Constructive Discrepancy Minimization for Convex Sets

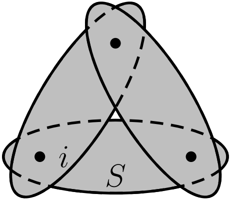
Thomas Rothvoss

UW Seattle

UNIVERSITY *of* WASHINGTON

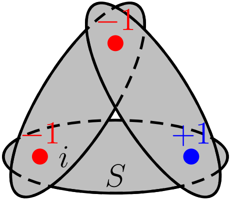
Discrepancy theory

- Set system $\mathcal{S} = \{S_1, \ldots, S_m\}, S_i \subseteq [n]$

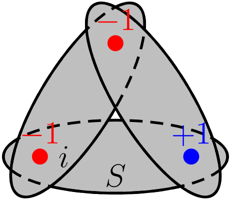


Discrepancy theory

- Set system $\mathcal{S} = \{S_1, \ldots, S_m\}, S_i \subseteq [n]$
- **Coloring** $\chi : [n] \to \{-1, +1\}$



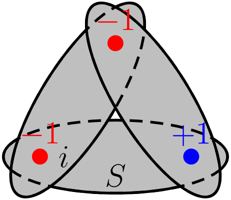
Discrepancy theory



- Set system $\mathcal{S} = \{S_1, \ldots, S_m\}, S_i \subseteq [n]$
- **Coloring** $\chi : [n] \to \{-1, +1\}$
- **Discrepancy**

$$\mathrm{disc}(\mathcal{S}) = \min_{\chi:[n]\to\{\pm 1\}} \max_{S\in\mathcal{S}} \Big| \sum_{i\in S} \chi(i) \Big|.$$

Discrepancy theory



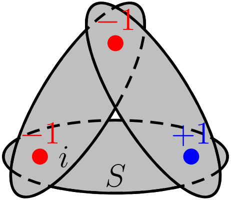
- Set system $\mathcal{S} = \{S_1, \dots, S_m\}, S_i \subseteq [n]$
- **Coloring** $\chi : [n] \to \{-1, +1\}$
- **Discrepancy**

$$\mathrm{disc}(\mathcal{S}) = \min_{\chi:[n]\to\{\pm 1\}} \max_{S \in \mathcal{S}} \Big| \sum_{i \in S} \chi(i) \Big|.$$

Known results:

- n sets, n elements: $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n})$ [Spencer '85]

Discrepancy theory



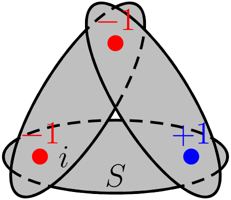
- Set system $\mathcal{S} = \{S_1, \ldots, S_m\}, S_i \subseteq [n]$
- **Coloring** $\chi : [n] \to \{-1, +1\}$
- **Discrepancy**

$$\mathrm{disc}(\mathcal{S}) = \min_{\chi:[n]\to\{\pm 1\}} \max_{S\in\mathcal{S}} \Big| \sum_{i\in S} \chi(i) \Big|.$$

Known results:

- n sets, n elements: $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n})$ [Spencer '85]
- Every element in $\leq t$ sets: $\mathrm{disc}(\mathcal{S}) < 2t$ [Beck & Fiala '81]

Discrepancy theory



- Set system $\mathcal{S} = \{S_1, \ldots, S_m\}, S_i \subseteq [n]$
- **Coloring** $\chi : [n] \to \{-1, +1\}$
- **Discrepancy**

$$\mathrm{disc}(\mathcal{S}) = \min_{\chi:[n]\to\{\pm 1\}} \max_{S\in\mathcal{S}} \Big| \sum_{i\in S} \chi(i) \Big|.$$

Known results:

- n sets, n elements: $\mathrm{disc}(\mathcal{S}) = O(\sqrt{n})$ [Spencer '85]
- Every element in $\leq t$ sets: $\mathrm{disc}(\mathcal{S}) < 2t$ [Beck & Fiala '81]

Main method: Find a **partial coloring** $\chi : [n] \to \{0, \pm 1\}$

- low discrepancy $\max_{S\in\mathcal{S}} |\chi(S)|$
- $|\mathrm{supp}(\chi)| \geq \Omega(n)$

Discrepancy theory (2)

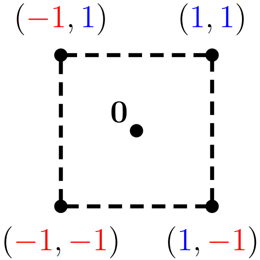
Lemma (Spencer)

For m set on $n \leq m$ elements there is a **partial coloring** of discrepancy $O(\sqrt{n \log \frac{2m}{n}})$.

