Discrepancy Theory & Graph Sparsification

Victor Reis

Joint work with Thomas Rothvoss

UNIVERSITY *of* WASHINGTON

Matrix Discrepancy

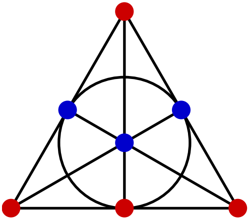
Given a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with rows $\boldsymbol{a}_{1\star}, \ldots, \boldsymbol{a}_{n\star} \in \mathbb{R}^n$, let

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

Matrix Discrepancy

Given a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with rows $\boldsymbol{a}_{1\star}, \ldots, \boldsymbol{a}_{n\star} \in \mathbb{R}^n$, let

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$



$$\mathrm{disc} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} = 3$$

Matrix Discrepancy

Given a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with rows $\boldsymbol{a}_{1\star}, \ldots, \boldsymbol{a}_{n\star} \in \mathbb{R}^n$, let

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

Matrix Discrepancy

Given a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with rows $\boldsymbol{a}_{1\star}, \ldots, \boldsymbol{a}_{n\star} \in \mathbb{R}^n$, let

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- $\operatorname{disc}(\boldsymbol{A}) \geqslant \min_{\boldsymbol{x} \in \{\pm 1\}^n} \sqrt{\frac{\sum_{i=1}^n \langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle^2}{n}} \geqslant \min_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sigma_{\min}(\boldsymbol{A})$

Matrix Discrepancy

Given a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with rows $\boldsymbol{a}_{1\star}, \dots, \boldsymbol{a}_{n\star} \in \mathbb{R}^n$, let

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- $\operatorname{disc}(\boldsymbol{A}) \geqslant \min_{\boldsymbol{x} \in \{\pm 1\}^n} \sqrt{\frac{\sum_{i=1}^n \langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle^2}{n}} \geqslant \min_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sigma_{\min}(\boldsymbol{A})$
- When $\boldsymbol{A}$ is Hadamard $\frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sqrt{n} = \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$ so

$$\exists \boldsymbol{A} : \operatorname{disc}(\boldsymbol{A}) \geqslant \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

Matrix Discrepancy

Given a square matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with rows $\boldsymbol{a}_{1\star}, \ldots, \boldsymbol{a}_{n\star} \in \mathbb{R}^n$, let

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- $\operatorname{disc}(\boldsymbol{A}) \geqslant \min_{\boldsymbol{x} \in \{\pm 1\}^n} \sqrt{\frac{\sum_{i=1}^n \langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle^2}{n}} \geqslant \min_{\boldsymbol{x} \neq \boldsymbol{0}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sigma_{\min}(\boldsymbol{A})$
- When $\boldsymbol{A}$ is Hadamard $\frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sqrt{n} = \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$ so

$$\exists \boldsymbol{A} : \operatorname{disc}(\boldsymbol{A}) \geqslant \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- Can we improve the upper bound?

Question

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- Can we improve the upper bound?

Question

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- Can we improve the upper bound?
- Answer: no

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \\ \star \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 1 \end{bmatrix}$$

Question

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- Can we improve the upper bound?
- Answer: no

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ \color{purple}{1} & \color{purple}{-1} & \color{purple}{1} & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \color{purple}{1} \\ \color{purple}{-1} \\ \color{purple}{1} \\ \star \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ \color{purple}{3} \\ 1 \end{bmatrix}$$

Question

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- Can we improve the upper bound?
- Answer: no

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \\ \star \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 1 \end{bmatrix}$$

- Here $\mathrm{disc}(\boldsymbol{A}) = \log 2n = \sqrt{\log 2n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$

Question

- For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$$

- Can we improve the upper bound?
- Answer: no

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \\ \star \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 3 \\ 1 \end{bmatrix}$$

- Here $\operatorname{disc}(\boldsymbol{A}) = \log 2n = \sqrt{\log 2n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_{i\star}\|_2$
- Change the question: here the columns have large norm!

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with <u>columns</u> $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with <u>columns</u> $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

- Hadamard matrices still give a nice lower bound:

$$\exists \boldsymbol{A} : \mathrm{disc}(\boldsymbol{A}) \geqslant \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with <u>columns</u> $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

- Hadamard matrices still give a nice lower bound:

$$\exists \boldsymbol{A} : \operatorname{disc}(\boldsymbol{A}) \geqslant \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$$

- Random $\boldsymbol{x} \sim \{\pm 1\}^n$?

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty$$

- Hadamard matrices still give a nice lower bound:

$$\exists \boldsymbol{A} : \mathrm{disc}(\boldsymbol{A}) \geqslant \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$$

- Random $\boldsymbol{x} \sim \{\pm 1\}^n$? $\mathbb{E}_{\boldsymbol{x} \in \{\pm 1\}^n} [\|\boldsymbol{A}\boldsymbol{x}\|_\infty] \geqslant \Omega(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty$$

- Hadamard matrices still give a nice lower bound:

$$\exists \boldsymbol{A} : \operatorname{disc}(\boldsymbol{A}) \geqslant \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$$

- Random $\boldsymbol{x} \sim \{\pm 1\}^n$? $\mathbb{E}_{\boldsymbol{x} \in \{\pm 1\}^n} [\|\boldsymbol{A}\boldsymbol{x}\|_\infty] \geqslant \Omega(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

- Not hard to show for any $\boldsymbol{A}$: $\mathbb{E}_{\boldsymbol{x} \in \{\pm 1\}^n} [\|\boldsymbol{A}\boldsymbol{x}\|_\infty] \leqslant \sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Banaszczyk ('98)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Banaszczyk ('98)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$

(Bansal-Dadush-Garg, FOCS '16)

We can find $\boldsymbol{x} \in \{\pm 1\}^n$ such that $\|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$ in polynomial time

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty$$

Banaszczyk ('98)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$

(Bansal-Dadush-Garg, FOCS '16)

We can find $\boldsymbol{x} \in \{\pm 1\}^n$ such that $\|\boldsymbol{A}\boldsymbol{x}\|_\infty \leqslant O(\sqrt{\log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$ in polynomial time

Komlós Conjecture ('85)

Is it true that $\mathrm{disc}(\boldsymbol{A}) \leqslant O(1) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$?

