Algorithmic Aspects of the Brascamp Lieb Inequality

Jacob Denson (Based on Gurvits, 2004)

University of Wisconsin Madison

September 28, 2021

Lieb's Theorem

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

Lieb's Theorem

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Lieb, 1990) To calculate BL(B,p), plug in Gaussians.

Lieb's Theorem

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \le \mathrm{BL}(B,p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Lieb, 1990) To calculate BL(B,p), plug in Gaussians.
- If $f_i(x) = e^{-\pi (A_i x)\cdot x}$ for $A_i \succ 0$, then

Lieb's Theorem

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B,p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Lieb, 1990) To calculate BL(B,p), plug in Gaussians.
- If $f_i(x) = e^{-\pi (A_i x)\cdot x}$ for $A_i \succ 0$, then

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx = \left(\sum p_i B_i^* A_i B_i\right)^{-1/2}$$

Lieb's Theorem

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B,p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Lieb, 1990) To calculate BL(B,p), plug in Gaussians.
- If $f_i(x) = e^{-\pi (A_i x)\cdot x}$ for $A_i \succ 0$, then

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx = (\sum p_i B_i^* A_i B_i)^{-1/2}$$

$$\prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i} = (\prod_{i=1}^{m} \det(A_i)^{p_i})^{-1/2}.$$

Thus

$$(\sum p_i B_i^* A_i B_i)^{-1/2} \leq \mathrm{BL}(B,p) \cdot (\prod_{i=1}^{m} \det(A_i)^{p_i})^{-1/2}$$

Lieb's Theore

- Rearranging gives

$$\mathrm{BL}(B, p) = \left(\sup_{A_1, \ldots, A_m \succ 0} \frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i B_i^* A_i B_i)} \right)^{1/2}.$$

Lieb's Theore

- Rearranging gives

$$\mathrm{BL}(B, p) = \left(\sup_{A_1,\ldots,A_m \succ 0} \frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i B_i^* A_i B_i)} \right)^{1/2}.$$

- Non convex objective, so tricky to optimize (NP hard).

Lieb's Theore

- Rearranging gives

$$\mathrm{BL}(B,p) = \left(\sup_{A_1,\ldots,A_m \succ 0} \frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i B_i^* A_i B_i)} \right)^{1/2}.$$

- Non convex objective, so tricky to optimize (NP hard).
- Checking Finiteness is also non-obvious.

Lieb's Theore

- Rearranging gives

$$\mathrm{BL}(B,p) = \left(\sup_{A_1,\ldots,A_m \succ 0} \frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i B_i^* A_i B_i)} \right)^{1/2}.$$

- Non convex objective, so tricky to optimize (NP hard).
- Checking Finiteness is also non-obvious.
- BCCT gives a family of conditions

$$\dim(V) \leq \sum p_i \dim(B_i V).$$

Since $0 \leq \dim(V) \leq n$ and $0 \leq \dim(B_i V) \leq n_i$, only finitely many linear conditions on the p_i.

Lieb's Theore

- Rearranging gives

$$\mathrm{BL}(B,p) = \left(\sup_{A_1,\ldots,A_m \succ 0} \frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i B_i^* A_i B_i)} \right)^{1/2}.$$

- Non convex objective, so tricky to optimize (NP hard).
- Checking Finiteness is also non-obvious.
- BCCT gives a family of conditions

$$\dim(V) \leq \sum p_i \dim(B_i V).$$

  Since $0 \leq \dim(V) \leq n$ and $0 \leq \dim(B_i V) \leq n_i$, only finitely many linear conditions on the p_i.
- But there can be exponentially many, so still tricky to compute in practice.

Geometric Brascamp Lieb

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}$$

Geometric Brascamp Lieb

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}$$

- A Brascamp-Lieb inequality is *geometric* if each B_i is a projection ($B_i B_i^* = I$), and $\sum_i p_i B_i^* B_i = I$.

Geometric Brascamp Lieb

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}$$

- A Brascamp-Lieb inequality is *geometric* if each B_i is a projection ($B_i B_i^* = I$), and $\sum_i p_i B_i^* B_i = I$.
- Hölder's inequality and Loomis-Whitney are special cases.

Geometric Brascamp Lieb

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}$$

- A Brascamp-Lieb inequality is *geometric* if each B_i is a projection ($B_i B_i^* = I$), and $\sum_i p_i B_i^* B_i = I$.
- Hölder's inequality and Loomis-Whitney are special cases.
- Plugging in $f_i(x) = e^{-\pi |x|^2}$ gives $\mathrm{BL}(B, p) \geq 1$.