- Run argument $\log n$ times
- Total discrepancy is

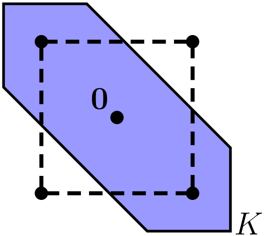
$$\lesssim \sqrt{n} + \sqrt{n/2} + \sqrt{n/2^2} + \ldots + 1 = O(\sqrt{n})$$

Thm of Spencer-Gluskin-Giannopolous



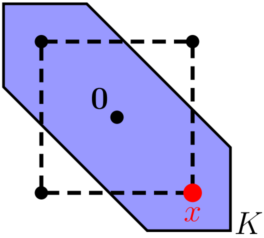
Thm of Spencer-Gluskin-Giannopolous

Goal: For $K := \left\{ x \in \mathbb{R}^n : |\sum_{j \in S_i} x_j| \leq 100\sqrt{n} \; \forall i \in [n] \right\}$



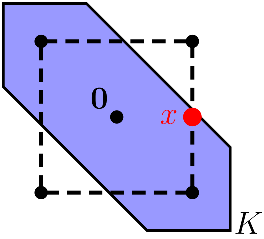
Thm of Spencer-Gluskin-Giannopolous

Goal: For $K := \left\{ x \in \mathbb{R}^n : |\sum_{j \in S_i} x_j| \leq 100\sqrt{n} \; \forall i \in [n] \right\}$ find a point $x \in \{-1, 1\}^n \cap K$



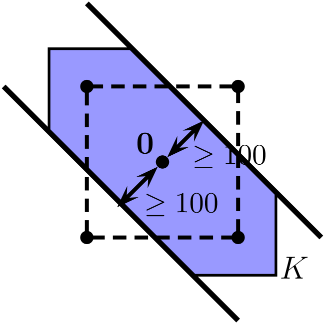
Thm of Spencer-Gluskin-Giannopolous

Goal: For $K := \left\{ x \in \mathbb{R}^n : |\sum_{j \in S_i} x_j| \leq 100\sqrt{n}\ \forall i \in [n] \right\}$ find a point $x \in \{-1, 0, 1\}^n \cap K$ with $|\mathrm{supp}(x)| \geq \Omega(n)$.



Thm of Spencer-Gluskin-Giannopolous

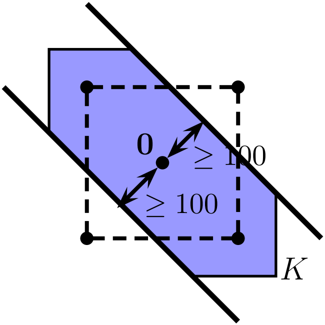
Goal: For $K := \left\{ x \in \mathbb{R}^n : |\sum_{j \in S_i} x_j| \leq 100\sqrt{n}\ \forall i \in [n] \right\}$ find a point $x \in \{-1, 0, 1\}^n \cap K$ with $|\text{supp}(x)| \geq \Omega(n)$.



- K is intersection of n **strips** of width 100

Thm of Spencer-Gluskin-Giannopolous

Goal: For $K := \left\{x \in \mathbb{R}^n : |\sum_{j \in S_i} x_j| \leq 100\sqrt{n} \; \forall i \in [n]\right\}$ find a point $x \in \{-1, 0, 1\}^n \cap K$ with $|\mathrm{supp}(x)| \geq \Omega(n)$.

