Ankit Garg
Leonid Gurvits
Rafael Oliveira
Avi Wigderson

Scaling Algorithms & Det. Approx. of Capacity and BL Constant

Today

Scaling Problems

Scaling Algorithms

Approx. Perm & Capacity

Brascamp-Lieb Primer

Conclusion & More

Scaling Problems - Motivation

Why should anyone care?

- *Communication Complexity:* gen. Forster's sign-rank lower bounds
- ***Algorithms:*** det. approx. to Perm. non-neg. matrices & mixed volume
- *Coding Theory:* lower bounds on LCC's over $\mathbb{R}$
- *Optimization:* Brascamp-Lieb & moment polytopes, non. comm. Duality
- *Operator Theory:* Paulsen problem
- *Quantum Information Theory:* Entanglement distillation
- ***Functional Analysis:*** Brascamp-Lieb inequalities
- *Algebraic Complexity:* non-commutative PIT, asymptotic Kronecker
- *Extremal combinatorics:* quantitative gen. of Sylvester-Gallai thms, asymptotic slice-rank
- Many more (invariant theory, representation theory, opt. transport…)

Matrix Scaling

$n \times n$ non-neg. matrix A is **doubly stochastic (DS)** if sum of rows/columns of A are equal to $\mathbf{1}$.

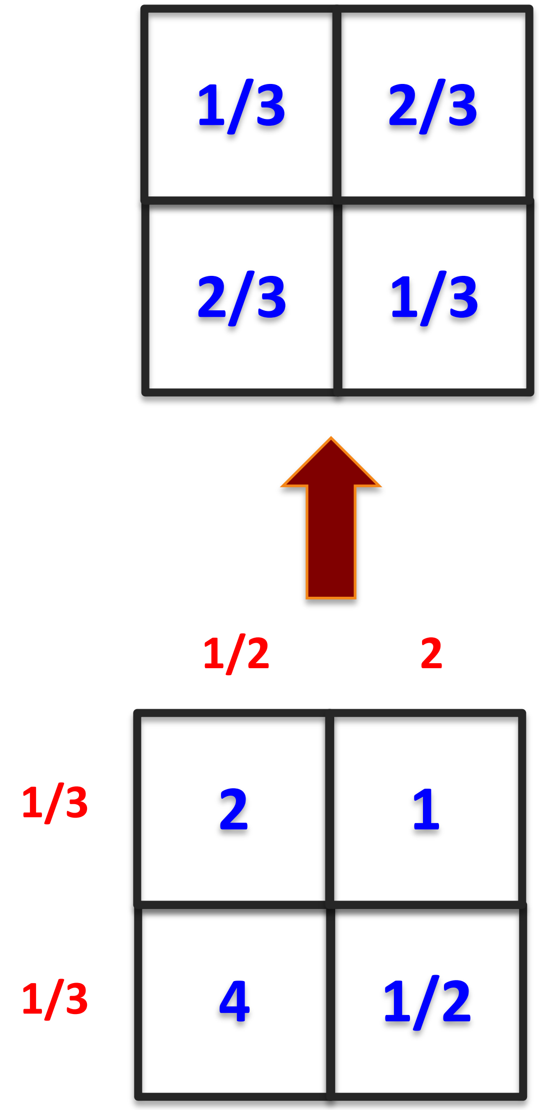
$\boldsymbol{B}$ is **scaling** of $\boldsymbol{A}$ if $\exists$ positive $\boldsymbol{x_1, \ldots, x_n, y_1, \ldots, y_n}$ such that $\boldsymbol{b_{ij} = x_i a_{ij} y_j}$.

$\boldsymbol{A}$ has DS scaling if there is DS scaling $\boldsymbol{B}$ of $\boldsymbol{A}$.

$$\boldsymbol{ds(A) = \sum_i (r_i - 1)^2 + \sum_j (c_j - 1)^2}$$

$\boldsymbol{A}$ has approx. DS scaling if $\forall \boldsymbol{\epsilon > 0}$ there is scaling $\boldsymbol{B_\epsilon}$ of $\boldsymbol{A}$ s.t. $\mathbf{ds}(\mathbf{B}_\epsilon) < \epsilon$.