Geometric Brascamp Lieb

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}$$

- A Brascamp-Lieb inequality is *geometric* if each B_i is a projection ($B_i B_i^* = I$), and $\sum_i p_i B_i^* B_i = I$.
- Hölder's inequality and Loomis-Whitney are special cases.
- Plugging in $f_i(x) = e^{-\pi|x|^2}$ gives $\mathrm{BL}(B, p) \geq 1$.
- (Barthe, 1998), generalizing (Ball, 1989), showed that these functions are extremizers, i.e. $\mathrm{BL}(B, p) = 1$.

Operator Rescaling

- What happens to $\mathrm{BL}(B, p)$ if we 'rescale / change coordinates'.

Operator Rescaling

- What happens to $\mathrm{BL}(B, p)$ if we ‘rescale / change coordinates’.
- Fix invertible matrices M and $M_1, \ldots, M_m$, and consider the Brascamp-Lieb inequality with matrices $B_i' = M_i B_i M$,

$$\int_{\mathbf{R}^n} \prod_{i=1}^m |f_i(M_i B_i M x)|^{p_i} \, dx \leq \mathrm{BL}(B', p) \cdot \prod_{i=1}^m \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

Operator Rescaling

- What happens to $\mathrm{BL}(B,p)$ if we 'rescale / change coordinates'.
- Fix invertible matrices M and $M_1,\ldots,M_m$, and consider the Brascamp-Lieb inequality with matrices $B_i' = M_i B_i M$,

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(M_i B_i M x)|^{p_i}\, dx \leq \mathrm{BL}(B',p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- Then

$$\begin{aligned}\mathrm{BL}(B',p) &= \sup_{A_1,\ldots,A_m \succ 0} \frac{\prod_i \det(A_i)^{p_i}}{\det(\sum_i p_i M^* B_i^* M_i^* A_i M_i B_i M)} \\ &= \sup_{A_1,\ldots,A_m \succ 0} \frac{\prod_i \det((M_i^{-1})^* A_i M_i^{-1})^{p_i}}{\det(M^*(\sum_i p_i B_i^* A_i B_i)M)} \\ &= \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B,p).\end{aligned}$$

Operator Rescaling

$$\mathrm{BL}(B', p) = \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B, p)$$

Operator Rescaling

$$\mathrm{BL}(B', p) = \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B, p)$$

- If (B', p) is geometric, $\mathrm{BL}(B', p) = 1$, so

$$\mathrm{BL}(B, p) = \det(M)^2 \prod_i \det(M_i)^{2p_i}.$$

Operator Rescaling

$$\mathrm{BL}(B', p) = \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B, p)$$

- If (B', p) is geometric, $\mathrm{BL}(B', p) = 1$, so

$$\mathrm{BL}(B, p) = \det(M)^2 \prod_i \det(M_i)^{2p_i}.$$

- (Bennett, Carbery, Christ, Tao, 2008) Geometric rescaling possible iff the Brascamp-Lieb equation has extremizers.

Operator Rescaling

$$\mathrm{BL}(B', p) = \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B, p)$$

- If (B', p) is geometric, $\mathrm{BL}(B', p) = 1$, so
$$\mathrm{BL}(B, p) = \det(M)^{2} \prod_i \det(M_i)^{2p_i}.$$
- (Bennett, Carbery, Christ, Tao, 2008) Geometric rescaling possible iff the Brascamp-Lieb equation has extremizers.
- (Garg, Gurvits, Oliveira, Wigderson, 2018) If $\mathrm{BL}(B, p) < \infty$, we can rescale to be *within* a ε of a geometric Brascamp-Lieb equation.

Operator Rescaling

$$\mathrm{BL}(B', p) = \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B, p)$$

- If (B', p) is geometric, $\mathrm{BL}(B', p) = 1$, so

$$\mathrm{BL}(B, p) = \det(M)^2 \prod_i \det(M_i)^{2p_i}.$$

- (Bennett, Carbery, Christ, Tao, 2008) Geometric rescaling possible iff the Brascamp-Lieb equation has extremizers.
- (Garg, Gurvits, Oliveira, Wigderson, 2018) If $\mathrm{BL}(B, p) < \infty$, we can rescale to be *within* a ε of a geometric Brascamp-Lieb equation.
  - We can do this algorithmically, i.e. a computer can compute a ε-approximate geometric rescaling in $\mathrm{Poly}(\mathrm{Bits}(B), \log(p), 1/\varepsilon)$ computations.