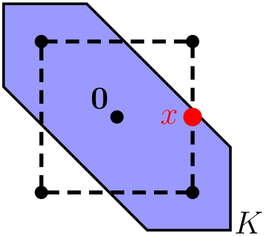


- K is intersection of n **strips** of width 100
- Gaussian measure

$$\gamma_n(K) \geq (\gamma_n(\text{width } 100 \text{ strip}))^n \geq e^{-n/100}$$

Thm of Spencer-Gluskin-Giannopolous

Goal: For $K := \left\{ x \in \mathbb{R}^n : |\sum_{j \in S_i} x_j| \leq 100\sqrt{n} \; \forall i \in [n] \right\}$ find a point $x \in \{-1, 0, 1\}^n \cap K$ with $|\mathrm{supp}(x)| \geq \Omega(n)$.



- K is intersection of n **strips** of width 100
- Gaussian measure

$$\gamma_n(K) \geq (\gamma_n(\text{width 100 strip}))^n \geq e^{-n/100}$$

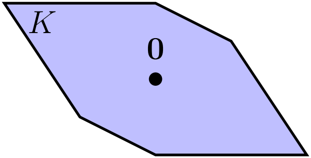
- Counting argument: any such K admits partial coloring

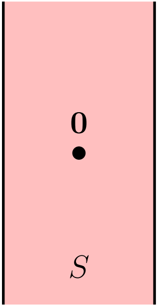
Gaussian measure

Lemma (Sidak-Kathri '67)

For K convex and symmetric and strip S,

$$\gamma_n(K \cap S) \geq \gamma_n(K) \cdot \gamma_n(S)$$



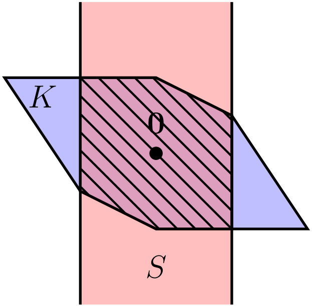


Gaussian measure

Lemma (Sidak-Kathri '67)

For K convex and symmetric and strip S,

$$\gamma_n(K \cap S) \geq \gamma_n(K) \cdot \gamma_n(S)$$



Partial coloring approaches

- **Spencer '85, Gluskin '87, Giannopolous '97:**
  - (+) very general
  - (−) non-constructive

Partial coloring approaches

- **Spencer '85, Gluskin '87, Giannopolous '97:**
  - (+) very general
  - (−) non-constructive
- **Bansal '10:** SDP-based random walk for Spencer's Thm
  - (+) poly-time
  - (−) custom-tailored to Spencers setting

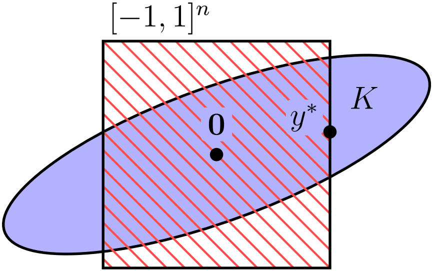
Partial coloring approaches

- **Spencer '85, Gluskin '87, Giannopolous '97:**
  - (+) very general
  - (−) non-constructive
- **Bansal '10:** SDP-based random walk for Spencer's Thm
  - (+) poly-time
  - (−) custom-tailored to Spencers setting
- **Lovett-Meka '12:**
  - (+) poly-time
  - (+) simple and elegant
  - (+/−) Works for any $K = \{x : |\langle x, v_i \rangle| \leq \lambda_i \forall i \in [m]\}$ with $\sum_{i=1}^{m} e^{-\lambda_i^2/16} \leq \frac{n}{16}$

Our main result

Theorem (R. 2014)

For a **convex symmetric set** $K \subseteq \mathbb{R}^n$ with $\gamma_n(K) \geq e^{-\delta n}$, one can find a $y \in K \cap [-1,1]^n$ with $|\{i : y_i = \pm 1\}| \geq \varepsilon n$ in **poly-time**.