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with <u>columns</u> $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Komlós Conjecture ('85)

Is it true that $\operatorname{disc}(\boldsymbol{A}) \leqslant O(1) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$?

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\text{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^n x_i \boldsymbol{a}_i \Big\|_\infty$$

Komlós Conjecture ('85)

Is it true that $\text{disc}(\boldsymbol{A}) \leqslant O(1) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$?

- Too hard! Maybe easier if increase upper bound: $\|\boldsymbol{a}_i\|_2 \leqslant \sqrt{n} \cdot \|\boldsymbol{a}_i\|_\infty$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Komlós Conjecture ('85)

Is it true that $\mathrm{disc}(\boldsymbol{A}) \leqslant O(1) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$?

- Too hard! Maybe easier if increase upper bound: $\|\boldsymbol{a}_i\|_2 \leqslant \sqrt{n} \cdot \|\boldsymbol{a}_i\|_\infty$
- Random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives $\mathrm{disc}(\boldsymbol{A}) \leqslant O(\sqrt{n \log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \Big\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \Big\|_\infty$$

Komlós Conjecture ('85)

Is it true that $\mathrm{disc}(\boldsymbol{A}) \leqslant O(1) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_2$?

- Too hard! Maybe easier if increase upper bound: $\|\boldsymbol{a}_i\|_2 \leqslant \sqrt{n} \cdot \|\boldsymbol{a}_i\|_\infty$
- Random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives $\mathrm{disc}(\boldsymbol{A}) \leqslant O(\sqrt{n \log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

Spencer's "Six Standard Devations Suffice" Theorem ('85)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant 6\sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\operatorname{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

Spencer's "Six Standard Devations Suffice" Theorem ('85)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\operatorname{disc}(\boldsymbol{A}) \leqslant 6\sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

Spencer's "Six Standard Devations Suffice" Theorem ('85)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant 6\sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

- Write $K = \{\boldsymbol{x} \in \mathbb{R}^n : |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle| \leqslant 6\sqrt{n} \; \forall i \in [n]\} = \cap_{i \in [n]} K_i$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^n x_i \boldsymbol{a}_i \right\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

Spencer's "Six Standard Devations Suffice" Theorem ('85)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant 6\sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

- Write $K = \{\boldsymbol{x} \in \mathbb{R}^n : |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle| \leqslant 6\sqrt{n}\ \forall i \in [n]\} = \cap_{i \in [n]} K_i$
- Let $\gamma_n(K) = \mathbb{P}[\text{gaussian} \in K]$ and bound $\gamma_n(K) \geqslant \prod_{i \in [n]} \gamma_n(K_i)$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

Spencer's "Six Standard Devations Suffice" Theorem ('85)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant 6\sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

- Write $K = \{\boldsymbol{x} \in \mathbb{R}^n : |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle| \leqslant 6\sqrt{n}\ \forall i \in [n]\} = \cap_{i \in [n]} K_i$
- Let $\gamma_n(K) = \mathbb{P}[\text{gaussian} \in K]$ and bound $\gamma_n(K) \geqslant \prod_{i \in [n]} \gamma_n(K_i)$
- Show $\gamma_n(K_i) \geqslant \Omega(1)$ for all $i \in [n]$ so that $\gamma_n(K) \geqslant 2^{-O(n)}$

Matrix Discrepancy

Given matrix $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ with columns $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n \in \mathbb{R}^n$, define

$$\mathrm{disc}(\boldsymbol{A}) := \min_{\boldsymbol{x} \in \{\pm 1\}^n} \|\boldsymbol{A}\boldsymbol{x}\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \left\| \sum_{i=1}^n x_i \boldsymbol{a}_i \right\|_\infty = \min_{\boldsymbol{x} \in \{\pm 1\}^n} \max_{i \in [n]} |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle|$$

Spencer's "Six Standard Devations Suffice" Theorem ('85)

For any $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ we have $\mathrm{disc}(\boldsymbol{A}) \leqslant 6\sqrt{n} \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

- Write $K = \{\boldsymbol{x} \in \mathbb{R}^n : |\langle \boldsymbol{a}_{i\star}, \boldsymbol{x} \rangle| \leqslant 6\sqrt{n}\ \forall i \in [n]\} = \cap_{i \in [n]} K_i$
- Let $\gamma_n(K) = \mathbb{P}[\text{gaussian} \in K]$ and bound $\gamma_n(K) \geqslant \prod_{i \in [n]} \gamma_n(K_i)$
- Show $\gamma_n(K_i) \geqslant \Omega(1)$ for all $i \in [n]$ so that $\gamma_n(K) \geqslant 2^{-O(n)}$
- Get partial coloring and recurse on subset of columns

Matrix Spencer

Given columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n \in \mathbb{R}^n$:

Spencer's "Six Standard Devations Suffice" Theorem ('85)

There exists $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{a}_i \right\|_\infty \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{a}_i\|_\infty$

Matrix Spencer

Given <u>matrices</u> $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n \in \mathbb{R}^{n \times n}$:

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$?

Matrix Spencer

Given matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n \in \mathbb{R}^{n \times n}$:

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$?

- Random $\boldsymbol{x} \sim \{\pm 1\}^n$ gives $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n \log 2n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$

Matrix Spencer

Given matrices $A_1, \ldots, A_n \in \mathbb{R}^{n \times n}$:

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $x \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i A_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|A_i\|_{op}$?

