A New Framework for Matrix Discrepancy: Partial Coloring Bounds via Mirror Descent

Victor Reis

STOC

June 2022

Joint work with Daniel Dadush (CWI) and Haotian Jiang (UW)

UNIVERSITY *of* WASHINGTON

Talk outline

- Introduction to vector balancing

Talk outline

- Introduction to vector balancing
- Measure bounds via polar covering

Talk outline

- Introduction to vector balancing
- Measure bounds via polar covering
- Warmup: Spencer setting

Talk outline

- Introduction to vector balancing
- Measure bounds via polar covering
- Warmup: Spencer setting
- Covering via mirror descent

Talk outline

- Introduction to vector balancing
- Measure bounds via polar covering
- Warmup: Spencer setting
- Covering via mirror descent
- Quantum relative entropy net construction

Talk outline

- Introduction to vector balancing
- Measure bounds via polar covering
- Warmup: Spencer setting
- Covering via mirror descent
- Quantum relative entropy net construction
- Open problems

Introduction to vector balancing

- Given $A, B \subseteq \mathbb{R}^d$, $\mathrm{vb}_n(A, B) :=$ smallest $C > 0$ so that for any $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$ there are signs $\boldsymbol{x} \in \{\pm 1\}^n$ with

$$\sum_{i=1}^{n} x_i \boldsymbol{v}_i \in C \cdot B.$$

Introduction to vector balancing

- Given $A, B \subseteq \mathbb{R}^d$, $\mathrm{vb}_n(A, B) :=$ smallest $C > 0$ so that for any $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$ there are signs $\boldsymbol{x} \in \{\pm 1\}^n$ with
$$\sum_{i=1}^{n} x_i \boldsymbol{v}_i \in C \cdot B.$$
- Applications to scheduling [BRS22] and RCTs [HSSZ19]

Introduction to vector balancing

- Given $A, B \subseteq \mathbb{R}^d$, $\mathrm{vb}_n(A, B) :=$ smallest $C > 0$ so that for any $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$ there are signs $\boldsymbol{x} \in \{\pm 1\}^n$ with

$$\sum_{i=1}^{n} x_i \boldsymbol{v}_i \in C \cdot B.$$

- Applications to scheduling [BRS22] and RCTs [HSSZ19]

- $\mathrm{vb}_n(B_2^n, B_2^n) \leqslant \sqrt{n}$ by taking uniformly random $\boldsymbol{x} \sim \{\pm 1\}^n$:

$$\mathbb{E}\left[\|x_1 \boldsymbol{v}_1 + \cdots + x_n \boldsymbol{v}_n\|_2^2\right] = \sum_{i=1}^{n} \|\boldsymbol{v}_i\|_2^2 \leqslant n$$

Introduction to vector balancing

- Given $A, B \subseteq \mathbb{R}^d$, $\mathrm{vb}_n(A, B) :=$ smallest $C > 0$ so that for any $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$ there are signs $\boldsymbol{x} \in \{\pm 1\}^n$ with

$$\sum_{i=1}^{n} x_i \boldsymbol{v}_i \in C \cdot B.$$

- Applications to scheduling [BRS22] and RCTs [HSSZ19]

- $\mathrm{vb}_n(B_2^n, B_2^n) \leqslant \sqrt{n}$ by taking uniformly random $\boldsymbol{x} \sim \{\pm 1\}^n$:

$$\mathbb{E}\left[\|x_1 \boldsymbol{v}_1 + \cdots + x_n \boldsymbol{v}_n\|_2^2\right] = \sum_{i=1}^{n} \|\boldsymbol{v}_i\|_2^2 \leqslant n$$

- Linear algebra shows $\mathrm{vb}_n(A, B) \leqslant 2 \cdot \mathrm{vb}_d(A, B)$: w.l.o.g. $n \leqslant d$.

Introduction to vector balancing

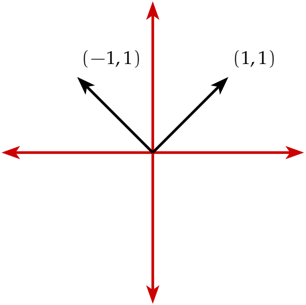
- $\mathrm{vb}_n(B_2^n, B_2^n) = \sqrt{n}$. What about $\mathrm{vb}_n(B_\infty^n, B_\infty^n)$?

Introduction to vector balancing

- $\mathrm{vb}_n(B_2^n, B_2^n) = \sqrt{n}$. What about $\mathrm{vb}_n(B_\infty^n, B_\infty^n)$?
- Random signs give $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \lesssim \sqrt{n \log(2n)}$

Introduction to vector balancing

- $\mathrm{vb}_n(B_2^n, B_2^n) = \sqrt{n}$. What about $\mathrm{vb}_n(B_\infty^n, B_\infty^n)$?
- Random signs give $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \lesssim \sqrt{n \log(2n)}$
- Cannot hope for $\sqrt{n}$: $\mathrm{vb}_2(B_\infty^2, B_\infty^2) = 2$



Introduction to vector balancing

- $\mathrm{vb}_n(B_2^n, B_2^n) = \sqrt{n}$. What about $\mathrm{vb}_n(B_\infty^n, B_\infty^n)$?
- Random signs give $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \lesssim \sqrt{n \log(2n)}$
- Cannot hope for $\sqrt{n}$: $C_2(B_\infty^2, B_\infty^2) = 2$

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

$\mathrm{vb}_n(B_\infty^n, B_\infty^n) \leqslant 6\sqrt{n}$ and more generally $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$.