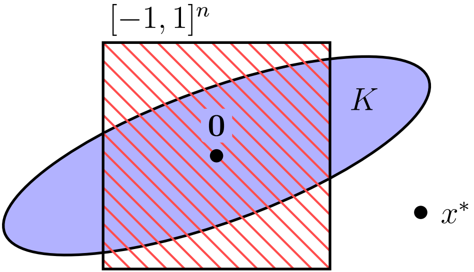
Our main result

Theorem (R. 2014)

For a **convex symmetric set** $K \subseteq \mathbb{R}^n$ with $\gamma_n(K) \geq e^{-\delta n}$, one can find a $y \in K \cap [-1,1]^n$ with $|\{i : y_i = \pm 1\}| \geq \varepsilon n$ in **poly-time**.

Algorithm:

(1) take a random $x^* \sim \gamma_n$



Our main result

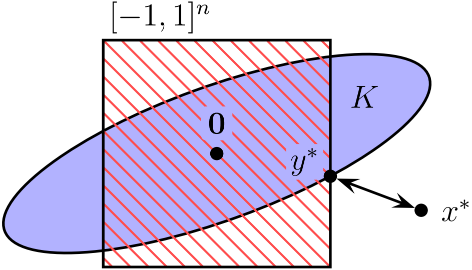
Theorem (R. 2014)

For a **convex symmetric set** $K \subseteq \mathbb{R}^n$ with $\gamma_n(K) \geq e^{-\delta n}$, one can find a $y \in K \cap [-1,1]^n$ with $|\{i : y_i = \pm 1\}| \geq \varepsilon n$ in **poly-time**.

Algorithm:

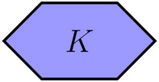
(1) take a random $x^* \sim \gamma_n$

(2) compute $y^* = \operatorname{argmin}\{\|x^* - y\|_2 \mid y \in K \cap [-1,1]^n\}$



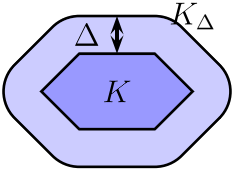
Isoperimetric inequality

- For set K



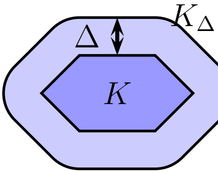
Isoperimetric inequality

- For set K, let $K_\Delta := \{x : d(x, K) \leq \Delta\}$



Isoperimetric inequality

- For set K, let $K_\Delta := \{x : d(x, K) \leq \Delta\}$

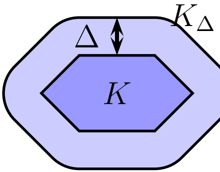


Lemma

$$\gamma_n(K) \geq e^{-\delta n} \implies \gamma_n(K_{3\sqrt{\delta n}}) \geq 1 - e^{-\delta n}$$

Isoperimetric inequality

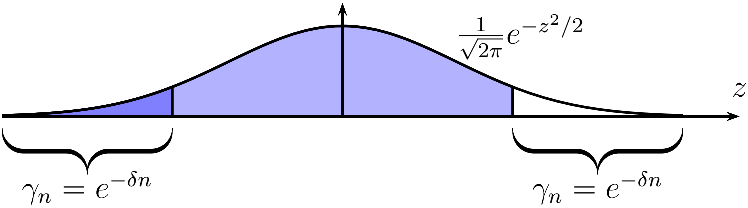
- For set K, let $K_\Delta := \{x : d(x, K) \leq \Delta\}$



Lemma

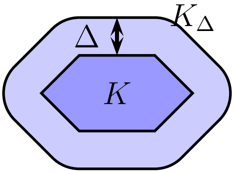
$\gamma_n(K) \geq e^{-\delta n} \implies \gamma_n(K_{3\sqrt{\delta n}}) \geq 1 - e^{-\delta n}$

- Isoperimetric inequality: worst case are half planes!