Operator Rescaling

$$\mathrm{BL}(B', p) = \det(M)^{-2} \prod_i \det(M_i)^{-2p_i} \cdot \mathrm{BL}(B, p)$$

- If (B', p) is geometric, $\mathrm{BL}(B', p) = 1$, so

$$\mathrm{BL}(B, p) = \det(M)^2 \prod_i \det(M_i)^{2p_i}.$$

- (Bennett, Carbery, Christ, Tao, 2008) Geometric rescaling possible iff the Brascamp-Lieb equation has extremizers.
- (Garg, Gurvits, Oliveira, Wigderson, 2018) If $\mathrm{BL}(B, p) < \infty$, we can rescale to be *within* a ε of a geometric Brascamp-Lieb equation.
  - We can do this algorithmically, i.e. a computer can compute a ε-approximate geometric rescaling in $\mathrm{Poly}(\mathrm{Bits}(B), \log(p), 1/\varepsilon)$ computations.
  - Conversely, we can determine if $\mathrm{BL}(B, p) = \infty$ in $\mathrm{Poly}(\mathrm{Bits}(B), \log(p))$ computations.

Computing Permanents (An Analogous Problem)

- Given an $n \times n$ matrix A with non-negative entries, define

$$\mathrm{Perm}(A) = \sum_{\sigma \in S_n} \prod_i A_{i\sigma(i)}.$$

Computing Permanents (An Analogous Problem)

- Given an $n \times n$ matrix A with non-negative entries, define

$$\mathrm{Perm}(A) = \sum_{\sigma \in S_n} \prod_i A_{i\sigma(i)}.$$

- This is (maybe suprisingly) NP hard to compute.

Computing Permanents (An Analogous Problem)

- Given an $n \times n$ matrix A with non-negative entries, define

$$\mathrm{Perm}(A) = \sum\nolimits_{\sigma \in S_n} \prod\nolimits_i A_{i\sigma(i)}.$$

- This is (maybe suprisingly) NP hard to compute.
  - If $R = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $C = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$, then

$$\begin{aligned} \mathrm{Perm}(RAC) &= (\lambda_1 \ldots \lambda_n)(\gamma_1 \ldots \gamma_n)\mathrm{Perm}(A) \\ &= \det(R) \cdot \det(C) \cdot \mathrm{Perm}(A). \end{aligned}$$

Computing Permanents (An Analogous Problem)

- Given an $n \times n$ matrix A with non-negative entries, define

$$\text{Perm}(A) = \sum_{\sigma \in S_n} \prod_i A_{i\sigma(i)}.$$

- This is (maybe suprisingly) NP hard to compute.
  - If $R = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and $C = \text{diag}(\gamma_1, \ldots, \gamma_n)$, then

$$\begin{aligned}\text{Perm}(RAC) &= (\lambda_1 \ldots \lambda_n)(\gamma_1 \ldots \gamma_n)\text{Perm}(A) \\ &= \det(R) \cdot \det(C) \cdot \text{Perm}(A).\end{aligned}$$

  - (Egorychev, 1981), (Falikman, 1981) If A is *doubly stochastic*, $e^{-n} \leq \text{Perm}(A) \leq 1$.

Computing Permanents (An Analogous Problem)

- Given an $n \times n$ matrix A with non-negative entries, define

$$\mathrm{Perm}(A) = \sum_{\sigma \in S_n} \prod_i A_{i\sigma(i)}.$$

- This is (maybe suprisingly) NP hard to compute.
  - If $R = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ and $C = \mathrm{diag}(\gamma_1, \ldots, \gamma_n)$, then

$$\begin{aligned}\mathrm{Perm}(RAC) &= (\lambda_1 \ldots \lambda_n)(\gamma_1 \ldots \gamma_n)\mathrm{Perm}(A) \\ &= \det(R) \cdot \det(C) \cdot \mathrm{Perm}(A).\end{aligned}$$

  - (Egorychev, 1981), (Falikman, 1981) If A is *doubly stochastic*, $e^{-n} \leq \mathrm{Perm}(A) \leq 1$.
  - If RAC is doubly stochastic, then

$$\mathrm{Perm}(A) \leq \det(R)^{-1} \det(C)^{-1} \leq e^n \cdot \mathrm{Perm}(A),$$

  so $\mathrm{Perm}(A) \approx \det(R)^{-1} \det(C)^{-1}$.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots .$$

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots.$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots.$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \dots, \gamma_{in})$, be the row / column sums of A_i.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots .$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \dots, \gamma_{in})$, be the row / column sums of A_i.
  - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \ldots .$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \ldots, \gamma_{in})$, be the row / column sums of A_i.
  - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.
- **Two key facts**:

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots.$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \dots, \gamma_{in})$, be the row / column sums of A_i.
  - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.
- **Two key facts**:

  (1) $\mathrm{Per}(A_i) \leq 1$ for all $i \geq 1$.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \ldots.$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \ldots, \gamma_{in})$, be the row / column sums of A_i.
  - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.
- **Two key facts**:
  (1) $\mathrm{Per}(A_i) \leq 1$ for all $i \geq 1$.
  (2) Let $\Delta_i = \sum_j (\gamma_{ij} - 1)^2$.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots.$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \dots, \gamma_{in})$, be the row / column sums of A_i.
  - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.
- **Two key facts**:
  (1) $\mathrm{Per}(A_i) \leq 1$ for all $i \geq 1$.
  (2) Let $\Delta_i = \sum_j (\gamma_{ij} - 1)^2$.
    - If $\Delta_i \leq 1$, then $\mathrm{Per}(A_{i+1}) \geq (1 + C\Delta_i) \cdot \mathrm{Per}(A_i)$.
    - If $\Delta_i \geq 1$, then $\mathrm{Per}(A_{i+1}) \geq (1 + C)\mathrm{Per}(A_i)$.