Introduction to vector balancing

- $\mathrm{vb}_n(B_2^n, B_2^n) = \sqrt{n}$. What about $\mathrm{vb}_n(B_\infty^n, B_\infty^n)$?
- Random signs give $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \lesssim \sqrt{n \log(2n)}$
- Cannot hope for $\sqrt{n}$: $C_2(B_\infty^2, B_\infty^2) = 2$

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

$\mathrm{vb}_n(B_\infty^n, B_\infty^n) \leqslant 6\sqrt{n}$ and more generally $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$.

- Tight for Hadamard columns: $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \geqslant \sqrt{n}$

Introduction to vector balancing

- $\mathrm{vb}_n(B_2^n, B_2^n) = \sqrt{n}$. What about $\mathrm{vb}_n(B_\infty^n, B_\infty^n)$?
- Random signs give $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \lesssim \sqrt{n \log(2n)}$
- Cannot hope for $\sqrt{n}$: $C_2(B_\infty^2, B_\infty^2) = 2$

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

$\mathrm{vb}_n(B_\infty^n, B_\infty^n) \leqslant 6\sqrt{n}$ and more generally $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$.

- Tight for Hadamard columns: $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \geqslant \sqrt{n}$
- Conjecture (Spielman) $\mathrm{vb}_n(B_\infty^n, B_\infty^n) \leqslant 2\sqrt{n}$

Matrix Spencer Conjecture

- $\mathrm{vb}_n(B_2^m, B_2^m) \leqslant \sqrt{n}$ and $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$

Matrix Spencer Conjecture

- $\mathrm{vb}_n(B_2^m, B_2^m) \leqslant \sqrt{n}$ and $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$
- Generalization for matrices:

$$S_p^m := \{\text{symmetric } \boldsymbol{A} \in \mathbb{R}^{m \times m} : \|\lambda(\boldsymbol{A})\|_p \leqslant 1\}$$

Matrix Spencer Conjecture

- $\mathrm{vb}_n(B_2^m, B_2^m) \leqslant \sqrt{n}$ and $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$
- Generalization for matrices:

$$S_p^m := \{\text{symmetric } \boldsymbol{A} \in \mathbb{R}^{m \times m} : \|\lambda(\boldsymbol{A})\|_p \leqslant 1\}$$

- For diagonal matrices $\|\boldsymbol{A}\|_{S_p^m} = \|\mathrm{diag}(\boldsymbol{A})\|_p$

Matrix Spencer Conjecture

- $\mathrm{vb}_n(B_2^m, B_2^m) \leqslant \sqrt{n}$ and $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$
- Generalization for matrices:

$$S_p^m := \{\text{symmetric } \boldsymbol{A} \in \mathbb{R}^{m \times m} : \|\lambda(\boldsymbol{A})\|_p \leqslant 1\}$$

- For diagonal matrices $\|\boldsymbol{A}\|_{S_p^m} = \|\mathrm{diag}(\boldsymbol{A})\|_p$
- $\mathrm{vb}_n(S_2^m, S_2^m) \leqslant \sqrt{n}$

Matrix Spencer Conjecture

- $\mathrm{vb}_n(B_2^m, B_2^m) \leqslant \sqrt{n}$ and $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$
- Generalization for matrices:

$$S_p^m := \{\text{symmetric } \boldsymbol{A} \in \mathbb{R}^{m \times m} : \|\lambda(\boldsymbol{A})\|_p \leqslant 1\}$$

- For diagonal matrices $\|\boldsymbol{A}\|_{S_p^m} = \|\mathrm{diag}(\boldsymbol{A})\|_p$
- $\mathrm{vb}_n(S_2^m, S_2^m) \leqslant \sqrt{n}$

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

$\mathrm{vb}_n(S_\infty^n, S_\infty^n) \lesssim \sqrt{n}$ and in general $\mathrm{vb}_n(S_\infty^m, S_\infty^m) \lesssim \sqrt{n \log(2\max(m,n)/n)}$

Matrix Spencer Conjecture

- $\mathrm{vb}_n(B_2^m, B_2^m) \leqslant \sqrt{n}$ and $\mathrm{vb}_n(B_\infty^m, B_\infty^m) \lesssim \sqrt{n \log(2m/n)}$
- Generalization for matrices:

$$S_p^m := \{\text{symmetric } \boldsymbol{A} \in \mathbb{R}^{m \times m} : \|\lambda(\boldsymbol{A})\|_p \leqslant 1\}$$

- For diagonal matrices $\|\boldsymbol{A}\|_{S_p^m} = \|\mathrm{diag}(\boldsymbol{A})\|_p$
- $\mathrm{vb}_n(S_2^m, S_2^m) \leqslant \sqrt{n}$

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

$\mathrm{vb}_n(S_\infty^n, S_\infty^n) \lesssim \sqrt{n}$ and in general $\mathrm{vb}_n(S_\infty^m, S_\infty^m) \lesssim \sqrt{n \log(2\max(m,n)/n)}$

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_\infty^m, S_\infty^m) \lesssim \sqrt{n \log(2m^2/n)}$ for all $1 \leqslant n \leqslant m^2$.