- Random $x \sim \{\pm 1\}^n$ gives $\left\| \sum_{i=1}^{n} x_i A_i \right\|_{op} \leqslant O(\sqrt{n \log 2n}) \cdot \max_{i \in [n]} \|A_i\|_{op}$
- Matrix concentration: $\left\| \sum_{i=1}^{n} x_i A_i \right\|_{op} \leqslant \underbrace{O(\sqrt{\log 2n})}_{\text{tight!}} \cdot \max_{i \in [n]} \left\| \sum_{i=1}^{n} A_i^2 \right\|_{op}^{1/2}$

Example

Take n matrices $(\boldsymbol{A}_i)_{i\in[n]}$ with $(\boldsymbol{A}_i)_{u,v} = \begin{cases} 1 & u - v \equiv i \mod n \\ 0 & \text{otherwise} \end{cases}$

Example

Take n matrices $(\boldsymbol{A}_i)_{i\in[n]}$ with $(\boldsymbol{A}_i)_{u,v} = \begin{cases} 1 & u - v \equiv i \mod n \\ 0 & \text{otherwise} \end{cases}$

$$\left\| \sum_{i=1}^{11} 1 \cdot \boldsymbol{A}_i \right\|_{\mathrm{op}} = \left\| \begin{matrix} 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \\ 1&1&1&1&1&1&1&1&1&1&1 \end{matrix} \right\|_{\mathrm{op}}$$

$$= 11$$

Example

Take n matrices $(\boldsymbol{A}_i)_{i \in [n]}$ with $(\boldsymbol{A}_i)_{u,v} = \begin{cases} 1 & u - v \equiv i \mod n \\ 0 & \text{otherwise} \end{cases}$

$$\left\| \sum_{i=1}^{11} \left(\frac{i}{11} \right) \cdot \boldsymbol{A}_i \right\|_{\text{op}} = \left\| \begin{matrix} + & + & - & + & + & + & - & - & - & + & - \\ - & + & + & - & + & + & + & - & - & - & + \\ + & - & + & + & - & + & + & + & - & - & - \\ - & + & - & + & + & - & + & + & + & - & - \\ - & - & + & - & + & + & - & + & + & + & - \\ - & - & - & + & - & + & + & - & + & + & + \\ + & - & - & - & + & - & + & + & - & + & + \\ + & + & - & - & - & + & - & + & + & - & + \\ + & + & + & - & - & - & + & - & + & + & - \\ - & + & + & + & - & - & - & + & - & + & + \\ + & - & + & + & + & - & - & - & + & - & + \end{matrix} \right\|_{\text{op}}$$

$$= 1 + \sqrt{11}$$

Matrix Spencer

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$?

Matrix Spencer

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$?

- Can write $K = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op} \right\}$

Matrix Spencer

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$?

- Can write $K = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op} \right\}$
- But this is not a polytope; how to lower bound $\gamma_n(K)$?

Matrix Spencer

Matrix Spencer Conjecture (Zouzias '12, Meka '14)

Does there exist $\boldsymbol{x} \in \{\pm 1\}^n$ with $\left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op}$?

- Can write $K = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{op} \leqslant O(\sqrt{n}) \cdot \max_{i \in [n]} \|\boldsymbol{A}_i\|_{op} \right\}$
- But this is not a polytope; how to lower bound $\gamma_n(K)$?

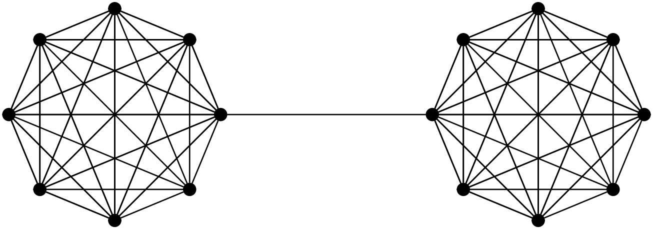
Broader question

How can we get measure lower bounds for more general convex bodies?

Graph Sparsification

Theorem (Batson-Spielman-Srivastava '08)

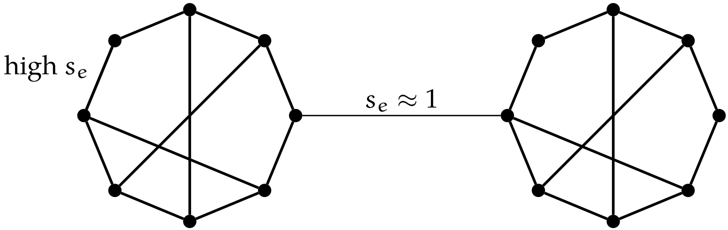
Given a graph $G = (V, E)$, we can find weights $\boldsymbol{s} \in \mathbb{R}^m_{\geqslant 0}$ in poly-time with $|\text{supp}(\boldsymbol{s})| \leqslant O(n/\varepsilon^2)$ so that $|\delta(U)| = (1 \pm \varepsilon) \cdot |\boldsymbol{s}(\delta(U))|$ for every $U \subseteq V$.



Graph Sparsification

Theorem (Batson-Spielman-Srivastava '08)

Given a graph $G = (V, E)$, we can find weights $\boldsymbol{s} \in \mathbb{R}^m_{\geqslant 0}$ in poly-time with $|\mathrm{supp}(\boldsymbol{s})| \leqslant O(n/\varepsilon^2)$ so that $|\delta(U)| = (1 \pm \varepsilon) \cdot |\boldsymbol{s}(\delta(U))|$ for every $U \subseteq V$.

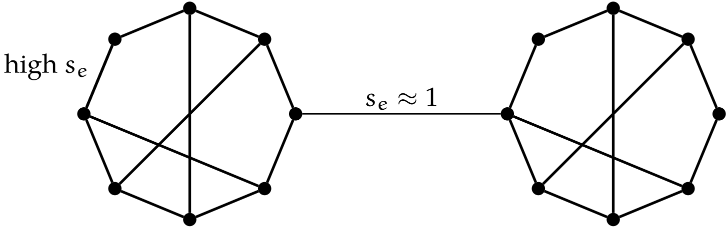


Graph Sparsification

Theorem (Batson-Spielman-Srivastava '08)

Given rank-one $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m \in \mathbb{R}^{n \times n}$ with $\sum_{i=1}^{m} \boldsymbol{A}_i = \boldsymbol{I}_n$, we can find weights $\boldsymbol{s} \in \mathbb{R}^m_{\geqslant 0}$ in poly-time such that $|\mathrm{supp}(\boldsymbol{s})| \leqslant O(n/\varepsilon^2)$ and

$$\Big\| \sum_{i=1}^{m} s_i \boldsymbol{A}_i \Big\|_{\mathrm{op}} \in [1-\varepsilon, 1+\varepsilon]$$



A new algorithm

Given: PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^{m} \boldsymbol{A}_i = \boldsymbol{I}_n$ and $\varepsilon > 0$.