- We continue this iteration, obtaining a sequence
$$A \to A_1 \to A_2 \to \dots.$$
- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \dots, \gamma_{in})$, be the row / column sums of A_i.
  - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.
- **Two key facts**:
  (1) $\mathrm{Per}(A_i) \le 1$ for all $i \ge 1$.
  (2) Let $\Delta_i = \sum_j (\gamma_{ij} - 1)^2$.
    - If $\Delta_i \le 1$, then $\mathrm{Per}(A_{i+1}) \ge (1 + C\Delta_i) \cdot \mathrm{Per}(A_i)$.
    - If $\Delta_i \ge 1$, then $\mathrm{Per}(A_{i+1}) \ge (1 + C)\mathrm{Per}(A_i)$.
- Thus $\mathrm{Per}(A_i)$ is bounded, monotonic, converges to $P \le 1$.

- We continue this iteration, obtaining a sequence

$$A \to A_1 \to A_2 \to \dots.$$

- Claim: If $\mathrm{Per}(A) > 0$, these matrices become arbitrarily close to being doubly stochastic as $i \to \infty$.
- Let $\gamma_i = (\gamma_{i1}, \dots, \gamma_{in})$, be the row / column sums of A_i.
    - i.e. so that $A_1 = \mathrm{diag}(1/\gamma_0)A$, $A_2 = A_1\mathrm{diag}(1/\gamma_1)$, etc.
- **Two key facts**:
    (1) $\mathrm{Per}(A_i) \leq 1$ for all $i \geq 1$.
    (2) Let $\Delta_i = \sum_j (\gamma_{ij} - 1)^2$.
        - If $\Delta_i \leq 1$, then $\mathrm{Per}(A_{i+1}) \geq (1 + C\Delta_i) \cdot \mathrm{Per}(A_i)$.
        - If $\Delta_i \geq 1$, then $\mathrm{Per}(A_{i+1}) \geq (1 + C)\mathrm{Per}(A_i)$.
- Thus $\mathrm{Per}(A_i)$ is bounded, monotonic, converges to $P \leq 1$.
- If $\mathrm{Per}(A_i) \geq P - \varepsilon$ for $\varepsilon \ll 1$, then

$$P \geq \mathrm{Per}(A_{i+1}) \geq (1+C\cdot\Delta_i)\cdot\mathrm{Per}(A_i) \geq (1+C\cdot\Delta_i)(P-\varepsilon).$$

Thus $\Delta_i \leq (C_0/P)\varepsilon$. Taking $\varepsilon \to 0$ shows $\Delta \to 0$.

Sinkhorn Iteration

- **Proof that Per$(A_i) \leq 1$:**

Sinkhorn Iteration

- **Proof that Per$(A_i) \leq 1$:**
  - Simple inductive argument: the hypothesis is that if the rows of a matrix B all sum up to less than one, then Per$(B) \leq 1$. Now expand along minors (ala the determinant).

Sinkhorn Iteration

- **Proof that Per(A_i) ≤ 1:**
  - Simple inductive argument: the hypothesis is that if the rows of a matrix B all sum up to less than one, then $\text{Per}(B) \leq 1$. Now expand along minors (ala the determinant).
- **Proof that Per(A_{i+1}) $\geq (1 + C\Delta_i) \cdot$ Per(A_i):**

Sinkhorn Iteration

- **Proof that Per**$(A_i) \leq 1$**:**
  - Simple inductive argument: the hypothesis is that if the rows of a matrix B all sum up to less than one, then $\mathrm{Per}(B) \leq 1$. Now expand along minors (ala the determinant).
- **Proof that Per**$(A_{i+1}) \geq (1 + C\Delta_i) \cdot$ **Per**(A_i)**:**
  - Write $\gamma_i = 1 + \delta_i$.

Sinkhorn Iteration

- **Proof that Per**$(A_i) \leq 1$**:**
  - Simple inductive argument: the hypothesis is that if the rows of a matrix B all sum up to less than one, then $\text{Per}(B) \leq 1$. Now expand along minors (ala the determinant).
- **Proof that Per**$(A_{i+1}) \geq (1 + C\Delta_i) \cdot$ **Per**(A_i)**:**
  - Write $\gamma_i = 1 + \delta_i$.
    - Then $\sum \delta_i^2 = \Delta_i$, and $\sum \delta_i = 0$.