Basic example

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

Basic example

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

$$\|\boldsymbol{A}_1 + \boldsymbol{A}_2\|_{S^2_\infty} = \left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\|_{S^2_\infty} = 2$$

$$\|\boldsymbol{A}_1 - \boldsymbol{A}_2\|_{S^2_\infty} = \left\| \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right\|_{S^2_\infty} = 1$$

More interesting example

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{A}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \boldsymbol{A}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

More interesting example

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{A}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \boldsymbol{A}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\Big\| \sum_{i=1}^{4} x_i \boldsymbol{A}_i \Big\|_{S_\infty^m} \geqslant \frac{1}{\sqrt{2}} \Big\| \sum_{i=1}^{4} x_i \boldsymbol{A}_i \Big\|_{S_2^m} = \frac{1}{\sqrt{2}} \cdot \sqrt{4 \cdot 2} = 2.$$

More interesting example

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \boldsymbol{A}_2 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \boldsymbol{A}_3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_4 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

$$\Big\| \sum_{i=1}^{4} x_i \boldsymbol{A}_i \Big\|_{S_\infty^m} \geqslant \frac{1}{\sqrt{2}} \Big\| \sum_{i=1}^{4} x_i \boldsymbol{A}_i \Big\|_{S_2^m} = \frac{1}{\sqrt{2}} \cdot \sqrt{4 \cdot 2} = 2.$$

- Generalizes for $n = m^2$: $\sqrt{n}$ lower bound

Partial Colorings

- Given $A, B \subseteq \mathbb{R}^d$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$, define the *discrepancy body*

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i \boldsymbol{v}_i \in B \right\}$$

Partial Colorings

- Given $A, B \subseteq \mathbb{R}^d$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$, define the *discrepancy body*

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i \boldsymbol{v}_i \in B \right\}$$

- Upper bound $\mathrm{vb}_n(A, B)$ by iteratively constructing *partial colorings*

$$\boldsymbol{x} \in (C \cdot K) \cap [-1, 1]^n : |\{i : |x_i| = 1\}| \gtrsim n$$

Partial Colorings

- Given $A, B \subseteq \mathbb{R}^d$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$, define the *discrepancy body*

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i \boldsymbol{v}_i \in B \right\}$$

- Upper bound $\mathrm{vb}_n(A, B)$ by iteratively constructing *partial colorings*

$$\boldsymbol{x} \in (C \cdot K) \cap [-1, 1]^n : |\{i : |x_i| = 1\}| \gtrsim n$$

- Step 1: Can find partial colorings as long as

$$\gamma_n(C \cdot K) := \Pr_{\boldsymbol{g} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)}[\boldsymbol{g} \in C \cdot K] \geqslant 2^{-O(n)}$$

Partial Colorings

- Given $A, B \subseteq \mathbb{R}^d$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$, define the *discrepancy body*

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i \boldsymbol{v}_i \in B \right\}$$

- Upper bound $\mathrm{vb}_n(A, B)$ by iteratively constructing *partial colorings*

$$\boldsymbol{x} \in (C \cdot K) \cap [-1, 1]^n : |\{i : |x_i| = 1\}| \gtrsim n$$

- Step 1: Can find partial colorings as long as

$$\gamma_n(C \cdot K) := \Pr_{\boldsymbol{g} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)}[\boldsymbol{g} \in C \cdot K] \geqslant 2^{-O(n)}$$

- Step 2: Suffices to cover *polar body* $(C \cdot K)^{\circ}$ with $2^{O(n)}$ balls or cubes

Partial Colorings

- Given $A, B \subseteq \mathbb{R}^d$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in A$, define the *discrepancy body*

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i \boldsymbol{v}_i \in B \right\}$$

- Upper bound $\mathrm{vb}_n(A, B)$ by iteratively constructing *partial colorings*

$$\boldsymbol{x} \in (C \cdot K) \cap [-1, 1]^n : |\{i : |x_i| = 1\}| \gtrsim n$$

- Step 1: Can find partial colorings as long as

$$\gamma_n(C \cdot K) := \Pr_{\boldsymbol{g} \sim N(\boldsymbol{0}, \boldsymbol{I}_n)}[\boldsymbol{g} \in C \cdot K] \geqslant 2^{-O(n)}$$

- Step 2: Suffices to cover *polar body* $(C \cdot K)^\circ$ with $2^{O(n)}$ balls or cubes
- Step 3: Construct such covering via optimization

Spencer's setting

- Given $\boldsymbol{V} \in \mathbb{R}^{m \times n}$ with columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in \mathsf{B}_\infty^m$, let

$$\mathsf{K} := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{v}_i \right\|_\infty \leqslant 1 \right\}$$

Spencer's setting

- Given $\boldsymbol{V} \in \mathbb{R}^{m \times n}$ with columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in B_\infty^m$, let

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{v}_i \right\|_\infty \leqslant 1 \right\}$$

- Here we have

$$\begin{aligned} K^\circ &= \left\{ (\langle \boldsymbol{v}_1, \boldsymbol{u} \rangle, \langle \boldsymbol{v}_2, \boldsymbol{u} \rangle, \ldots, \langle \boldsymbol{v}_n, \boldsymbol{u} \rangle) : \boldsymbol{u} \in B_1^m \right\} \\ &= \mathrm{conv}\left(\left\{ \pm (\langle \boldsymbol{v}_1, \boldsymbol{e}_j \rangle, \langle \boldsymbol{v}_2, \boldsymbol{e}_j \rangle, \ldots, \langle \boldsymbol{v}_n, \boldsymbol{e}_j \rangle) : j \in [m] \right\} \right) \end{aligned}$$

Spencer's setting

- Given $\boldsymbol{V} \in \mathbb{R}^{m \times n}$ with columns $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in B_\infty^m$, let

$$K := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{v}_i \right\|_\infty \leqslant 1 \right\}$$

