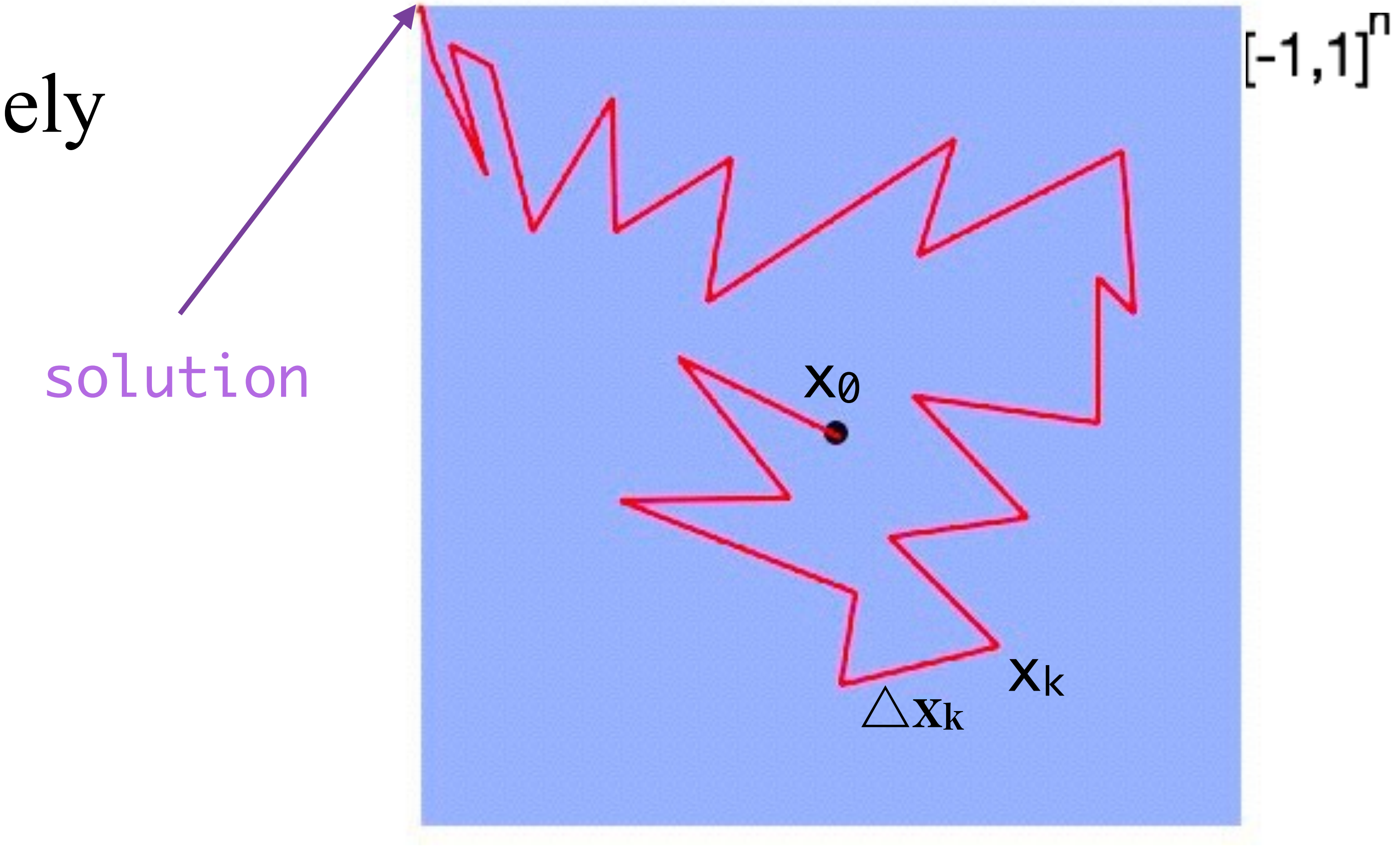
Earlier approaches

- Based on partial coloring
    - Color half the points globally
    - Discrepancy = $\sqrt{t}$
    - Repeat log n times
    - Total discrepancy = $\sqrt{t} \log n$