1. When does A have approx. DS scaling?
2. Can we find it efficiently?



Matrix Scaling – examples (alg. & geom.)

	$\sqrt{2}$	1
$\sqrt{2}-1$	1	1
$(2+\sqrt{2})^{-1}$	1	2

↓

$2-\sqrt{2}$	$\sqrt{2}-1$
$\sqrt{2}-1$	$2-\sqrt{2}$

	$1/\epsilon$	ϵ
ϵ	1	1
$1/\epsilon$	0	1

↓

1	ϵ^2
0	1

Matrix Scaling – Algorithm S

Problem: $A \in M_n(\mathbb{R}_{\geq 0})$, $\epsilon > 0$, is there ϵ-scaling to DS? If yes, find it.

Algorithm S [Kruithof'37, …, Sinkhorn'64]:

Repeat k times:

1. Normalize rows of A (make row sums equal)
2. Normalize columns of A (make col sums equal)

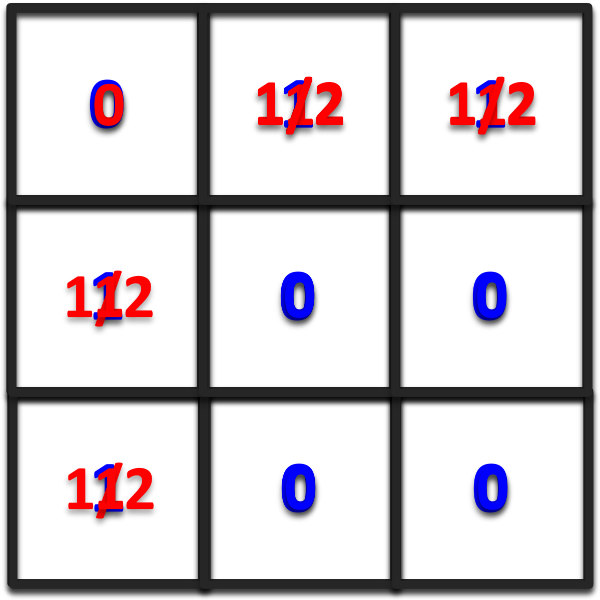
If at any point $\mathbf{ds}(A) < \epsilon$, output the scaling so far.

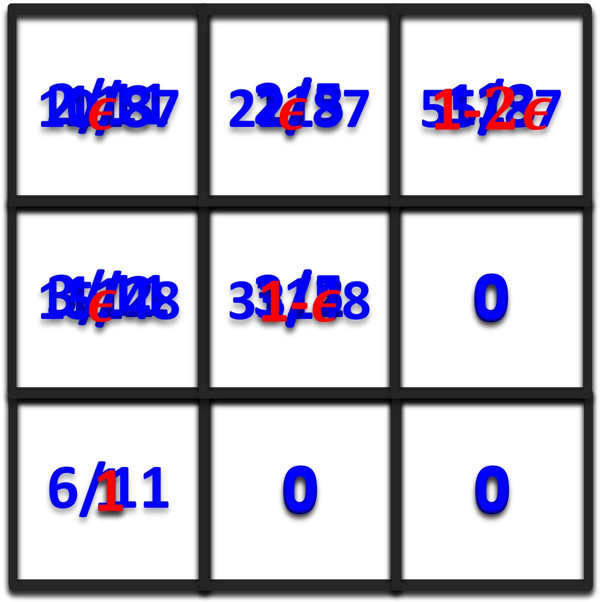
Else, output: **no scaling.**

Questions:

- Are we making progress at all?
- How do we know when to stop? (Which k?)
- Is there $\epsilon_0 > 0$ s.t. if $\mathbf{ds}(A) < \epsilon_0$ can get DS for any $\epsilon > 0$?