(0) Initialize weights $s_i := 1$ for $i \in [m]$

- Repeat until $\boldsymbol{s}$ sufficiently sparse:

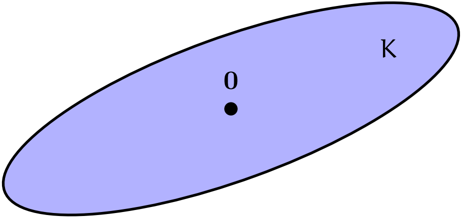
A new algorithm

Given: PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$ and $\varepsilon > 0$.

(0) Initialize weights $s_i := 1$ for $i \in [m]$

- Repeat until $\boldsymbol{s}$ sufficiently sparse:

(1) Consider the convex body $K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$



A new algorithm

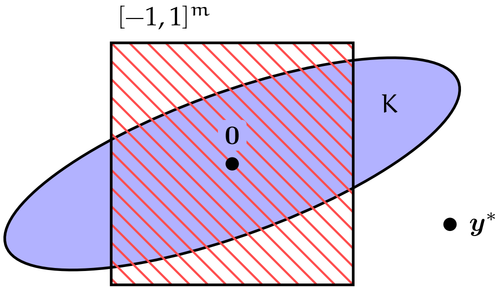
Given: PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$ and $\varepsilon > 0$.

(0) Initialize weights $s_i := 1$ for $i \in [m]$

- Repeat until $\boldsymbol{s}$ sufficiently sparse:

(1) Consider the convex body $K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$

(2) Sample $\boldsymbol{y}^* \sim N(\boldsymbol{0}, \boldsymbol{I}_m)$



A new algorithm

Given: PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$ and $\varepsilon > 0$.

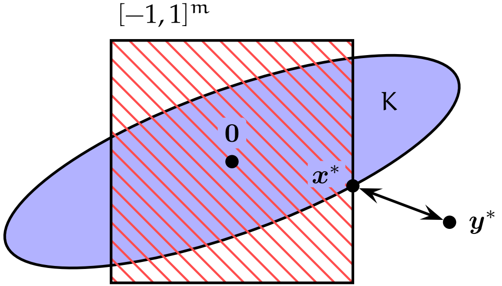
(0) Initialize weights $s_i := 1$ for $i \in [m]$

- Repeat until $\boldsymbol{s}$ sufficiently sparse:

(1) Consider the convex body $K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$

(2) Sample $\boldsymbol{y}^* \sim N(\boldsymbol{0}, \boldsymbol{I}_m)$

(3) Let $\boldsymbol{x}^*$ be the projection of $\boldsymbol{y}^*$ onto $K \cap [-1, 1]^m\}$



A new algorithm

Given: PSD matrices $\boldsymbol{A}_1, \dots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$ and $\varepsilon > 0$.

(0) Initialize weights $s_i := 1$ for $i \in [m]$

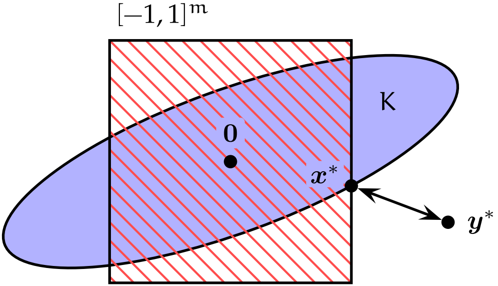
- Repeat until $\boldsymbol{s}$ sufficiently sparse:

(1) Consider the convex body $K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$

(2) Sample $\boldsymbol{y}^* \sim N(\boldsymbol{0}, \boldsymbol{I}_m)$

(3) Let $\boldsymbol{x}^*$ be the projection of $\boldsymbol{y}^*$ onto $K \cap [-1,1]^m\}$

- $\boldsymbol{x}^*$ will have $\Omega(m)$ many $\{-1,1\}$ coordinates with high probability



A new algorithm

Given: PSD matrices $\boldsymbol{A}_1, \dots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$ and $\varepsilon > 0$.

(0) Initialize weights $s_i := 1$ for $i \in [m]$

- Repeat until $\boldsymbol{s}$ sufficiently sparse:

(1) Consider the convex body $K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$

(2) Sample $\boldsymbol{y}^* \sim N(\boldsymbol{0}, \boldsymbol{I}_m)$

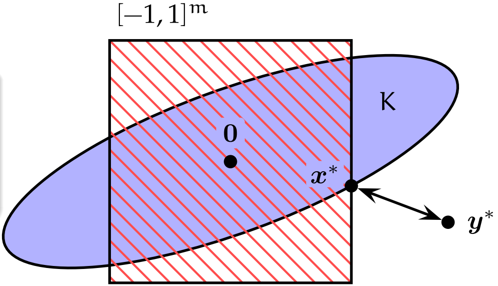
(3) Let $\boldsymbol{x}^*$ be the projection of $\boldsymbol{y}^*$ onto $K \cap [-1,1]^m\}$

- $\boldsymbol{x}^*$ will have $\Omega(m)$ many $\{-1,1\}$ coordinates with high probability

(4) Update weights $s_i \leftarrow s_i(1 + x_i^*)$.

(R., Rothvoss '19)

$O(\log m)$ iterations suffice and output $1 \pm O(\varepsilon)$ sparsifier with high probability.



Main technical results

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$\mathsf{K} = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{\text{op}} \leqslant \sqrt{n/m} \right\}$$

Main technical results

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$\mathsf{K} = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \Big\| \sum_{i=1}^m x_i \boldsymbol{A}_i \Big\|_{\mathrm{op}} \leqslant \sqrt{n/m} \right\}$$

Question: Do we have $\gamma_m(\mathsf{K}) \geqslant 2^{-O(m)}$?