- Obstacle: Partial coloring can give discrepancy $\sqrt{t}$ to set S while coloring only few points of S

Main idea

- Random walk cleverly
- Use SDP to get a finer control over how discrepancy of each set evolves
- Choose updates steps wisely



Overcoming $\sqrt{t}\log n$ barrier

a) Zero discrepancy for big sets $(> 100t)$

↘ at most n/100 of these

m≥n

Also possible in earlier approaches

Overcoming $\sqrt{t}\log n$ barrier

Partial coloring can give discrepancy $\sqrt{t}$ to set S
while coloring only few elements of S

Overcoming $\sqrt{t}\log n$ barrier

Partial coloring can give discrepancy $\sqrt{t}$ to set S
while coloring only few elements of S

b)Proportional discrepancy

$$\left(\sum_{i\in S}\Delta x_k(i)\right)^2 \leq \sum_{i\in S}\Delta x_k(i)^2$$

$(\Delta\text{discrepancy})^2$

`energy injected into S at time k`

energy of $S = \sum_{i\in S} x_k(i)^2$

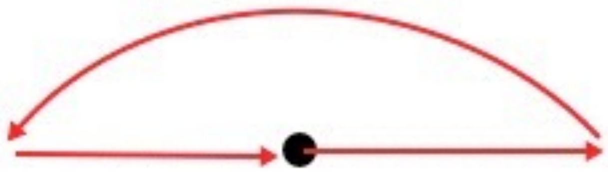
Overcoming $\sqrt{t}\log n$ barrier

So if total energy injected into set $S \approx |S| \approx t$
we get $\mathrm{disc}(S) \approx \sqrt{t}$

Overcoming $\sqrt{t}\log n$ barrier

So if total energy injected into set $S \approx |S| \approx t$
we get $\mathrm{disc}(S) \approx \sqrt{t}$

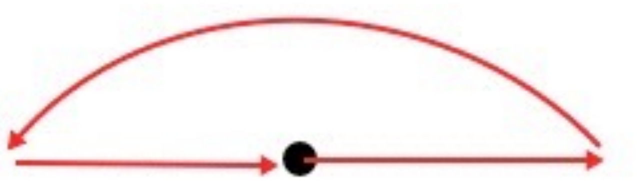
But total energy injected >> final energy !!



Overcoming $\sqrt{t}\log n$ barrier

So if total energy injected into set $S \approx |S| \approx t$
we get $\mathrm{disc}(S) \approx \sqrt{t}$

But total energy injected >> final energy !!

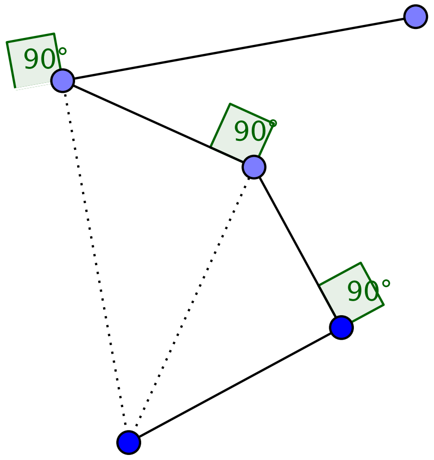


c)Orthogonality constraints

$$\left(\sum_{i\in S} x_{k-1}(i)\Delta x_k(i)\right)^2 \leq \sum_{i\in S} \Delta x_k(i)^2$$

$(\text{fluctuations})^2$

energy injected into S at time k



Summing it up

1. sets only start incurring discrepancy when their size at most *100t*
2. after that, $(\text{disc})^2 <$ total energy injected
3. and total energy injected $< |S| <$ *100t*

Thus, discrepancy of S $\leq 10\sqrt{t}$

Lose a factor of $\sqrt{\log n}$ in union bound

Capturing this via SDP

zero discrepancy for big sets:

A_k = elements not yet frozen

$$\| \sum_{i \in S} u_i \|_2^2 = 0 \qquad \text{if } |S \cap A_k| > 100t$$

Capturing this via SDP

zero discrepancy for big sets:

A_k = elements not yet frozen

$$\|\sum_{i\in S} u_i\|_2^2 = 0 \qquad \text{if } |S \cap A_k| > 100t$$

$(\text{disc})^2$ < energy injected:

$$\|\sum_{i\in S} u_i\|_2^2 \leq 2\sum_{i\in S}\|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

Capturing this via SDP

zero discrepancy for big sets:

A_k = elements not yet frozen

$$\|\sum_{i\in S} u_i\|_2^2 = 0 \qquad \text{if } |S \cap A_k| > 100t$$

(disc)^2 < energy injected:

$$\|\sum_{i\in S} u_i\|_2^2 \leq 2\sum_{i\in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

total energy injected $\approx$ final energy:

$$\|\sum_{i\in S} x_{k-1}(i) u_i\|_2^2 \leq 2\sum_{i\in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

Capturing this via SDP

zero discrepancy for big sets:

A_k = elements not yet frozen

$$\|\sum_{i\in S} u_i\|_2^2 = 0 \qquad \text{if } |S \cap A_k| > 100t$$

(disc)^2 < energy injected:

$$\|\sum_{i\in S} u_i\|_2^2 \le 2\sum_{i\in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \le 100t$$

total energy injected ≈ final energy:

$$\|\sum_{i\in S} x_{k-1}(i) u_i\|_2^2 \le 2\sum_{i\in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \le 100t$$

make large progress:

$$maximize \sum_{i\in A_k} \|u_i\|_2^2$$

Algorithm

```
$x_0 = 0^n$
for k=1,…,T :
    solve SDP
    $x_k(i) = x_{k-1}(i) + \gamma\langle r_k, u_i\rangle$
    if for some i,  $|x_k(i)| \geq 1 - 1/n$
        freeze i
```

small step size

random ±

remove it from A_k

T is chosen large enough s.t. all elements are frozen by time T w.h.p.

Proof strategy

1. at each time *k*, SDP has value $\Omega(A_k)$

   thus we make a large move at each step

2. all properties are satisfied and $\mathrm{disc(S)} = \sqrt{t \log n}$

   whp till time T

3. all elements are frozen by time T whp

SDP has large value

$$maximize \sum_{i \in A_k} \|u_i\|_2^2$$ subject to

$$\|\sum_{i \in S} u_i\|_2^2 = 0 \qquad \text{if } |S \cap A_k| > 100t$$

$$\|\sum_{i \in S} u_i\|_2^2 \leq 2\sum_{i \in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

$$\|\sum_{i \in S} x_{k-1}(i) u_i\|_2^2 \leq 2\sum_{i \in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

$$\|u_i\|_2^2 \leq 1$$

A_k = elements not yet frozen

SDP has large value

$$\mathit{maximize} \sum_{i \in A_k} \|u_i\|_2^2$$ subject to

$$\|\sum_{i \in S} u_i\|_2^2 = 0 \qquad \text{if } |S \cap A_k| > 100t$$

$$\|\sum_{i \in S} u_i\|_2^2 \leq 2 \sum_{i \in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

$$\|\sum_{i \in S} x_{k-1}(i) u_i\|_2^2 \leq 2 \sum_{i \in S} \|u_i\|_2^2 \qquad \text{if } |S \cap A_k| \leq 100t$$

$$\|u_i\|_2^2 \leq 1$$

A_k = elements not yet frozen

Theorem: strong duality holds.