Isoperimetric inequality

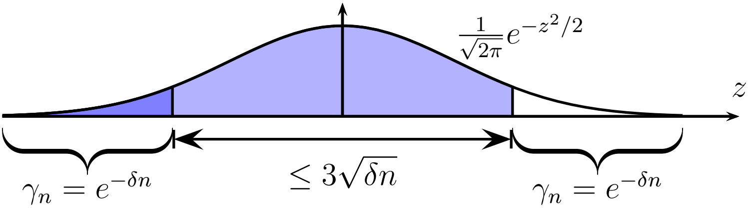
- For set K, let $K_\Delta := \{x : d(x, K) \le \Delta\}$



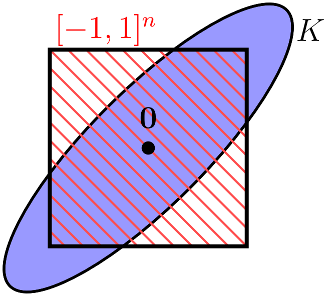
Lemma

$\gamma_n(K) \ge e^{-\delta n} \implies \gamma_n(K_{3\sqrt{\delta n}}) \ge 1 - e^{-\delta n}$

- Isoperimetric inequality: worst case are half planes!

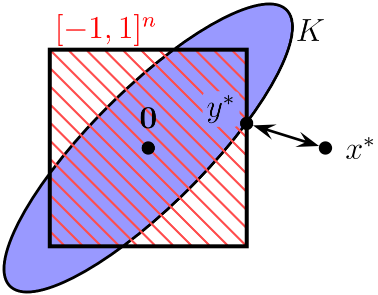


Analysis



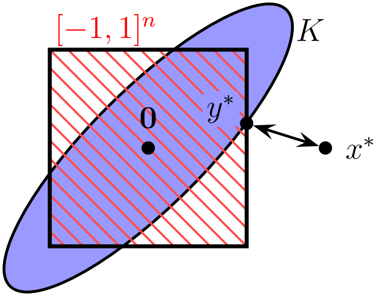
Analysis

- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$



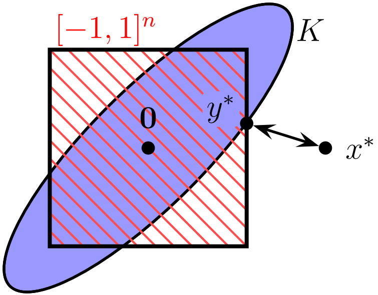
Analysis

- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$



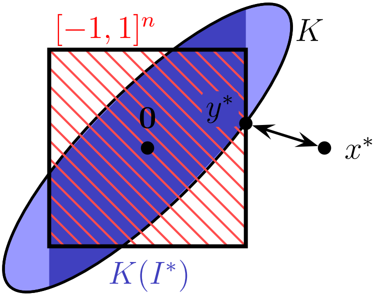
Analysis

- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$
- Def. $I^* := \{i : |y_i^*| = 1\}$
  Suppose $|I^*| \leq \varepsilon n$



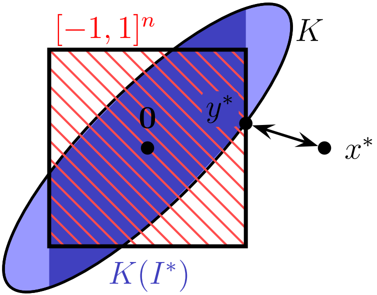
Analysis

- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$
- Def. $I^* := \{i : |y_i^*| = 1\}$
  Suppose $|I^*| \leq \varepsilon n$
- $K(I^*) := K \cap \{|x_i| \leq 1 : i \in I^*\}$



Analysis

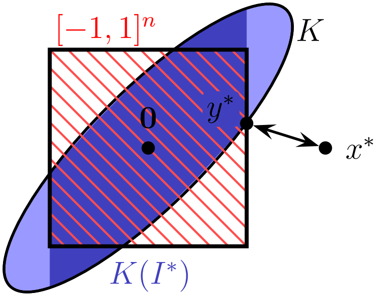
- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$
- Def. $I^* := \{i : |y_i^*| = 1\}$
  Suppose $|I^*| \leq \varepsilon n$
- $K(I^*) := K \cap \{|x_i| \leq 1 : i \in I^*\}$
- $\|x^* - y^*\|_2 = d(y^*, K(I^*))$