Main technical results

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{\mathrm{op}} \leqslant \sqrt{n/m} \right\}$$

Question: Do we have $\gamma_m(K) \geqslant 2^{-O(m)}$? We don't know!

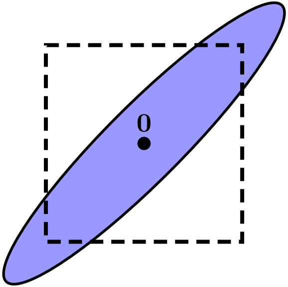
Main technical results

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

Question: Do we have $\gamma_m(K) \geqslant 2^{-O(m)}$? We don't know!

- Weaker condition to get $\Omega(m)$ coordinates $\boldsymbol{x}_i \in \{-1, 1\}$:



Main technical results

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

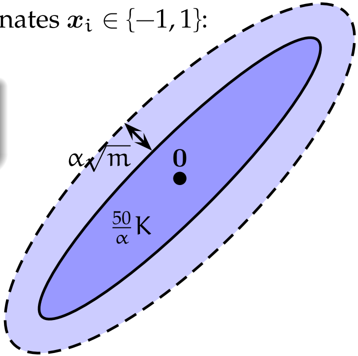
$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

Question: Do we have $\gamma_m(K) \geqslant 2^{-O(m)}$? We don't know!

- Weaker condition to get $\Omega(m)$ coordinates $\boldsymbol{x}_i \in \{-1, 1\}$:

Theorem (R., Rothvoss '19)

$\gamma_m(\frac{50}{\alpha}K + \alpha\sqrt{m}B_2^m) \geqslant \frac{1}{2} \ \forall \alpha \in (0, 1)$



Main technical results

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^m x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

Question: Do we have $\gamma_m(K) \geqslant 2^{-O(m)}$? We don't know!

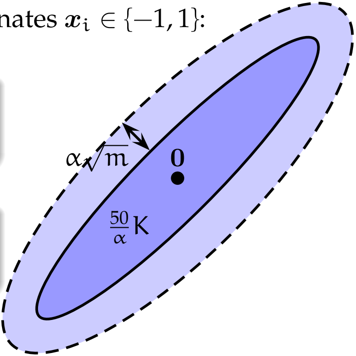
- Weaker condition to get $\Omega(m)$ coordinates $\boldsymbol{x}_i \in \{-1, 1\}$:

Theorem (R., Rothvoss '19)

$\gamma_m(\frac{50}{\alpha}K + \alpha\sqrt{m}B_2^m) \geqslant \frac{1}{2}\ \forall \alpha \in (0,1)$

Corollary (R., Rothvoss '19)

The **mean width** of K is $\Omega(\sqrt{m})$



Corollary: Spectral Partial Coloring

Corollary (R., Rothvoss '19)

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^m \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \Big\| \sum_{i=1}^m x_i \boldsymbol{A}_i \Big\|_{op} \leqslant \sqrt{n/m} \right\}$$

We can find a point $\boldsymbol{x}^* \in 500K \cap [-1,1]^m$ with $\geqslant \frac{m}{9000}$ indices $\boldsymbol{x}_i^* \in \{-1, 1\}$

Proof idea

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^{m} x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

$$B_2^m = \{\boldsymbol{x} \in \mathbb{R}^m : \|\boldsymbol{x}\|_2 \leqslant 1\}$$

Theorem (R., Rothvoss '19)

We have $\gamma_m(\frac{50}{\alpha}K + \alpha\sqrt{m}B_2^m) \geqslant \frac{1}{2}$ for any choice of $\alpha \in (0, 1)$

- That is:

  a Gaussian $\sim N(\boldsymbol{0}, \boldsymbol{I}_m)$ is close to $\frac{50}{\alpha}K$ with prob $1/2$

Proof idea

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^{m} x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

$$B_2^m = \{ \boldsymbol{x} \in \mathbb{R}^m : \|\boldsymbol{x}\|_2 \leqslant 1 \}$$

Theorem (R., Rothvoss '19)

We have $\gamma_m(\frac{50}{\alpha}K + \alpha\sqrt{m}B_2^m) \geqslant \frac{1}{2}$ for any choice of $\alpha \in (0, 1)$

- That is:

  a Gaussian $\sim N(\mathbf{0}, \boldsymbol{I}_m)$ is close to $\frac{50}{\alpha}K$ with prob $1/2$
- Equivalently:

  construct random $\boldsymbol{x}$ with $\mathbb{P}[\boldsymbol{x} \in \frac{50}{\alpha}K$ and $\boldsymbol{x}$ close to Gaussian$] \geqslant 1/2$

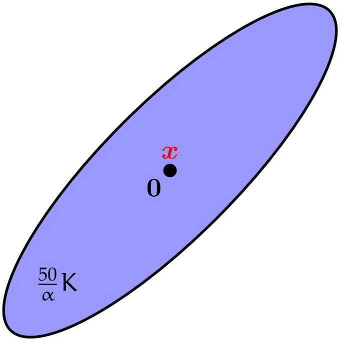
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



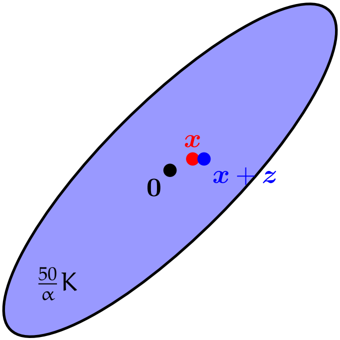
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

   (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

   (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



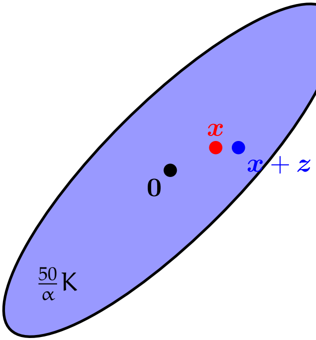
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