Dual SDP

$$\text{Minimize} \sum_{i \in A_k} b_i \qquad \text{s.t.}$$

$$\sum_{i \in A_k} b_i e_i e_i^T + \sum_{S \in \text{big sets}} \alpha_S v_S v_S^T \quad + \quad \sum_{S \in \text{small sets}} \left(\beta_S (v_S v_S^T - 2I_S) + \beta_S^x (x_S x_S^T - 2I_S) \right) \succeq \; I$$

$$\begin{array}{rcll} b_i & \geq & 0 & \forall i \in A_k \\ \alpha_S & \in & \mathbb{R} & \forall \text{big sets } S \\ \beta_S, \beta_S^x & \geq & 0 & \forall \text{small sets } S \end{array}$$

indicator vector of set S

Dual SDP

$$\text{Minimize } \sum_{i\in[n]} b_i \qquad \text{s.t.}$$

$$\sum_{i\in[n]} b_i e_i e_i^T + \sum_{S\in\text{big sets}} \alpha_S v_S v_S^T \quad + \quad \sum_{S\in\text{small sets}} \left(\beta_S(v_S v_S^T - 2I_S) + \beta_S^x(x_S x_S^T - 2I_S)\right) \succeq \; I$$

Dual SDP

$$\text{Minimize} \sum_{i \in [n]} b_i \qquad \text{s.t.}$$

$$\sum_{i \in [n]} b_i e_i e_i^T + \sum_{S \in \text{big sets}} \alpha_S v_S v_S^T \quad + \sum_{S \in \text{small sets}} \left(\beta_S (v_S v_S^T - 2I_S) + \beta_S^x (x_S x_S^T - 2I_S)\right) \succeq \; I$$

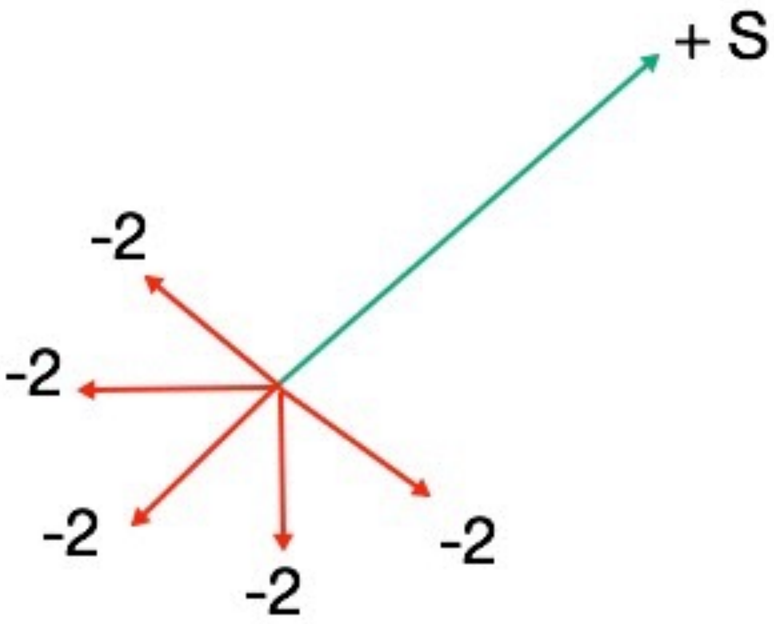
$$\text{dim} \leq \frac{n}{100}$$

Dual SDP

$$\text{Minimize} \sum_{i\in[n]} b_i \qquad \text{s.t.}$$

$$\sum_{i\in[n]} b_i e_i e_i^T + \sum_{S\in\text{big sets}} \alpha_S v_S v_S^T \ + \sum_{S\in\text{small sets}} \big(\beta_S(v_S v_S^T - 2I_S) + \beta_S^x(x_S x_S^T - 2I_S)\big) \succeq \ I$$