- Here we have

$$\begin{aligned} K^\circ &= \left\{ (\langle \boldsymbol{v}_1, \boldsymbol{u} \rangle, \langle \boldsymbol{v}_2, \boldsymbol{u} \rangle, \ldots, \langle \boldsymbol{v}_n, \boldsymbol{u} \rangle) : \boldsymbol{u} \in B_1^m \right\} \\ &= \mathrm{conv}\left(\left\{ \pm (\langle \boldsymbol{v}_1, \boldsymbol{e}_j \rangle, \langle \boldsymbol{v}_2, \boldsymbol{e}_j \rangle, \ldots, \langle \boldsymbol{v}_n, \boldsymbol{e}_j \rangle) : j \in [m] \right\} \right) \end{aligned}$$

Approximate Carathéodory Theorem

Given $\boldsymbol{x} \in \mathrm{conv}(\{\boldsymbol{a}_j\}_{j \in S}) \subseteq B_2^m$ and $k \in \mathbb{N}$, there exist $i_1, \ldots, i_k \in [n]$ with

$$\left\| \boldsymbol{x} - \frac{1}{k} \sum_{j=1}^{k} \boldsymbol{a}_{i_j} \right\|_2 \leqslant \frac{1}{\sqrt{k}}$$

Spencer's setting

Approximate Carathéodory Theorem

Given $\boldsymbol{x} \in \operatorname{conv}(\{\boldsymbol{a}_j\}_{j \in S}) \subseteq B_2^m$ and $k \in \mathbb{N}$, there exist $i_1, \ldots, i_k \in [n]$ with

$$\left\| \boldsymbol{x} - \frac{1}{k} \sum_{j=1}^{k} \boldsymbol{a}_{i_j} \right\|_2 \leqslant \frac{1}{\sqrt{k}}$$

- For $k := \lfloor n / \log(2m/n) \rfloor$ we get $\approx \binom{2m}{n/\log(2m/n)} \leqslant 2^{O(n)}$ centers

Spencer's setting

Approximate Carathéodory Theorem

Given $\boldsymbol{x} \in \operatorname{conv}(\{\boldsymbol{a}_j\}_{j \in S}) \subseteq B_2^m$ and $k \in \mathbb{N}$, there exist $i_1, \ldots, i_k \in [n]$ with

$$\left\| \boldsymbol{x} - \frac{1}{k} \sum_{j=1}^{k} \boldsymbol{a}_{i_j} \right\|_2 \leqslant \frac{1}{\sqrt{k}}$$

- For $k := \lfloor n / \log(2m/n) \rfloor$ we get $\approx \binom{2m}{n/\log(2m/n)} \leqslant 2^{O(n)}$ centers
- In Spencer's setting the rows have ℓ_2 norm $\leqslant \sqrt{n}$, implying

$$N\left(\frac{1}{\sqrt{n}} K^{\circ}, \frac{\sqrt{\log(2m/n)}}{\sqrt{n}} B_2^n \right) \leqslant 2^{O(n)}$$

Spencer's setting

Approximate Carathéodory Theorem

Given $\boldsymbol{x} \in \mathrm{conv}(\{\boldsymbol{a}_j\}_{j\in S}) \subseteq B_2^m$ and $k \in \mathbb{N}$, there exist $i_1, \ldots, i_k \in [n]$ with

$$\Big\|\boldsymbol{x} - \frac{1}{k}\sum_{j=1}^{k} \boldsymbol{a}_{i_j}\Big\|_2 \leqslant \frac{1}{\sqrt{k}}$$

- For $k := \lfloor n/\log(2m/n)\rfloor$ we get $\approx \binom{2m}{n/\log(2m/n)} \leqslant 2^{O(n)}$ centers
- In Spencer's setting the rows have ℓ_2 norm $\leqslant \sqrt{n}$, implying

$$N\Big(\frac{1}{\sqrt{n}}K^\circ, \frac{\sqrt{\log(2m/n)}}{\sqrt{n}}B_2^n\Big) \leqslant 2^{O(n)}$$

$$\implies \gamma_n\Big(\sqrt{n\log(2m/n)}K\Big) \geqslant 2^{-O(n)}.$$

Example: Spencer's setting

$$\boldsymbol{v}_1 = (1, 0) \text{ and } \boldsymbol{v}_2 = (0, 1)$$

Example: Spencer's setting

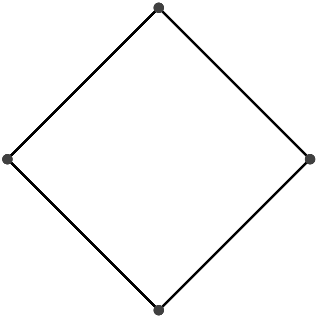
$$\boldsymbol{v}_1 = (1, 0) \text{ and } \boldsymbol{v}_2 = (0, 1)$$

- Here $K^\circ = \{(u_1, u_2) : \boldsymbol{u} \in B_1^2\} = B_1^2 = \mathrm{conv}\{\pm\boldsymbol{v}_1, \pm\boldsymbol{v}_2\}$

Example: Spencer's setting

$$\boldsymbol{v}_1 = (1, 0) \text{ and } \boldsymbol{v}_2 = (0, 1)$$

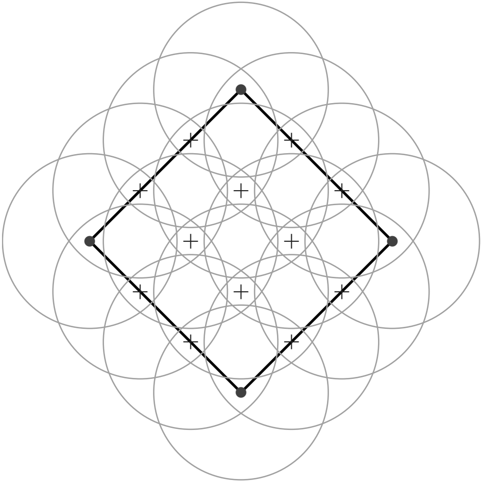
- Here $K^\circ = \{(u_1, u_2) : \boldsymbol{u} \in B_1^2\} = B_1^2 = \mathrm{conv}\{\pm\boldsymbol{v}_1, \pm\boldsymbol{v}_2\}$