Sinkhorn Iteration

- **Proof that Per**$(A_i) \leq 1$**:**
  - Simple inductive argument: the hypothesis is that if the rows of a matrix B all sum up to less than one, then $\text{Per}(B) \leq 1$. Now expand along minors (ala the determinant).
- **Proof that Per**$(A_{i+1}) \geq (1 + C\Delta_i) \cdot$ **Per**(A_i)**:**
  - Write $\gamma_i = 1 + \delta_i$.
    - Then $\sum \delta_i^2 = \Delta_i$, and $\sum \delta_i = 0$.
  - Since $1 + t \leq \exp(t - t^2/2 + t^3/3)$,

$$\begin{aligned}
\text{Per}(A_i)/\text{Per}(A_{i+1}) &= \gamma_1 \dots \gamma_n \\
&= (1+\delta_1)\dots(1+\delta_n) \\
&\leq \exp\left(\sum \delta_i - \sum \delta_i^2/2 + \sum \delta_i^3/3\right) \\
&\leq \exp(0 - \Delta/2 + \Delta^{3/2}/3) \\
&= 1 - \Delta/2 + O(\Delta^{3/2}).
\end{aligned}$$

And now, back to our regularly scheduled programming

$$\mathrm{BL}(B, p) = \sup_{A_1,\ldots,A_m \succ 0} \sqrt{\frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i \cdot B_i^* A_i B_i)}}.$$

And now, back to our regularly scheduled programming

$$\mathrm{BL}(B, p) = \sup_{A_1, \ldots, A_m \succ 0} \sqrt{\frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i \cdot B_i^* A_i B_i)}}.$$

- Goal: Rescale our inputs so that

And now, back to our regularly scheduled programming

$$\mathrm{BL}(B,p) = \sup_{A_1,\ldots,A_m \succ 0} \sqrt{\frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i \cdot B_i^* A_i B_i)}}.$$

- Goal: Rescale our inputs so that
  - (Isotropy) $\sum p_i B_i^* B_i = I$.

And now, back to our regularly scheduled programming

$$\mathrm{BL}(B, p) = \sup_{A_1,\ldots,A_m \succ 0} \sqrt{\frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i \cdot B_i^* A_i B_i)}}.$$

- Goal: Rescale our inputs so that
  - (Isotropy) $\sum p_i B_i^* B_i = I$.
  - (Projection) $B_i B_i^* = I$ for each i.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.
    - Then $\sum p_i (B_i')^* B_i' = 1$, i.e. isotropy holds.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.
    - Then $\sum p_i (B_i')^* B_i' = 1$, i.e. isotropy holds.
  - (Projection Normalization)

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.
    - Then $\sum p_i (B_i')^* B_i' = 1$, i.e. isotropy holds.
  - (Projection Normalization)
    - Let $M_i = B_i B_i^*$.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.
    - Then $\sum p_i (B_i')^* B_i' = 1$, i.e. isotropy holds.
  - (Projection Normalization)
    - Let $M_i = B_i B_i^*$.
    - Replace B_i with $B_i' = M_i^{-1/2} B_i$.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.
    - Then $\sum p_i (B_i')^* B_i' = 1$, i.e. isotropy holds.
  - (Projection Normalization)
    - Let $M_i = B_i B_i^*$.
    - Replace B_i with $B_i' = M_i^{-1/2} B_i$.
    - Then $(B_i')^* B_i' = I$ for each i.

Iteration of Operator Rescaling

- Sinkhorn: Alternately apply the following two procedures:
  - (Isotropy Normalization)
    - Let $M = \sum_i p_i B_i^* B_i$.
    - Replace B_i with $B_i' = B_i M^{-1/2}$.
    - Then $\sum p_i (B_i')^* B_i' = 1$, i.e. isotropy holds.
  - (Projection Normalization)
    - Let $M_i = B_i B_i^*$.
    - Replace B_i with $B_i' = M_i^{-1/2} B_i$.
    - Then $(B_i')^* B_i' = I$ for each i.
  - We obtain a sequence $B \to B_1 \to B_2 \to \ldots$.

Iteration of Operator Rescaling

- **Three Key Facts Ensuring Convergence:**

Iteration of Operator Rescaling

- **Three Key Facts Ensuring Convergence:**
  (1) $\mathrm{BL}(B_i, p) \geq 1$ for all $i \geq 1$.

Iteration of Operator Rescaling

- **Three Key Facts Ensuring Convergence:**
  (1) $\mathrm{BL}(B_i, p) \geq 1$ for all $i \geq 1$.
  (2) :
    - If $1 + \varepsilon \leq \mathrm{BL}(B_i, p) \leq 2$,
      $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_1 \varepsilon^k)\mathrm{BL}(B_i, p)$.
    - If $\mathrm{BL}(B_i, p) > 2$, then $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_2)\mathrm{BL}(B_i, p)$.