Analysis

- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$
- Def. $I^* := \{i : |y_i^*| = 1\}$
  Suppose $|I^*| \leq \varepsilon n$
- $K(I^*) := K \cap \{|x_i| \leq 1 : i \in I^*\}$
- $\|x^* - y^*\|_2 = d(y^*, K(I^*))$
- $K(I^*)$ still large:

$$\gamma_n(K(I^*)) \geq \gamma_n(K) \cdot (\gamma_n(\text{strip of width 1}))^{\varepsilon n} \geq e^{-2\delta n}$$

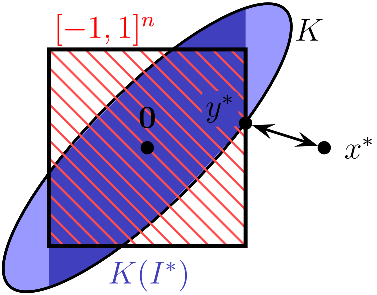


Analysis

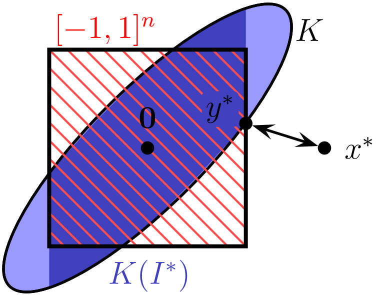
- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$
- Def. $I^* := \{i : |y_i^*| = 1\}$
  Suppose $|I^*| \leq \varepsilon n$
- $K(I^*) := K \cap \{|x_i| \leq 1 : i \in I^*\}$
- $\|x^* - y^*\|_2 = d(y^*, K(I^*))$
- $K(I^*)$ still large:

$$\gamma_n(K(I^*)) \geq \gamma_n(K) \cdot (\gamma_n(\text{strip of width } 1))^{\varepsilon n} \geq e^{-2\delta n}$$

- W.h.p. $d(x^*, K(I^*)) \leq \sqrt{3\delta} \cdot \sqrt{n} \ll \frac{1}{5}\sqrt{n}$



Analysis



- W.h.p. $\|x^* - y^*\|_2 \geq \frac{1}{5}\sqrt{n}$
- For any Q with $\gamma_n(Q) \geq e^{-o(n)}$:
  $\Pr[d(x^*, Q) \geq \frac{1}{5} \cdot \sqrt{n}] \leq e^{-\Omega(n)}$
- Def. $I^* := \{i : |y_i^*| = 1\}$
  Suppose $|I^*| \leq \varepsilon n$
- $K(I^*) := K \cap \{|x_i| \leq 1 : i \in I^*\}$
- $\|x^* - y^*\|_2 = d(y^*, K(I^*))$
- $K(I^*)$ still large:

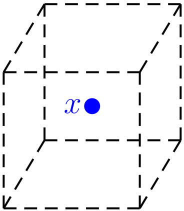
$$\gamma_n(K(I^*)) \geq \gamma_n(K) \cdot (\gamma_n(\text{strip of width 1}))^{\varepsilon n} \geq e^{-2\delta n}$$

- W.h.p. $d(x^*, K(I^*)) \leq \sqrt{3\delta} \cdot \sqrt{n} \ll \frac{1}{5}\sqrt{n}$
- Union bound over all $|I| \leq \varepsilon n$:

$$\Pr\Big[\bigcup_{|I| \leq \varepsilon n} d(x^*, K(I)) < \frac{1}{5}\sqrt{n}\Big] \leq e^{-\Omega(n)} \qquad \square$$

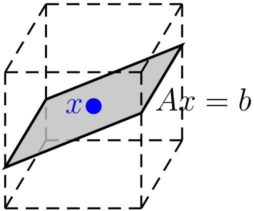
Application to approximation algorithms

- **Classical technique:** For $K = \{x \in \mathbb{R}^n \mid Ax = b\}$ with $\frac{n}{2}$ constraints $\rightarrow$ basic solution x' has $\leq \frac{n}{2}$ fractional variables.