$$\text{dim} \le \frac{n}{100}$$



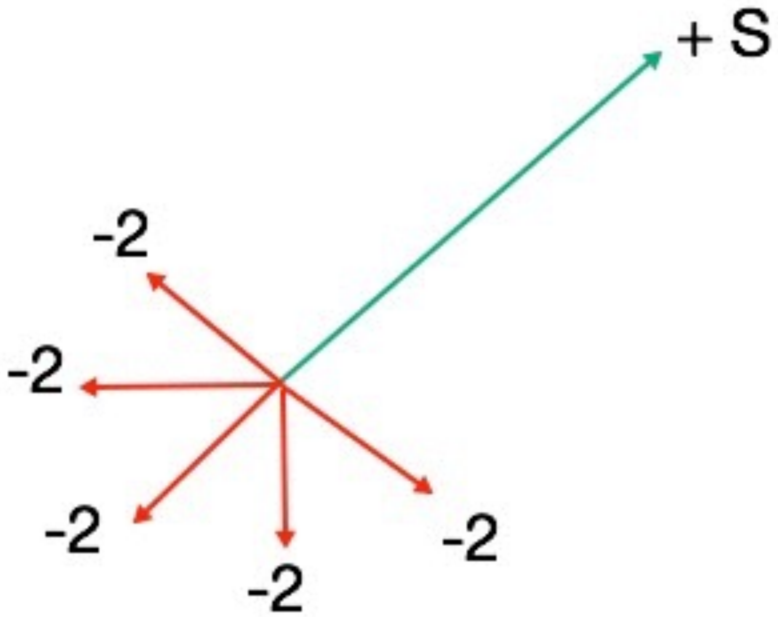
Dual SDP

$$\text{Minimize} \sum_{i\in[n]} b_i \qquad \text{s.t.}$$

$$\sum_{i\in[n]} b_i e_i e_i^T + \sum_{S\in\text{big sets}} \alpha_S v_S v_S^T \; + \sum_{S\in\text{small sets}} \big(\beta_S(v_S v_S^T - 2I_S) + \beta_S^x(x_S x_S^T - 2I_S)\big) \succeq \; I$$

$\dim \leq \frac{n}{100}$

$\preceq 0 \text{ on } \dim \geq n/2$



Dual SDP

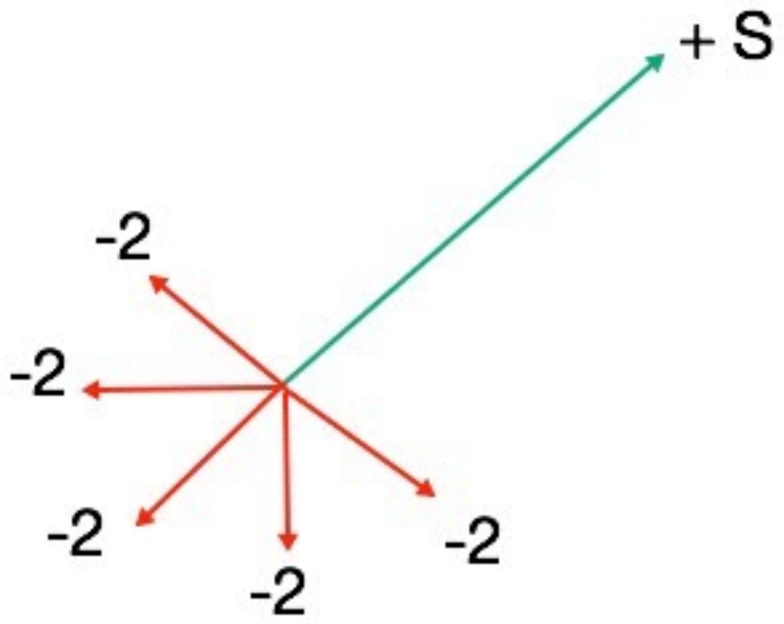
$$\text{Minimize} \sum_{i\in[n]} b_i \qquad \text{s.t.}$$

$$\sum_{i\in[n]} b_i e_i e_i^T + \sum_{S\in\text{big sets}} \alpha_S v_S v_S^T + \sum_{S\in\text{small sets}} \left(\beta_S(v_S v_S^T - 2I_S) + \beta_S^x(x_S x_S^T - 2I_S)\right) \succeq I$$

$$\frac{n}{2} - \frac{n}{100}$$

$$\dim \leq \frac{n}{100}$$

$$\preceq 0 \text{ on } \dim \geq n/2$$



Theorem: Let $B = \sum_v \left(vv^T - 2diag(vv^T)\right)$. Then, $y^T By \leq 0$ for $y \in W$ and $dim(W) \geq n/2$.