Iteration of Operator Rescaling

- **Three Key Facts Ensuring Convergence:**
  (1) $\mathrm{BL}(B_i, p) \geq 1$ for all $i \geq 1$.
  (2) :
    - If $1 + \varepsilon \leq \mathrm{BL}(B_i, p) \leq 2$,
      $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_1 \varepsilon^k)\mathrm{BL}(B_i, p)$.
    - If $\mathrm{BL}(B_i, p) > 2$, then $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_2)\mathrm{BL}(B_i, p)$.

  (3) If isotropy or projection holds, and $\mathrm{BL}(B_i, p) \leq 2$, then

$$\| \sum p_i B_i^* B_i - I \| + \sum_i \| B_i^* B_i - I \| \lesssim \log(\mathrm{BL}(B, p)).$$

Iteration of Operator Rescaling

- **Three Key Facts Ensuring Convergence:**
  (1) $\mathrm{BL}(B_i, p) \geq 1$ for all $i \geq 1$.
  (2) :
    - If $1 + \varepsilon \leq \mathrm{BL}(B_i, p) \leq 2$,
      $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_1\varepsilon^k)\mathrm{BL}(B_i, p)$.
    - If $\mathrm{BL}(B_i, p) > 2$, then $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_2)\mathrm{BL}(B_i, p)$.

  (3) If isotropy or projection holds, and $\mathrm{BL}(B_i, p) \leq 2$, then

$$\| \sum p_i B_i^* B_i - I\| + \sum_i \|B_i^* B_i - I\| \lesssim \log(\mathrm{BL}(B, p)).$$

- Thus convergence to the family of geometric Brascamp-Lieb datum occurs as with Sinkhorn iteration provided that $\mathrm{BL}(B, p) < \infty$.

Iteration of Operator Rescaling

- **Three Key Facts Ensuring Convergence:**
  (1) $\mathrm{BL}(B_i, p) \geq 1$ for all $i \geq 1$.
  (2) :
    - If $1 + \varepsilon \leq \mathrm{BL}(B_i, p) \leq 2$, $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_1\varepsilon^k)\mathrm{BL}(B_i, p)$.
    - If $\mathrm{BL}(B_i, p) > 2$, then $\mathrm{BL}(B_{i+1}, p) \leq (1 - C_2)\mathrm{BL}(B_i, p)$.

  (3) If isotropy or projection holds, and $\mathrm{BL}(B_i, p) \leq 2$, then

$$\|\sum p_i B_i^* B_i - I\| + \sum_i \|B_i^* B_i - I\| \lesssim \log(\mathrm{BL}(B, p)).$$

- Thus convergence to the family of geometric Brascamp-Lieb datum occurs as with Sinkhorn iteration provided that $\mathrm{BL}(B, p) < \infty$.
- Obtain (1), (2), and (3) by studying *positive operators*.

Another Viewpoint: Positive Operators

- For simplicity, assume that all spaces have the same ambient dimension (all B_i are square matrices).

Another Viewpoint: Positive Operators

- For simplicity, assume that all spaces have the same ambient dimension (all B_i are square matrices).
- $\mathrm{BL}(B, p) < \infty$ can only hold if $\sum p_i = 1$.

Another Viewpoint: Positive Operators

- For simplicity, assume that all spaces have the same ambient dimension (all B_i are square matrices).
- $\mathrm{BL}(B, p) < \infty$ can only hold if $\sum p_i = 1$.
- Also assume all A_i are equal, and let us consider optimizing the quantity

$$\inf_{A \succ 0} \frac{\det(\sum p_i B_i^* A B_i)}{\det(A)}$$

analogous to

$$\mathrm{BL}(B, p) = \sup_{A_1, \ldots, A_m \succ 0} \sqrt{\frac{\prod_i \det(A_i)^{p_i}}{\det(\sum p_i \cdot B_i^* A_i B_i)}}.$$

Positive Operators

- A linear map $T : M_n \to M_m$ is *completely positive* if there are $m \times n$ matrices $B_1, \ldots, B_K$ such that

$$T(A) = \sum B_i A B_i^*.$$

For simplicity, focus on the case $n = m$.

Positive Operators

- A linear map $T : M_n \to M_m$ is *completely positive* if there are $m \times n$ matrices $B_1, \ldots, B_K$ such that

$$T(A) = \sum B_i A B_i^*.$$

  For simplicity, focus on the case $n = m$.
- Important example: $T(A) = \sum p_i B_i^* A B_i$.

Positive Operators

- A linear map $T : M_n \to M_m$ is *completely positive* if there are $m \times n$ matrices $B_1, \dots, B_K$ such that

$$T(A) = \sum B_i A B_i^*.$$

  For simplicity, focus on the case $n = m$.
- Important example: $T(A) = \sum p_i B_i^* A B_i$.
- Given T, we have $T^*(A) = \sum B_i^* A B_i$.