Algorithm S – Two Examples





Question: How can we distinguish between these two cases?

Observation: In first example, have no matchings (and Hall blocker).

Are these the only bad cases?

Algorithm S – Analysis [LSW'00]

Algorithm S:

Repeat k times:

1. Normalize rows of A
2. Normalize columns of A

If at any point $ds(A) \leq \epsilon$, output the scaling so far.

Else, output: **no scaling.**

Analysis [LSW'00]:

1. $Per(A) > 0 \Rightarrow Per(A) > \nu^{-n}$
2. $\mathrm{ds}(\mathrm{A}) > \epsilon \ \Rightarrow Per(A)$ grows by $\exp(O(\epsilon))$ after normalization
3. $Per(A) \leq 1$ for any normalized matrix

Within $\mathbf{k} = poly(n/\epsilon)$ iterations we will get our scaling!

$\mathrm{Per}(A) > 0 \Leftrightarrow A$ has matching (also no Hall blocker), so correct.

Bounding ϵ_0

$\mathrm{Per}(A) = 0 \Leftrightarrow A$ has no matching (and a Hall blocker).

See board.

Quantum Operators – Definition

A **quantum operator** is any map $\mathbf{T}: \mathbf{M}_n(\mathbb{C}) \to \boldsymbol{M}_n(\mathbb{C})$ given by $(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m)$ s.t.

$$\boldsymbol{T}(\boldsymbol{X}) = \sum_{1 \le i \le m} \boldsymbol{A}_i \boldsymbol{X} \boldsymbol{A}_i^{\dagger}$$

Such maps take psd matrices to psd matrices.

Dual of $\mathbf{T}(\mathbf{X})$ is map $\mathbf{T}^*: \mathbf{M}_n(\mathbb{C}) \to \boldsymbol{M}_n(\mathbb{C})$ given by:

$$\boldsymbol{T}^*(\boldsymbol{X}) = \sum_{1 \le i \le m} \boldsymbol{A}_i^{\dagger} \boldsymbol{X} \boldsymbol{A}_i$$

- Analog of scaling?
- Doubly stochastic?

Operator Scaling

A quantum operator $\mathbf{T}: \mathbf{M}_n(\mathbb{C}) \rightarrow \boldsymbol{M}_n(\mathbb{C})$ is **doubly stochastic (DS)** if $\boldsymbol{T}(\boldsymbol{I}) = \boldsymbol{T}^*(\boldsymbol{I}) = \boldsymbol{I}$.

Scaling of $\boldsymbol{T}(\boldsymbol{X})$ consists of $\boldsymbol{L}, \boldsymbol{R} \in \boldsymbol{GL}_n(\mathbb{C})$ s.t.

$$(\boldsymbol{A}_1, \dots, \boldsymbol{A}_m) \rightarrow (\boldsymbol{L}\boldsymbol{A}_1\boldsymbol{R}, \dots, \boldsymbol{L}\boldsymbol{A}_m\boldsymbol{R})$$

Distance to doubly-stochastic:

$$\boldsymbol{ds}(\boldsymbol{T}) \stackrel{\text{def}}{=} \|\boldsymbol{T}(\boldsymbol{I}) - \boldsymbol{I}\|_{\boldsymbol{F}}^{\mathbf{2}} + \|\boldsymbol{T}^*(\boldsymbol{I}) - \boldsymbol{I}\|_{\boldsymbol{F}}^{\mathbf{2}}$$

$\boldsymbol{T}(\boldsymbol{X})$ has approx. DS scaling if $\forall \boldsymbol{\epsilon} > \mathbf{0}$, $\exists$ scaling $\boldsymbol{L}_{\boldsymbol{\epsilon}}, \boldsymbol{R}_{\boldsymbol{\epsilon}}$ s.t. operator $\boldsymbol{T}_{\boldsymbol{\epsilon}}(\boldsymbol{X})$ given by $(\mathbf{L}_{\boldsymbol{\epsilon}}\boldsymbol{A}_1\boldsymbol{R}_{\boldsymbol{\epsilon}}, \dots, \boldsymbol{L}_{\boldsymbol{\epsilon}}\boldsymbol{A}_m\boldsymbol{R}_{\boldsymbol{\epsilon}})$ has $\boldsymbol{ds}(\boldsymbol{T}_{\boldsymbol{\epsilon}}) \leq \boldsymbol{\epsilon}$.