Proof:

$$M = \begin{array}{cc} \begin{matrix} \quad i \quad & \qquad \end{matrix} & \\ \begin{pmatrix} \vdots & \\ & \\ v_i & \cdots \\ & \\ & \end{pmatrix} & v \end{array}$$

$$D = \begin{array}{cc} \begin{pmatrix} \vdots & \\ & \\ \sqrt{\sum_v v_i^2} & \cdots \\ & \\ & \end{pmatrix} & i \end{array}$$

diagonal

$$\begin{aligned} y^T By &\leq 0 \\ &\Updownarrow \\ \|My\|_2^2 &\leq 2\|Dy\|_2^2 \\ &\Updownarrow \\ \|MD^{-1}y\|_2^2 &\leq 2\|y\|_2^2 \end{aligned}$$

columns of length ≤ 1

Lemma: Given $m \times n$ matrix M will all columns of ℓ_2-norm ≤ 1. Then, $\|My\|_2^2 \leq 2\|y\|_2^2$ for all $y \in W$ where $dim(W) \geq n/2$.

Lemma: Given $m \times n$ matrix M will all columns of ℓ_2-norm ≤ 1. Then, $\|My\|_2^2 \leq 2\|y\|_2^2$ for all $y \in W$ where $dim(W) \geq n/2$.

Proof:

Sum of squared-singular values of M is at most n.

Lemma: Given $m \times n$ matrix M will all columns of ℓ_2-norm ≤ 1. Then, $\|My\|_2^2 \leq 2\|y\|_2^2$ for all $y \in W$ where $dim(W) \geq n/2$.

Proof:

Sum of squared-singular values of M is at most n.

$$\sum_i \sigma_i^2 = Tr[M^T M] = \|M\|_F \leq n$$

Lemma: Given $m \times n$ matrix M will all columns of ℓ_2-norm ≤ 1. Then, $\|My\|_2^2 \leq 2\|y\|_2^2$ for all $y \in W$ where $dim(W) \geq n/2$.

Proof:

Sum of squared-singular values of M is at most n.

$$\sum_i \sigma_i^2 = Tr[M^T M] = \|M\|_F \leq n$$

Thus, half of the singular values are at most $\sqrt{2}$.

■

Bounding discrepancy

- How does *disc$_k$(S)* evolve?

$$disc_k(S) = disc_{k-1}(S) + \textstyle\sum_{i \in S} \gamma \langle r_k, u_i \rangle$$

Bounding discrepancy

- How does *disc$_k$(S)* evolve?

$$disc_k(S) = disc_{k-1}(S) + \sum_{i \in S} \gamma \langle r_k, u_i \rangle$$

- It is a martingale.

- Use martingale concentration inequalities

`Freedman's inequality: very strong !`

Bounding discrepancy

Freedman: Given a martingale $X_0, \ldots, X_T$, with $|X_t - X_{t-1}| \leq M$ and $W = \sum_t \mathbb{E}_{t-1}[(X_t - X_{t-1})^2]$,

$$\Pr[X_T \geq \lambda \text{ and } W \leq \sigma^2] \leq 2exp\left(-\frac{\lambda^2}{2(\sigma^2 + M\lambda/3)}\right)$$

Bounding discrepancy

Freedman: Given a martingale $X_0, \ldots, X_T$, with $|X_t - X_{t-1}| \leq M$ and $W = \sum_t \mathbb{E}_{t-1}[(X_t - X_{t-1})^2]$,

$$\Pr[X_T \geq \lambda \text{ and } W \leq \sigma^2] \leq 2exp\left(-\frac{\lambda^2}{2(\sigma^2 + M\lambda/3)}\right)$$

$disc_k(S)$ is a martingale with:

- M almost zero
- W roughly equal to total energy injected

Bounding discrepancy

Freedman: Given a martingale $X_0, \ldots, X_T$, with $|X_t - X_{t-1}| \leq M$ and $W = \sum_t \mathbb{E}_{t-1}[(X_t - X_{t-1})^2]$,

$$\Pr[X_T \geq \lambda \text{ and } W \leq \sigma^2] \leq 2exp\left(-\frac{\lambda^2}{2(\sigma^2 + M\lambda/3)}\right)$$