Example: Spencer's setting

$$\boldsymbol{v}_1 = (1, 0) \text{ and } \boldsymbol{v}_2 = (0, 1)$$

- Here $K^\circ = \{(u_1, u_2) : \boldsymbol{u} \in B_1^2\} = B_1^2 = \text{conv}\{\pm\boldsymbol{v}_1, \pm\boldsymbol{v}_2\}$



Obstacle for Matrix Spencer

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

Obstacle for Matrix Spencer

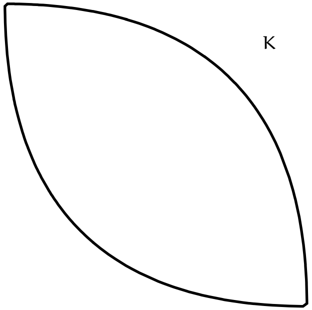
$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

$$\mathsf{K} := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{S_\infty^m} \leqslant 1 \right\}$$

Obstacle for Matrix Spencer

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

$$\mathsf{K} := \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{S_\infty^m} \leqslant 1 \right\}$$



Obstacle for Matrix Spencer

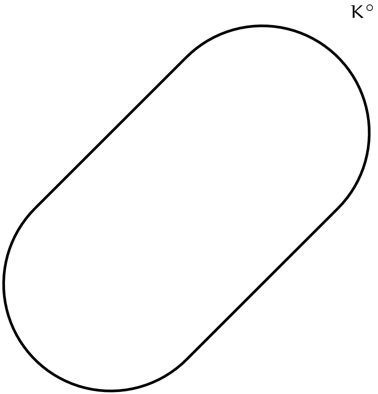
$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

$$\mathsf{K}^\circ = \left\{ (\langle \boldsymbol{A}_1, \boldsymbol{U} \rangle, \langle \boldsymbol{A}_2, \boldsymbol{U} \rangle, \ldots, \langle \boldsymbol{A}_n, \boldsymbol{U} \rangle) : \boldsymbol{U} \in S_1^m \right\}$$

Obstacle for Matrix Spencer

$$\boldsymbol{A}_1 = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 0 \end{pmatrix} \text{ and } \boldsymbol{A}_2 = \begin{pmatrix} 0 & 1/2 \\ 1/2 & 1 \end{pmatrix}$$

$$\mathsf{K}^\circ = \left\{ (\langle \boldsymbol{A}_1, \boldsymbol{U} \rangle, \langle \boldsymbol{A}_2, \boldsymbol{U} \rangle, \ldots, \langle \boldsymbol{A}_n, \boldsymbol{U} \rangle) : \boldsymbol{U} \in S_1^m \right\}$$



Mirror descent framework

- Fix a domain $\mathcal{D} \subseteq \mathbb{R}^m$ and a subset $\mathcal{X} \subseteq \overline{\mathcal{D}}$, start $\boldsymbol{x}_0 \in \mathcal{X} \cap \mathcal{D}$

Mirror descent framework

- Fix a domain $\mathcal{D} \subseteq \mathbb{R}^m$ and a subset $\mathcal{X} \subseteq \overline{\mathcal{D}}$, start $\boldsymbol{x}_0 \in \mathcal{X} \cap \mathcal{D}$
- L-Lipschitz convex $f : \mathcal{X} \to \mathbb{R}$ and a ρ-strongly convex $\Phi : \mathcal{D} \to \mathbb{R}$

Mirror descent framework

- Fix a domain $\mathcal{D} \subseteq \mathbb{R}^m$ and a subset $\mathcal{X} \subseteq \overline{\mathcal{D}}$, start $\boldsymbol{x}_0 \in \mathcal{X} \cap \mathcal{D}$
- L-Lipschitz convex $f : \mathcal{X} \to \mathbb{R}$ and a ρ-strongly convex $\Phi : \mathcal{D} \to \mathbb{R}$
- Iterate

$$\nabla\Phi(\boldsymbol{y}_{t+1}) := \nabla\Phi(\boldsymbol{x}_t) - \eta \boldsymbol{g}_t$$

$$\boldsymbol{x}_{t+1} := \operatorname{argmin}_{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{D}} D_\Phi(\boldsymbol{x}, \boldsymbol{y}_{t+1}),$$

where $\boldsymbol{g}_t \in \partial f(\boldsymbol{x}_t)$ and $D_\Phi(\boldsymbol{x}, \boldsymbol{y}) := \Phi(\boldsymbol{x}) - \Phi(\boldsymbol{y}) - \langle \nabla\Phi(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle$.

Mirror descent framework

- Fix a domain $\mathcal{D} \subseteq \mathbb{R}^m$ and a subset $\mathcal{X} \subseteq \overline{\mathcal{D}}$, start $\boldsymbol{x}_0 \in \mathcal{X} \cap \mathcal{D}$
- L-Lipschitz convex $f : \mathcal{X} \to \mathbb{R}$ and a ρ-strongly convex $\Phi : \mathcal{D} \to \mathbb{R}$
- Iterate

$$\nabla\Phi(\boldsymbol{y}_{t+1}) := \nabla\Phi(\boldsymbol{x}_t) - \eta \boldsymbol{g}_t$$

$$\boldsymbol{x}_{t+1} := \operatorname{argmin}_{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{D}} D_\Phi(\boldsymbol{x}, \boldsymbol{y}_{t+1}),$$

where $\boldsymbol{g}_t \in \partial f(\boldsymbol{x}_t)$ and $D_\Phi(\boldsymbol{x}, \boldsymbol{y}) := \Phi(\boldsymbol{x}) - \Phi(\boldsymbol{y}) - \langle \nabla\Phi(\boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle$.