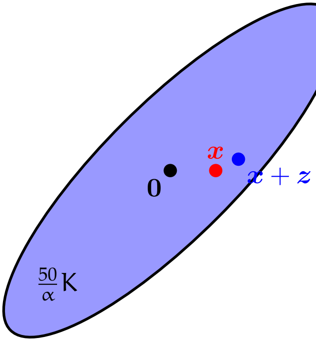
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \boldsymbol{0}, \boldsymbol{z} := \boldsymbol{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\boldsymbol{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\boldsymbol{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{X})$



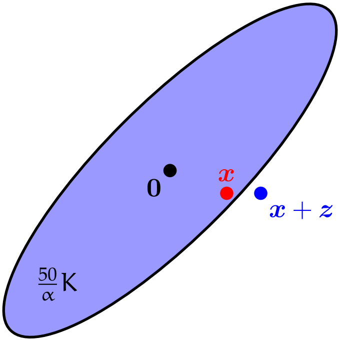
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



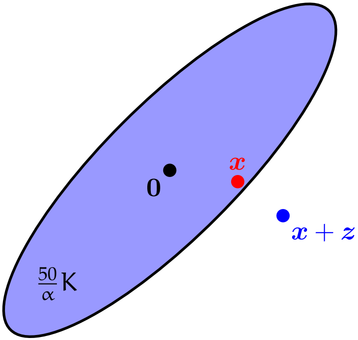
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



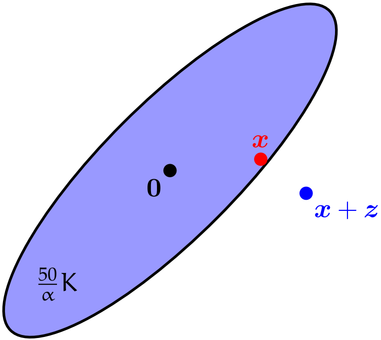
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



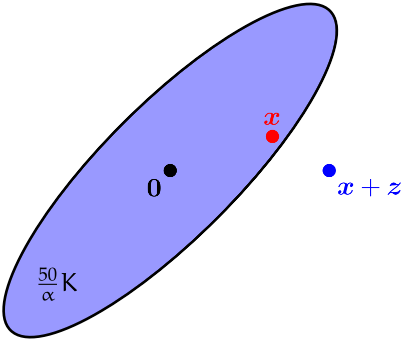
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \boldsymbol{0}, \boldsymbol{z} := \boldsymbol{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\boldsymbol{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1 - \alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\boldsymbol{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\boldsymbol{0}, \boldsymbol{I}_m - \boldsymbol{X})$



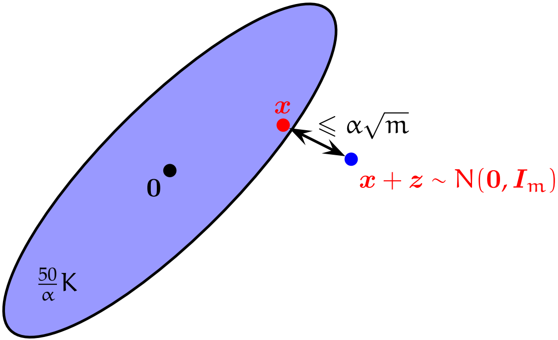
A coupling argument

(1) Set $\delta :=$ tiny step size and $\boldsymbol{x} := \mathbf{0}, \boldsymbol{z} := \mathbf{0}$

(2) FOR $T := \delta^{-2}$ iterations:

  (3) Find $\mathbf{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

  (4) Update $\boldsymbol{x} := \boldsymbol{x} + \delta \cdot N(\mathbf{0}, \boldsymbol{X})$ and $\boldsymbol{z} := \boldsymbol{z} + \delta \cdot N(\mathbf{0}, \boldsymbol{I}_m - \boldsymbol{X})$



One update step

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^{m} x_i \boldsymbol{A}_i \right\|_{\mathrm{op}} \leqslant \sqrt{n/m} \right\}$$

- Want $\boldsymbol{x} \in K$: can reduce to upper bounding $\lambda_{\max}(\sum_{i=1}^{m} x_i \boldsymbol{A}_i)$
- Use **potential function**

$$\boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \quad \text{and} \quad \Phi(\boldsymbol{x}) := \mathrm{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

One update step

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^{m} x_i \boldsymbol{A}_i \right\|_{\mathrm{op}} \leqslant \sqrt{n/m} \right\}$$

- Want $\boldsymbol{x} \in K$: can reduce to upper bounding $\lambda_{\max}(\sum_{i=1}^m x_i \boldsymbol{A}_i)$
- Use **potential function**

$$\boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \quad \text{and} \quad \Phi(\boldsymbol{x}) := \mathrm{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

- Keep $\Phi(\boldsymbol{x})$ bounded above so that at each step

$$\boldsymbol{A}(\boldsymbol{x}) \succ 0 \implies \lambda_{\max}\left(\sum_{i=1}^{m} x_i \boldsymbol{A}_i \right) < C + D\|\boldsymbol{x}\|_2^2$$

One update step

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^{m} x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

- Want $\boldsymbol{x} \in K$: can reduce to upper bounding $\lambda_{max}(\sum_{i=1}^{m} x_i \boldsymbol{A}_i)$
- Use **potential function**

$$\boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \quad \text{and} \quad \Phi(\boldsymbol{x}) := \mathrm{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

- Keep $\Phi(\boldsymbol{x})$ bounded above so that at each step

$$\boldsymbol{A}(\boldsymbol{x}) \succ 0 \implies \lambda_{max}\left(\sum_{i=1}^{m} x_i \boldsymbol{A}_i \right) < C + D\|\boldsymbol{x}\|_2^2$$