Capacity of Operators

$$\inf_{A \succ 0} \frac{\det(\sum p_i B_i^* A B_i)}{\det(A)}$$

Capacity of Operators

$$\inf_{A \succ 0} \frac{\det(\sum p_i B_i^* A B_i)}{\det(A)}$$

- The *capacity* of a positive operator $T : M_n \to M_n$ is

$$\mathrm{Cap}(T) = \inf_{A \succ 0} \frac{\det(TA)}{\det(A)}.$$

Capacity of Operators

$$\inf_{A \succ 0} \frac{\det(\sum p_i B_i^* A B_i)}{\det(A)}$$

- The *capacity* of a positive operator $T : M_n \to M_n$ is

$$\mathrm{Cap}(T) = \inf_{A \succ 0} \frac{\det(TA)}{\det(A)}.$$

- Capacity introduced in (Gurvits, 2004) in $n = m$ case, in general introduced in (Garg et al, 2018).

Capacity of Operators

$$\inf_{A \succ 0} \frac{\det(\sum p_i B_i^* A B_i)}{\det(A)}$$

- The *capacity* of a positive operator $T : M_n \to M_n$ is

$$\mathrm{Cap}(T) = \inf_{A \succ 0} \frac{\det(TA)}{\det(A)}.$$

- Capacity introduced in (Gurvits, 2004) in $n = m$ case, in general introduced in (Garg et al, 2018).
- For any Brascamp-Lieb data (B, p), there exists a positive $T : M_n \to M_m$ such that $\mathrm{Cap}(T) = 1/\mathrm{BL}(B, p)^2$.

Capacity of Operators

$$\inf_{A \succ 0} \frac{\det(\sum p_i B_i^* A B_i)}{\det(A)}$$

- The *capacity* of a positive operator $T : M_n \to M_n$ is

$$\mathrm{Cap}(T) = \inf_{A \succ 0} \frac{\det(TA)}{\det(A)}.$$

- Capacity introduced in (Gurvits, 2004) in $n = m$ case, in general introduced in (Garg et al, 2018).
- For any Brascamp-Lieb data (B, p), there exists a positive $T : M_n \to M_m$ such that $\mathrm{Cap}(T) = 1/\mathrm{BL}(B, p)^2$.
- Positive operators are well studied in the quantum information theory literature, so reduction of BL to this theory gives new insights.

Doubly Stochastic Positive Operators

- (Isotropy) Let $T(A) = \sum p_i B_i^* A B_i$.

Doubly Stochastic Positive Operators

- (Isotropy) Let $T(A) = \sum p_i B_i^* A B_i$.
  - $\sum p_i B_i^* B_i = I$ holds iff $T(I) = I$.

Doubly Stochastic Positive Operators

- (Isotropy) Let $T(A) = \sum p_i B_i^* A B_i$.
  - $\sum p_i B_i^* B_i = I$ holds iff $T(I) = I$.
- If (B, p) is a Brascamp-Lieb datum with associated operator $T : M_n \to M_m$, then (B, p) is geometric if and only if T is *doubly stochastic*. For $n = m$ this means $T(I) = I$ and $T^*(I) = I$.

Operator Rescaling for Positive Operators

- If T is doubly stochastic, $\mathrm{Cap}(T) = 1$.

Operator Rescaling for Positive Operators

- If T is doubly stochastic, $\mathrm{Cap}(T) = 1$.
- We can rescale. If

$$T_{M_1 M_2}(A) = M_2^* T(M_1^* A M_1) M_2,$$

then $\mathrm{Cap}(T_{M_1,M_2}) = \det(M_1)^2 \det(M_2)^2 \cdot \mathrm{Cap}(T)$.

Operator Rescaling for Positive Operators

- If T is doubly stochastic, $\text{Cap}(T) = 1$.
- We can rescale. If

$$T_{M_1 M_2}(A) = M_2^* T(M_1^* A M_1) M_2,$$

  then $\text{Cap}(T_{M_1,M_2}) = \det(M_1)^2 \det(M_2)^2 \cdot \text{Cap}(T)$.
- Sinkhorn iteration (alternately iterating $T \mapsto T_{I, T(I)^{-1/2}}$ and $T \mapsto T_{T^*(I)^{-1/2}, I}$) yields a method for rescaling any T with $\text{Cap}(T) > 0$ to be arbitrarily close to a doubly stochastic operator, allowing us to approximate $\text{Cap}(T)$.

