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

Daniel Dadush (CWI), Haotian Jiang (UW) and Victor Reis (UW)

Introduction to vector balancing

- Given \( A, B \subseteq \mathbb{R}^d \), \( v_B(A, B) := \text{smallest } C > 0 \) so that for any \( v_1, \ldots, v_n \in A \) there are signs \( x \in \{ \pm 1 \}^n \) with
  \[
  \sum_{i=1}^n x_i v_i \in C : B.
  \]
- Applications to scheduling [BRSZ22] and RCTs [HSSZ19]
- Random signs give \( v_B(B^*_{n^2}, B^*_n) \leq \sqrt{n} \log(2n) \) by Chernoff + union bound

Matrix Spencer Conjecture

- Generalization for matrices: \( S^p_n := \{ \text{symmetric } A \in \mathbb{R}^{n \times n} : \| A \|_p \leq 1 \} \)
- For diagonal matrices \( \| A \|_\infty = \| \text{diag}(A) \|_\infty \)

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

\( v_B(S^p_n, S^p_2) \leq \sqrt{n} \) in general, \( v_B(S^p_n, S^p_2) \leq \sqrt{n} \log^{O(n^{1/3}/\sqrt{n})} \)

Theorem (Dadush, Jiang, Reis'22; Hopkins, Raghavendra, Shetty'22)

\( v_B(S^p_n, S^p_2) \leq \sqrt{n} \log^{O(n^{1/3}/\sqrt{n})} \) for all \( 1 \leq n \leq m^2 \).

Partial coloring bound

- Given \( A, B \subseteq \mathbb{R}^d \) and \( v_1, \ldots, v_n \in A \), define the discrepancy body
  \( K := \{ x \in \mathbb{R}^d : \sum_{i=1}^n x_i v_i \in B \} \)
- Upper bound \( v_B(A, B) \) by iteratively constructing partial colorings
  \( x \in (C \setminus K) \cap [-1, 1]^n : \{ |x_i| = 1 \} \geq n \)
- Can find partial colorings if \( \gamma_C(C \setminus K) = \text{Pr}_{x \sim \mathcal{N}(0, 1)}(y \in C \setminus K)^2 > 2^{-O(n)} \)
- Sufficient to cover polar body \( (C \setminus K)^* \) with \( 2^{O(n)} \) balls \( B^*_{n^2} \) or cubes \( B^*_{n^2} \)
- Can construct such covering via optimization

Mirror descent setting

- \( X = \{ x \in \mathbb{R}^{n \times n} : X \geq 0, \text{tr}(X) = 1 \} \) and \( \Phi(X) = \text{tr}(X \log X) \)
- \( f_\rho(x) := \max_{y \in \mathbb{R}^n} \{ (x, y) \} \) \( ||y||_{\Phi} \)-Lipschitz for \( A_i \in S^p_n \)
- \( D_\Phi(X, Y) = S(X||Y) = \text{tr}(X \log X - Y \log Y) \)

Mirror descent: Matrix Spencer setting

- Closed formula:
  \( U_{i+1} = \exp \left( \log(U_i) - \eta \sum_{t=0}^i G_t \right) \tr(\exp(\log(U_i) - \eta \sum_{t=0}^i G_t)) \)
  with \( G_t \in \partial(U(t) \subseteq \{ \pm A_1, \ldots, \pm A_n \}) \)
- Key observation: \( U_i \) does not depend on the order of the gradients!
- After \( T := n \) iterations, we have at most
  \[ \sum_{t=0}^T \left( n + 1 \right) \leq (n+1) \cdot \left( \frac{3n}{n} \right) \leq 2^{O(n)} \text{ centers } U_i, \]
- Each cube is scaled by
  \[ \frac{\sqrt{S(U_i) n}}{n} \leq \frac{\log n}{n} \]
  giving a \( \sqrt{n \log n} \) partial coloring bound.
- Pick \( 2^{3^{O(n)}} \) starting points \( U_0 \), total of \( 2^{3^{O(n)}} \) \( 2^{O(n)} \) \( S^p_n \) centers!

Mirror descent: Quantum relative entropy net

Key observation (Dadush, Jiang, Reis’22)

Let \( X, Y \in X \) satisfy \( \|X - Y\|_{\log} \leq \epsilon \) for some \( \epsilon > 1/m \).
That \( S(X || Y) \leq \log(2me) \), where \( Y^* := Y + \frac{1}{2} \frac{X - Y}{m} \)

Covering \( S^p_n \) with \( S^p_{n^2} \) [HPY17]

- Combining the two results gives a net with quantum relative entropy
  \( \log(2m \cdot m/n) = \log(2m/n) \)

Further applications

Theorem (Dadush, Jiang, Reis’22)

\( v_B(S^{m n^2} B^*_{n^2}, S^{m^2} B^*_{n^2}) \leq \sqrt{n} \log^{O(n^{1/3}/\sqrt{n})}/n \) for h-block diagonal matrices.

Sketch: interpolate between two nets

\[
\text{N}_{\left( S^p_n, \frac{m n^2}{n} S^p_n \right)} \leq 2^{O(n)} \quad \text{and} \quad \text{N}_{\left( B_{n^2}, \frac{\log(2m/n)}{n} B^*_{n^2} \right)} \leq 2^{O(n)}.
\]

Theorem (Dadush, Jiang, Reis’22)

\( v_B(S^n S^n B^*_{n^2}, S^n B^*_{m^2}) \leq \sqrt{n} \min\{\log(2m/n), 1 \} \min(1, n/m) \log^{O(n^{1/3}/\sqrt{n})}/n \}

Sketch: use a different mirror map \( \Phi(X) := \frac{1}{m} \sqrt{X} 1_{\leq m/2}X 1_{> m/2} \)