Main Lemma

If $\Phi(\boldsymbol{x}) \leqslant \frac{D m^2 \alpha^2}{10}$, there is $\boldsymbol{0} \prec \boldsymbol{X} \prec \boldsymbol{I}_m$ with $\mathrm{Tr}[\boldsymbol{X}] \geqslant (1 - \alpha^2)m$ so that

$$\mathop{\mathbb{E}}_{\boldsymbol{y} \sim N(\boldsymbol{0}, \boldsymbol{X})} [\Phi(\boldsymbol{x} + \delta \boldsymbol{y})] \leqslant \Phi(\boldsymbol{x})$$

One update step

- Recall

$$\boldsymbol{A} := \boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \text{ and } \Phi(\boldsymbol{x}) := \operatorname{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

One update step

- Recall

$$\boldsymbol{A} := \boldsymbol{A}(\boldsymbol{x}) := (\mathsf{C} + \mathsf{D}\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \text{ and } \Phi(\boldsymbol{x}) := \operatorname{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

- Pick $\boldsymbol{X}$: $\boldsymbol{y} \sim N(\boldsymbol{0}, \boldsymbol{X})$ is orthogonal to $\boldsymbol{x}$ and any linear term

One update step

- Recall

$$\boldsymbol{A} := \boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \text{ and } \Phi(\boldsymbol{x}) := \operatorname{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

- Pick $\boldsymbol{X}$: $\boldsymbol{y} \sim N(\mathbf{0}, \boldsymbol{X})$ is orthogonal to $\boldsymbol{x}$ and any linear term
- Set $\boldsymbol{B} := \sum_{i=1}^{m} y_i \boldsymbol{A}_i - \delta D\|\boldsymbol{y}\|_2^2 \boldsymbol{I}_n$.
- Change of potential function is

$$\operatorname{tr}[(\boldsymbol{A} - \delta\boldsymbol{B})^{-1}] - \operatorname{tr}[\boldsymbol{A}^{-1}]$$

One update step

- Recall

$$\boldsymbol{A} := \boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \text{ and } \Phi(\boldsymbol{x}) := \mathrm{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

- Pick $\boldsymbol{X}$: $\boldsymbol{y} \sim N(\boldsymbol{0}, \boldsymbol{X})$ is orthogonal to $\boldsymbol{x}$ and any linear term
- Set $\boldsymbol{B} := \sum_{i=1}^{m} y_i \boldsymbol{A}_i - \delta D\|\boldsymbol{y}\|_2^2 \boldsymbol{I}_n$.
- Change of potential function is

$$\begin{aligned} & \mathrm{tr}[(\boldsymbol{A} - \delta\boldsymbol{B})^{-1}] - \mathrm{tr}[\boldsymbol{A}^{-1}] \\ \overset{\text{Taylor}}{\leqslant} \quad & \delta \cdot \mathrm{tr}[\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}] + 2\delta^2 \cdot \mathrm{tr}[\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}] \end{aligned}$$

One update step

- Recall

$$\boldsymbol{A} := \boldsymbol{A}(\boldsymbol{x}) := (C + D\|\boldsymbol{x}\|_2^2) \cdot \boldsymbol{I}_n - \sum_{i=1}^{m} x_i \boldsymbol{A}_i \text{ and } \Phi(\boldsymbol{x}) := \mathrm{tr}[\boldsymbol{A}(\boldsymbol{x})^{-1}]$$

- Pick $\boldsymbol{X}$: $\boldsymbol{y} \sim N(\boldsymbol{0}, \boldsymbol{X})$ is orthogonal to $\boldsymbol{x}$ and any linear term
- Set $\boldsymbol{B} := \sum_{i=1}^{m} y_i \boldsymbol{A}_i - \delta D\|\boldsymbol{y}\|_2^2 \boldsymbol{I}_n$.
- Change of potential function is

$$\begin{aligned} & \mathrm{tr}[(\boldsymbol{A} - \delta\boldsymbol{B})^{-1}] - \mathrm{tr}[\boldsymbol{A}^{-1}] \\ \overset{\text{Taylor}}{\leqslant} \; & \delta \cdot \mathrm{tr}[\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}] + 2\delta^2 \cdot \mathrm{tr}[\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{A}^{-1}] \\ \approx \; & -\delta^2 D\|\boldsymbol{y}\|_2^2 \mathrm{tr}[\boldsymbol{A}^{-2}] + 2\delta^2 \mathrm{tr}\Big[\boldsymbol{A}^{-1}\Big(\sum_{i=1}^{m} y_i \boldsymbol{A}_i\Big)\boldsymbol{A}^{-1}\Big(\sum_{j=1}^{m} y_j \boldsymbol{A}_j\Big)\boldsymbol{A}^{-1}\Big] \end{aligned}$$

One update step

Key Claim

$$\mathop{\mathbb{E}}_{\boldsymbol{y}\sim N(\boldsymbol{0},\boldsymbol{X})}\left[\operatorname{tr}\left[\boldsymbol{A}^{-1}\left(\sum_{i=1}^{m} y_i\boldsymbol{A}_i\right)\boldsymbol{A}^{-1}\left(\sum_{j=1}^{m} y_j\boldsymbol{A}_j\right)\boldsymbol{A}^{-1}\right]\right] \leqslant \frac{2}{\alpha^2 m}\operatorname{tr}[\boldsymbol{A}^{-1}]\operatorname{tr}[\boldsymbol{A}^{-2}]$$

One update step

Key Claim

$$\mathop{\mathbb{E}}_{\boldsymbol{y}\sim N(\boldsymbol{0},\boldsymbol{X})}\left[\operatorname{tr}\left[\boldsymbol{A}^{-1}\left(\sum_{i=1}^{m} y_i \boldsymbol{A}_i\right)\boldsymbol{A}^{-1}\left(\sum_{j=1}^{m} y_j \boldsymbol{A}_j\right)\boldsymbol{A}^{-1}\right]\right] \leqslant \frac{2}{\alpha^2 m}\operatorname{tr}[\boldsymbol{A}^{-1}]\operatorname{tr}[\boldsymbol{A}^{-2}]$$

- Key idea: Pick $y \perp \boldsymbol{e}_i$ for i such that $\operatorname{tr}[\boldsymbol{A}^{-1}\boldsymbol{A}_i] > \frac{2}{\alpha^2 m}\operatorname{tr}[\boldsymbol{A}^{-1}]$