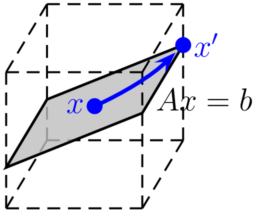
Application to approximation algorithms

- **Classical technique:** For $K = \{x \in \mathbb{R}^n \mid Ax = b\}$ with $\frac{n}{2}$ constraints $\to$ basic solution x' has $\leq \frac{n}{2}$ fractional variables.

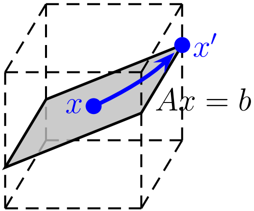


Application to approximation algorithms

- **Classical technique:** For $K = \{x \in \mathbb{R}^n \mid Ax = b\}$ with $\frac{n}{2}$ constraints $\rightarrow$ basic solution x' has $\leq \frac{n}{2}$ fractional variables.

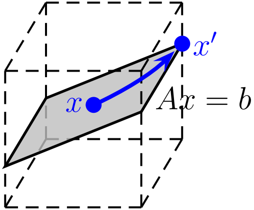


Application to approximation algorithms



- **Classical technique:** For $K = \{x \in \mathbb{R}^n \mid Ax = b\}$ with $\frac{n}{2}$ constraints $\rightarrow$ basic solution x' has $\leq \frac{n}{2}$ fractional variables.
- **Lovett-Meka rounding:**
  - x' has $n(\frac{1}{2} + \varepsilon)$ fractional variables
  - **Additional:** For $|\langle v_i, x - x' \rangle| \leq \lambda_i$ for unit vectors satisfying $\sum_{i=1}^{m} e^{-\lambda_i^2/16} \leq c(\varepsilon) \cdot n$

Application to approximation algorithms



- **Classical technique:** For $K = \{x \in \mathbb{R}^n \mid Ax = b\}$ with $\frac{n}{2}$ constraints $\to$ basic solution x' has $\leq \frac{n}{2}$ fractional variables.

- **Lovett-Meka rounding:**
  - x' has $n(\frac{1}{2} + \varepsilon)$ fractional variables
  - **Additional:** For $|\langle v_i, x - x'\rangle| \leq \lambda_i$ for unit vectors satisfying $\sum_{i=1}^{m} e^{-\lambda_i^2/16} \leq c(\varepsilon) \cdot n$

Better approximation algorithms for

- Bin Packing [R. '13]
- Broadcast scheduling [Bansal, Charikar, Krishnaswamy, Li '14]

Open problems

Setting: Given a set system $S_1, \ldots, S_m \subseteq [n]$ where each element is in at most t sets.

Beck-Fiala Conjecture: $\text{disc} \leq O(\sqrt{t})$

Open problems

Setting: Given a set system $S_1, \ldots, S_m \subseteq [n]$ where each element is in at most t sets.

Beck-Fiala Conjecture: $\text{disc} \leq O(\sqrt{t})$

Known bounds:

- $2t$ (constructive [Beck-Fiala '81])
- $O(\sqrt{t}\log n)$ (constructive)
- $O(\sqrt{t \cdot \log n})$ (non-constructive [Banaszczyk '97])

Open problems

Setting: Given a set system $S_1, \ldots, S_m \subseteq [n]$ where each element is in at most t sets.

Beck-Fiala Conjecture: $\text{disc} \leq O(\sqrt{t})$

Known bounds:

- $2t$ (constructive [Beck-Fiala '81])
- $O(\sqrt{t}\log n)$ (constructive)
- $O(\sqrt{t \cdot \log n})$ (non-constructive [Banaszczyk '97])

Open problem:

- Show that $\text{disc} \leq o(t)$

Open problems

Setting: Given a set system $S_1, \ldots, S_m \subseteq [n]$ where each element is in at most t sets.

Beck-Fiala Conjecture: $\text{disc} \leq O(\sqrt{t})$

Known bounds:

- $2t$ (constructive [Beck-Fiala '81])
- $O(\sqrt{t} \log n)$ (constructive)
- $O(\sqrt{t \cdot \log n})$ (non-constructive [Banaszczyk '97])

Open problem:

- Show that $\text{disc} \leq o(t)$

Thanks for your attention