$disc_k(S)$ is a martingale with:

- M almost zero
- W roughly equal to total energy injected

$$\Pr[disc_T(S) \geq \lambda] \approx exp\left(-\lambda^2/100t\right)$$

Put $\lambda = O(\sqrt{t \log n})$ and union bound over $\leq nt$ sets

Komlós setting

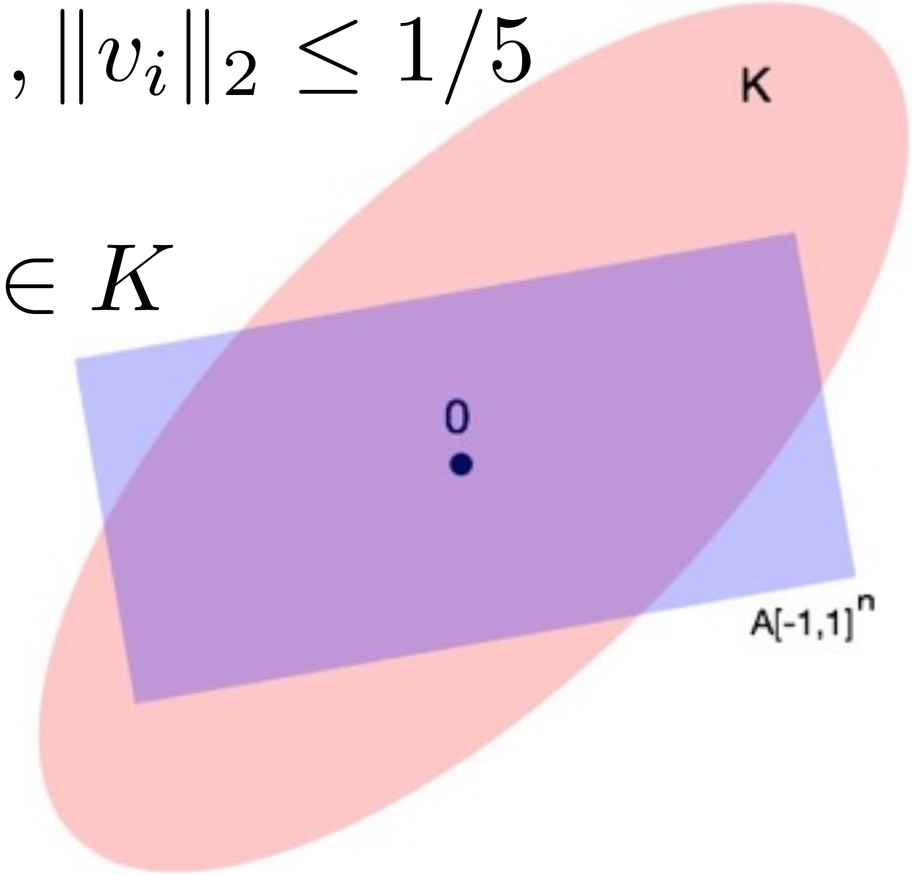
- Matrix A with columns of length ≤ 1. Find $x \in \{-1,1\}^n$ to minimise $\|Ax\|_\infty$

- Our algorithm also extends here with guarantee $O(\sqrt{\log n})$.

$$\Pr[(A\underset{\substack{\big\downarrow \\ \texttt{final coloring}}}{x_T})_i \geq \lambda] \approx exp\left(-\lambda^2/100\right)$$

Open Problem: General Banaszczyk

Theorem: Given $v_1, v_2, \ldots, v_n \in \mathbb{R}^m, \|v_i\|_2 \leq 1/5$ and convex body K, $\gamma_m(K) \geq 1/2$. There exists coloring ϵ_i s.t. $\sum_{i=1}^n \epsilon_i v_i \in K$



- ✦ We only proved the special case of $K = \sqrt{\log n}$ cube

- ✦ An algorithm for general K will imply
  - constructive approximation for hereditary discrepancy [Matousek,Nikolov,Talwar'15]
  - improved algorithm for Steinitz conjecture

Thank you for your attention

Any questions?