Mirror descent (Bubéck ’15)

After T iterations with $\eta := \frac{1}{L}\sqrt{\frac{2\rho D_\Phi(\boldsymbol{x}^*, \boldsymbol{x}_0)}{T}}$ we get a sequence $\boldsymbol{x}_t$ with

$$\min_{t<T} f(\boldsymbol{x}_t) - f(\boldsymbol{x}^*) \leqslant L \cdot \sqrt{\frac{2D_\Phi(\boldsymbol{x}^*, \boldsymbol{x}_0)}{\rho T}},$$

where $\boldsymbol{x}^* := \operatorname{argmin}_{\boldsymbol{x} \in \mathcal{X} \cap \mathcal{D}} f(\boldsymbol{x})$.

Mirror descent: Matrix Spencer setting

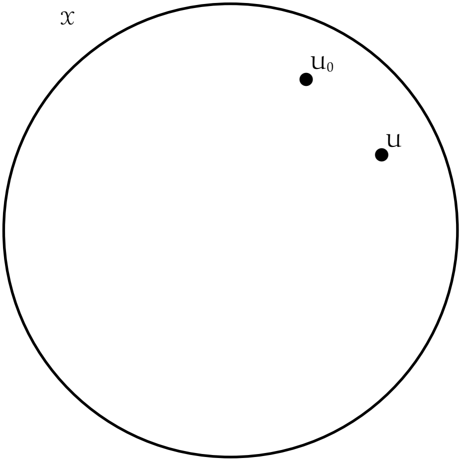
- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$

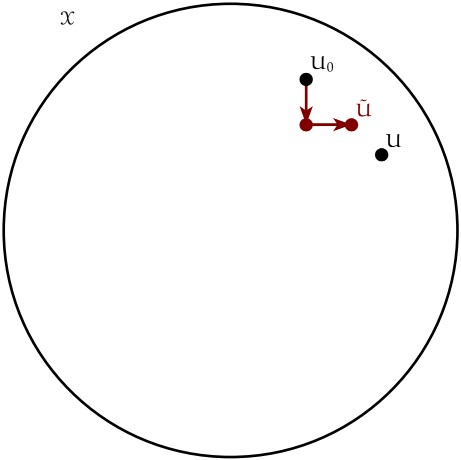
Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\|\cdot\|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$



Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \operatorname{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \operatorname{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\|\cdot\|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$



Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- Recall $K^\circ = \{\underbrace{(\langle \boldsymbol{A}_1, \boldsymbol{U} \rangle, \langle \boldsymbol{A}_2, \boldsymbol{U} \rangle, \ldots, \langle \boldsymbol{A}_n, \boldsymbol{U} \rangle)}_{=: \boldsymbol{A}(\boldsymbol{U})} : \boldsymbol{U} \in S_1^m\}$ and note

$$f_{\boldsymbol{U}}(\tilde{\boldsymbol{U}}) \leqslant C \implies \boldsymbol{A}(\boldsymbol{U}) \in C \cdot B_\infty^n + \boldsymbol{A}(\tilde{\boldsymbol{U}})$$

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \operatorname{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \operatorname{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- Recall $K^\circ = \{\underbrace{(\langle \boldsymbol{A}_1, \boldsymbol{U} \rangle, \langle \boldsymbol{A}_2, \boldsymbol{U} \rangle, \ldots, \langle \boldsymbol{A}_n, \boldsymbol{U} \rangle)}_{=:\boldsymbol{A}(\boldsymbol{U})} : \boldsymbol{U} \in S_1^m\}$ and note

$$f_{\boldsymbol{U}}(\tilde{\boldsymbol{U}}) \leqslant C \implies \boldsymbol{A}(\boldsymbol{U}) \in C \cdot B_\infty^n + \boldsymbol{A}(\tilde{\boldsymbol{U}})$$

- How to bound the number of centers $\boldsymbol{A}(\tilde{\boldsymbol{U}})$ starting from some $\boldsymbol{U}_0$?

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \operatorname{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \operatorname{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- $D_\Phi(\boldsymbol{X}, \boldsymbol{Y}) = S(\boldsymbol{X} \| \boldsymbol{Y}) = \mathrm{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}))$

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m\times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i\in[n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U}\rangle|$ is $\|\cdot\|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- $D_\Phi(\boldsymbol{X}, \boldsymbol{Y}) = S(\boldsymbol{X}\|\boldsymbol{Y}) = \mathrm{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}))$
- Closed formula:

$$\boldsymbol{U}_{t+1} = \frac{\exp\left(\log \boldsymbol{U}_0 - \eta \sum_{i=0}^t \boldsymbol{G}_i\right)}{\mathrm{tr}\Big(\exp\left(\log \boldsymbol{U}_0 - \eta \sum_{i=0}^t \boldsymbol{G}_i\right)\Big)},$$

where $\boldsymbol{G}_t \in \partial f(\boldsymbol{X}_t) \subseteq \{\pm\boldsymbol{A}_1, \ldots, \pm\boldsymbol{A}_n\}$.