One update step

Key Claim

$$\mathop{\mathbb{E}}_{\boldsymbol{y}\sim N(\mathbf{0},\boldsymbol{X})}\left[\operatorname{tr}\left[\boldsymbol{A}^{-1}\left(\sum_{i=1}^{m} y_i \boldsymbol{A}_i\right)\boldsymbol{A}^{-1}\left(\sum_{j=1}^{m} y_j \boldsymbol{A}_j\right)\boldsymbol{A}^{-1}\right]\right] \leqslant \frac{2}{\alpha^2 m}\operatorname{tr}[\boldsymbol{A}^{-1}]\operatorname{tr}[\boldsymbol{A}^{-2}]$$

- Key idea: Pick $y \perp \boldsymbol{e}_i$ for i such that $\operatorname{tr}[\boldsymbol{A}^{-1}\boldsymbol{A}_i] > \frac{2}{\alpha^2 m}\operatorname{tr}[\boldsymbol{A}^{-1}]$
- By Markov: at most $\frac{\alpha^2 m}{2}$ such constraints

One Update Step

- Putting everything together..

$$
\begin{aligned}
& \mathbb{E}[\mathrm{tr}[(\boldsymbol{A} - \delta \boldsymbol{B})^{-1}]] - \mathrm{tr}[\boldsymbol{A}^{-1}] \\
\leqslant & -\delta^2 D\, \mathbb{E}[\|\boldsymbol{y}\|_2^2] \mathrm{tr}[\boldsymbol{A}^{-2}] + 2\delta^2\, \mathbb{E}\left[\mathrm{tr}\left[\boldsymbol{A}^{-1}\Big(\sum_{i=1}^{m} y_i \boldsymbol{A}_i\Big)\boldsymbol{A}^{-1}\Big(\sum_{i=1}^{m} y_i \boldsymbol{A}_i\Big)\boldsymbol{A}^{-1}\right]\right] \\
\leqslant & \quad \delta^2 \cdot \mathrm{tr}[\boldsymbol{A}^{-2}] \cdot \Big(-0.4Dm + \frac{4}{\alpha^2 m} \mathrm{tr}[\boldsymbol{A}^{-1}]\Big) \leqslant 0
\end{aligned}
$$

One Update Step

- Putting everything together..

$$
\begin{aligned}
&\mathbb{E}[\mathrm{tr}[(\boldsymbol{A}-\delta\boldsymbol{B})^{-1}]] - \mathrm{tr}[\boldsymbol{A}^{-1}] \\
\leqslant\, & -\delta^2 D\,\mathbb{E}[\|\boldsymbol{y}\|_2^2]\mathrm{tr}[\boldsymbol{A}^{-2}] + 2\delta^2\,\mathbb{E}\left[\mathrm{tr}\left[\boldsymbol{A}^{-1}\Big(\sum_{i=1}^{m} y_i\boldsymbol{A}_i\Big)\boldsymbol{A}^{-1}\Big(\sum_{i=1}^{m} y_i\boldsymbol{A}_i\Big)\boldsymbol{A}^{-1}\right]\right] \\
\leqslant\, & \delta^2\cdot\mathrm{tr}[\boldsymbol{A}^{-2}]\cdot\Big(-0.4Dm + \frac{4}{\alpha^2 m}\mathrm{tr}[\boldsymbol{A}^{-1}]\Big) \leqslant 0
\end{aligned}
$$

- Only $\leqslant \alpha^2 m$ constraints needed: $\mathrm{tr}[\boldsymbol{X}] \geqslant (1-\alpha^2)m$

Conclusion

Given PSD matrices $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m$ with $\sum_{i=1}^{m} \boldsymbol{A}_i = \boldsymbol{I}_n$, define the set

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^m \mid \left\| \sum_{i=1}^{m} x_i \boldsymbol{A}_i \right\|_{op} \leqslant \sqrt{n/m} \right\}$$

Theorem (R., Rothvoss '19)

We have $\gamma_m(\frac{50}{\alpha} K + \alpha\sqrt{m} B_2^m) \geqslant \frac{1}{2}\ \forall \alpha \in (0,1)$

Corollary (R., Rothvoss '19)

The **mean width** of K is $\Omega(\sqrt{m})$

- Do we have $\gamma_m(K) \geqslant 2^{-O(m)}$? Does the Theorem already imply this?

More Open Problems

- Can we show similar measure lower bounds for Matrix Spencer?
- Between Kómlos and Spencer:

$$\text{Given } p \geqslant 2\text{, show } \mathrm{disc}(\boldsymbol{A}) \leqslant O(\sqrt{n}) \text{ for } \|\boldsymbol{a}_i\|_p \leqslant n^{1/p}$$

(the same as Kómlos for $p = 2$, the same as Spencer for $p = \infty$)

- Circulant Kómlos:

$$\text{Given unit } \boldsymbol{v} \in \mathbb{R}^n\text{, find } \boldsymbol{x} \in \{\pm 1\}^n \text{ with } \max_{j \in [n]} \left| \sum_{i=1}^{n} x_i v_{i+j} \right| \leqslant O(1)$$

More Open Problems

- Can we show similar measure lower bounds for Matrix Spencer?
- Between Kómlos and Spencer:

$$\text{Given } p \geqslant 2, \text{ show } \operatorname{disc}(\boldsymbol{A}) \leqslant O(\sqrt{n}) \text{ for } \|\boldsymbol{a}_i\|_p \leqslant n^{1/p}$$

(the same as Kómlos for $p = 2$, the same as Spencer for $p = \infty$)

- Circulant Kómlos:

$$\text{Given unit } \boldsymbol{v} \in \mathbb{R}^n, \text{ find } \boldsymbol{x} \in \{\pm 1\}^n \text{ with } \max_{j \in [n]} \left| \sum_{i=1}^{n} x_i v_{i+j} \right| \leqslant O(1)$$

Thanks for your attention!

give me anonymous feedback:
www.admonymous.co/voreis