Rank Decreasing Operators

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

Rank Decreasing Operators

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \le \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Bennett et al, 2008) implies that $\mathrm{BL}(B, p) < \infty$ if and only if $\sum p_i n_i = n$, and for any subspace $V \subset \mathbf{R}^n$,

$$\dim(V) \le \sum p_i \dim(B_i V).$$

Rank Decreasing Operators

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Bennett et al, 2008) implies that $\mathrm{BL}(B, p) < \infty$ if and only if $\sum p_i n_i = n$, and for any subspace $V \subset \mathbf{R}^n$,

$$\dim(V) \leq \sum p_i \dim(B_i V).$$

- An operator $T : M_n \to M_n$ is *rank non-decreasing* if for any $A \succeq 0$, $\mathrm{Rank}(TA) \geq \mathrm{Rank}(A)$.

Rank Decreasing Operators

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Bennett et al, 2008) implies that $\mathrm{BL}(B, p) < \infty$ if and only if $\sum p_i n_i = n$, and for any subspace $V \subset \mathbf{R}^n$,

$$\dim(V) \leq \sum p_i \dim(B_i V).$$

- An operator $T : M_n \to M_n$ is *rank non-decreasing* if for any $A \succeq 0$, $\mathrm{Rank}(TA) \geq \mathrm{Rank}(A)$.
- (Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.

Rank Decreasing Operators

$$\int_{\mathbf{R}^n} \prod_{i=1}^{m} |f_i(B_i x)|^{p_i} \, dx \leq \mathrm{BL}(B, p) \cdot \prod_{i=1}^{m} \|f_i\|_{L^1(\mathbf{R}^{n_i})}^{p_i}.$$

- (Bennett et al, 2008) implies that $\mathrm{BL}(B, p) < \infty$ if and only if $\sum p_i n_i = n$, and for any subspace $V \subset \mathbf{R}^n$,

$$\dim(V) \leq \sum p_i \dim(B_i V).$$

- An operator $T : M_n \to M_n$ is *rank non-decreasing* if for any $A \succeq 0$, $\mathrm{Rank}(TA) \geq \mathrm{Rank}(A)$.
- (Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.
- Generalized in (Garg et al, 2018). For $T : M_n \to M_m$, $\mathrm{Cap}(T) > 0$ if and only if T is *fractional rank non-decreasing*.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\text{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\text{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.
- Given an orthonormal basis $U = \{u_1, \ldots, u_N\}$, we define the *decoherence operator* $D_U(A) = \sum \langle Au_i, u_i \rangle \cdot u_i u_i^*$.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.
- Given an orthonormal basis $U = \{u_1, \ldots, u_N\}$, we define the *decoherence operator* $D_U(A) = \sum \langle Au_i, u_i \rangle \cdot u_i u_i^*$.
- Define $T_U = D_U \circ T$.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.
- Given an orthonormal basis $U = \{u_1, \ldots, u_N\}$, we define the *decoherence operator* $D_U(A) = \sum \langle A u_i, u_i \rangle \cdot u_i u_i^*$.
- Define $T_U = D_U \circ T$.
- Result follows from the following two facts:

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\text{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.
- Given an orthonormal basis $U = \{u_1, \ldots, u_N\}$, we define the *decoherence operator* $D_U(A) = \sum \langle A u_i, u_i \rangle \cdot u_i u_i^*$.
- Define $T_U = D_U \circ T$.
- Result follows from the following two facts:
  (1) T is rank non-decreasing if and only if T_U is rank non-decreasing for all U.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.
- Given an orthonormal basis $U = \{u_1, \dots, u_N\}$, we define the *decoherence operator* $D_U(A) = \sum \langle Au_i, u_i \rangle \cdot u_i u_i^*$.
- Define $T_U = D_U \circ T$.
- Result follows from the following two facts:
  1. T is rank non-decreasing if and only if T_U is rank non-decreasing for all U.
  2. $\mathrm{Cap}(T) = \inf_U \mathrm{Cap}(T_U)$.

Proof Idea

(Gurvits, 2004) $T : M_n \to M_n$ is rank non-decreasing if and only if $\mathrm{Cap}(T) > 0$.

- Results from (Gurvits and Samorodnitsky, 2002) show the result is true if $T(X) = \sum X_{ii} A_i$, where $A_i \succeq 0$. The general case can be reduced to this case.
- Given an orthonormal basis $U = \{u_1, \ldots, u_N\}$, we define the *decoherence operator* $D_U(A) = \sum \langle Au_i, u_i \rangle \cdot u_i u_i^*$.
- Define $T_U = D_U \circ T$.
- Result follows from the following two facts:
  (1) T is rank non-decreasing if and only if T_U is rank non-decreasing for all U.
  (2) $\mathrm{Cap}(T) = \inf_U \mathrm{Cap}(T_U)$.
- To prove (1) and (2), use a simple trick: Given $A \succeq 0$, find U diagonalizing $T(A)$. Then $T(A) = T_U(A)$.

Thanks For Listening!