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- $D_\Phi(\boldsymbol{X}, \boldsymbol{Y}) = S(\boldsymbol{X} \| \boldsymbol{Y}) = \mathrm{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}))$
- Closed formula:

$$\boldsymbol{U}_{t+1} = \frac{\exp\left(\log \boldsymbol{U}_0 - \eta \sum_{i=0}^t \boldsymbol{G}_i\right)}{\mathrm{tr}\Big(\exp\left(\log \boldsymbol{U}_0 - \eta \sum_{i=0}^t \boldsymbol{G}_i\right)\Big)},$$

  where $\boldsymbol{G}_t \in \partial f(\boldsymbol{X}_t) \subseteq \{\pm \boldsymbol{A}_1, \ldots, \pm \boldsymbol{A}_n\}$.

- Key observation: $\boldsymbol{U}_t$ does not depend on the *order* of the gradients!

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \operatorname{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \operatorname{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- $D_\Phi(\boldsymbol{X}, \boldsymbol{Y}) = S(\boldsymbol{X} \| \boldsymbol{Y}) = \operatorname{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}))$
- Closed formula:

$$\boldsymbol{U}_{t+1} = \frac{\exp\left(\log \boldsymbol{U}_0 - \eta \sum_{i=0}^{t} \boldsymbol{G}_i\right)}{\operatorname{tr}\left(\exp\left(\log \boldsymbol{U}_0 - \eta \sum_{i=0}^{t} \boldsymbol{G}_i\right)\right)},$$

  where $\boldsymbol{G}_t \in \partial f(\boldsymbol{X}_t) \subseteq \{\pm \boldsymbol{A}_1, \ldots, \pm \boldsymbol{A}_n\}$.
- Key observation: $\boldsymbol{U}_t$ does not depend on the *order* of the gradients!
- After $T := n$ iterations, we have at most

$$\sum_{t=0}^{n} \binom{t + 2n - 1}{2n - 1} \leqslant (n+1) \cdot \binom{3n}{n} \leqslant 2^{O(n)} \text{ centers } \boldsymbol{U}_t.$$

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \operatorname{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \operatorname{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- Each starting point $\boldsymbol{U}_0$ generates $2^{O(n)}$ centers of cubes

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \mathrm{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \mathrm{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- Each starting point $\boldsymbol{U}_0$ generates $2^{O(n)}$ centers of cubes
- Each cube is scaled by

$$\sqrt{\frac{S(\boldsymbol{U} \| \boldsymbol{U}_0)}{n}} \leqslant \sqrt{\frac{\log n}{n}},$$

giving a $\sqrt{n \log n}$ partial coloring bound.

Mirror descent: Matrix Spencer setting

- $\mathcal{X} := \{\boldsymbol{X} \in \mathbb{R}^{m \times m} : \boldsymbol{X} \succeq 0, \operatorname{tr}(\boldsymbol{X}) = 1\}$ and $\Phi(\boldsymbol{X}) := \operatorname{tr}(\boldsymbol{X} \log \boldsymbol{X})$
- $f_{\boldsymbol{U}}(\boldsymbol{X}) := \max_{i \in [n]} |\langle \boldsymbol{A}_i, \boldsymbol{X} - \boldsymbol{U} \rangle|$ is $\| \cdot \|_{S_1^m}$-Lipschitz for $\boldsymbol{A}_i \in S_\infty^m$
- Each starting point $\boldsymbol{U}_0$ generates $2^{O(n)}$ centers of cubes
- Each cube is scaled by

$$\sqrt{\frac{S(\boldsymbol{U} \| \boldsymbol{U}_0)}{n}} \leqslant \sqrt{\frac{\log n}{n}},$$

  giving a $\sqrt{n \log n}$ partial coloring bound.
- Pick $2^{O(n)}$ starting points $\boldsymbol{U}_0$, total of $2^{O(n)} \cdot 2^{O(n)} \leqslant 2^{O(n)}$ centers!

Quantum relative entropy net

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

Quantum relative entropy net

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

Covering S_1^m with S_∞^m [HPV17]

$N\left(S_1^m, \frac{m}{n} S_\infty^m\right) \leqslant 2^{O(n)}$.

Quantum relative entropy net

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

Covering S_1^m with S_∞^m [HPV17]

$N\left(S_1^m, \frac{m}{n} S_\infty^m\right) \leqslant 2^{O(n)}$.

- Combining the two results gives a net with quantum relative entropy

$$\log(2m \cdot m/n) = \log(2m^2/n).$$

Quantum relative entropy net proof

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

Quantum relative entropy net proof

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

$$S(\boldsymbol{X} \| \boldsymbol{Y}') = \operatorname{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}'))$$

Quantum relative entropy net proof

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

$$\begin{aligned} S(\boldsymbol{X} \| \boldsymbol{Y}') &= \operatorname{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}')) \\ &\leqslant \operatorname{tr}(\boldsymbol{X} \cdot (\log(\boldsymbol{Y} + \varepsilon \boldsymbol{I}_m) - \log \boldsymbol{Y}')) \end{aligned}$$

Quantum relative entropy net proof

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

$$\begin{aligned} S(\boldsymbol{X} \| \boldsymbol{Y}') &= \operatorname{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}')) \\ &\leqslant \operatorname{tr}(\boldsymbol{X} \cdot (\log(\boldsymbol{Y} + \varepsilon \boldsymbol{I}_m) - \log \boldsymbol{Y}')) \\ &\leqslant \underbrace{\|\boldsymbol{X}\|_{S_1^m}}_{=1} \cdot \|\log(\boldsymbol{Y} + \varepsilon \boldsymbol{I}_m) - \log \boldsymbol{Y}'\|_{S_\infty^m} \end{aligned}$$

Quantum relative entropy net proof

Key lemma

Let $\boldsymbol{X}, \boldsymbol{Y} \in \mathcal{X}$ satisfy $\|\boldsymbol{X} - \boldsymbol{Y}\|_{S_\infty^m} \leqslant \varepsilon$ for some $\varepsilon \geqslant 1/m$. Then

$$S(\boldsymbol{X} \| \boldsymbol{Y}') \leqslant \log(2m\varepsilon),$$

where $\boldsymbol{Y}' := \frac{1}{2} \cdot \boldsymbol{Y} + \frac{1}{2} \cdot \frac{\boldsymbol{I}_m}{m}$.

$$\begin{aligned}
S(\boldsymbol{X} \| \boldsymbol{Y}') &= \operatorname{tr}(\boldsymbol{X}(\log \boldsymbol{X} - \log \boldsymbol{Y}')) \\
&\leqslant \operatorname{tr}(\boldsymbol{X} \cdot (\log(\boldsymbol{Y} + \varepsilon \boldsymbol{I}_m) - \log \boldsymbol{Y}')) \\
&\leqslant \underbrace{\|\boldsymbol{X}\|_{S_1^m}}_{=1} \cdot \|\log(\boldsymbol{Y} + \varepsilon \boldsymbol{I}_m) - \log \boldsymbol{Y}'\|_{S_\infty^m} \\
&\leqslant \log\left(2 \cdot \left\| \frac{\boldsymbol{Y} + \varepsilon \boldsymbol{I}_m}{\boldsymbol{Y} + \frac{\boldsymbol{I}_m}{m}} \right\|_{S_\infty^m}\right) \leqslant \log(2m\varepsilon).
\end{aligned}$$

Further applications

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_\infty^{m,h}, S_\infty^{m,h}) \lesssim \sqrt{n \log(hm/n)}$ for h-block diagonal matrices.

Further applications

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_\infty^{m,h}, S_\infty^{m,h}) \lesssim \sqrt{n \log(hm/n)}$ for h-block diagonal matrices.

- Sketch: interpolate between two nets

$$N\left(S_1^m, \frac{m}{n} S_\infty^m\right) \leqslant 2^{O(n)} \text{ and } N\left(B_1^m, \frac{\log(2m/n)}{n} B_\infty^m\right) \leqslant 2^{O(n)}.$$

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_p^m, S_q^m) \lesssim \sqrt{n \cdot \min(p, \log(2m^2/n))} \cdot \min(1, m/n)^{1/p - 1/q}$.

Further applications

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_\infty^{m,h}, S_\infty^{m,h}) \lesssim \sqrt{n \log(hm/n)}$ for h-block diagonal matrices.

- Sketch: interpolate between two nets

$$N\left(S_1^m, \frac{m}{n} S_\infty^m\right) \leqslant 2^{O(n)} \text{ and } N\left(B_1^m, \frac{\log(2m/n)}{n} B_\infty^m\right) \leqslant 2^{O(n)}.$$

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_p^m, S_q^m) \lesssim \sqrt{n \cdot \min(p, \log(2m^2/n))} \cdot \min(1, m/n)^{1/p-1/q}.$

Further applications

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_\infty^{m,h}, S_\infty^{m,h}) \lesssim \sqrt{n \log(hm/n)}$ for h-block diagonal matrices.

- Sketch: interpolate between two nets

$$N\left(S_1^m, \frac{m}{n} S_\infty^m\right) \leqslant 2^{O(n)} \text{ and } N\left(B_1^m, \frac{\log(2m/n)}{n} B_\infty^m\right) \leqslant 2^{O(n)}.$$

Theorem (Dadush, Jiang, R. '22)

$\mathrm{vb}_n(S_p^m, S_q^m) \lesssim \sqrt{n \cdot \min(p, \log(2m^2/n))} \cdot \min(1, m/n)^{1/p-1/q}$.

- Sketch: use a different mirror map $\Phi(\boldsymbol{X}) := \frac{1}{2(p-1)} \|\boldsymbol{X}\|_{S_p^m}^2$

Open problems

- Can we show a $\Omega(\log(2m^2/n))$ lower bound for entropy nets?

Open problems

- Can we show a $\Omega(\log({}^{2m^2}/_{n}))$ lower bound for entropy nets?
- Given $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n \in \mathbb{R}^{n \times n}$ can we show measure bound for

$$\mathsf{K} = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{S_\infty^n} \leqslant \left\| \left(\sum_{i=1}^{n} \boldsymbol{A}_i^2 \right)^{1/2} \right\|_{S_\infty^n} \right\}$$

Open problems

- Can we show a $\Omega(\log(2m^2/n))$ lower bound for entropy nets?
- Given $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n \in \mathbb{R}^{n\times n}$ can we show measure bound for

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{S_\infty^n} \leqslant \left\| \left(\sum_{i=1}^{n} \boldsymbol{A}_i^2 \right)^{1/2} \right\|_{S_\infty^n} \right\}$$

- Tight coverings for $\mathrm{vb}_n(B_p^m, B_q^m)$?

Open problems

- Can we show a $\Omega(\log(2m^2/n))$ lower bound for entropy nets?
- Given $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_n \in \mathbb{R}^{n \times n}$ can we show measure bound for

$$K = \left\{ \boldsymbol{x} \in \mathbb{R}^n : \left\| \sum_{i=1}^{n} x_i \boldsymbol{A}_i \right\|_{S_\infty^n} \leqslant \left\| \left(\sum_{i=1}^{n} \boldsymbol{A}_i^2 \right)^{1/2} \right\|_{S_\infty^n} \right\}$$

- Tight coverings for $\mathrm{vb}_n(B_p^m, B_q^m)$?

Thanks for your attention!