1. When does $(\boldsymbol{A}_1, \dots, \boldsymbol{A}_m)$ have approx. DS scaling?
2. Can we find it efficiently?

Generalizes Matrix Scaling

Take quantum operator

$$T_A = (\sqrt{a_{11}} \cdot E_{11}, \sqrt{a_{12}} \cdot E_{12}, \dots, \sqrt{a_{nn}} \cdot E_{nn})$$

and dual

$$T_A^* = (\sqrt{a_{11}} \cdot E_{11}, \sqrt{a_{21}} \cdot E_{12}, \dots, \sqrt{a_{nn}} \cdot E_{nn})$$

$$T_A(I) = \sum a_{ij} E_{ij} E_{ij}^\dagger = \sum a_{ij} E_{ii} = diag(r_1, \dots, r_n)$$

$$T_A^*(I) = \sum a_{ji} E_{ij} E_{ij}^\dagger = \sum a_{ji} E_{ii} = diag(c_1, \dots, c_n)$$

Distance to doubly-stochastic:

$$ds(T_A) \stackrel{\text{def}}{=} \|T(I) - I\|_F^2 + \|T^*(I) - I\|_F^2 = ds(A)$$

Operator Scaling – Algorithm G

Problem: operator $\mathbf{T} = (A_1, \ldots, A_m)$, $\epsilon > 0$, can T be ϵ-scaled to double stochastic? If yes, find scaling.

Algorithm G [Gurvits' 04]:

Repeat k times:

1. Left normalize $T(X)$, i.e.,$(A_1, \ldots, A_m) \leftarrow (LA_1, \ldots, LA_m)$ s.t. $T(I) = I$.
2. Right normalize $\mathbf{T}(\mathbf{X})$, i.e., $(A_1, \ldots, A_m) \leftarrow (A_1 R, \ldots, A_m R)$ s.t. $T^*(I) = I$.

If at any point $\mathbf{ds}(\mathbf{T}) \leq \epsilon$, output the current scaling.
Else output **no scaling**.

- Which k should we choose?

Algorithm G – Analysis

Algorithm G:

Repeat $\boldsymbol{k}$ times:

1. Left normalize: $(\boldsymbol{A_1}, \dots, \boldsymbol{A_m}) \leftarrow (\boldsymbol{RA_1}, \dots, \boldsymbol{RA_m})$ s.t. $\boldsymbol{T(I) = I}$.
2. Right normalize: $(\boldsymbol{A_1}, \dots, \boldsymbol{A_m}) \leftarrow (\boldsymbol{A_1C}, \dots, \boldsymbol{A_mC})$ s.t. $\boldsymbol{T^*(I) = I}$.

If at any point $\boldsymbol{ds(T) \le \epsilon}$, output current scaling.

Else output **no scaling**.

Potential Function (Capacity) [Gur'04]:

$$\boldsymbol{cap(T) = inf\left\{\frac{det\big(T(X)\big)}{det(X)} : \ X \succ 0\right\}}.$$

Analysis [Gur'04, GGOW'15]:

1. $\boldsymbol{cap(T) > 0 \Rightarrow cap(T) > e^{-poly(n)}}$ **(GGOW'15)**
2. $\boldsymbol{\mathrm{ds}(T) > \epsilon \Rightarrow cap(T)}$ grows by $\exp(\boldsymbol{O(\epsilon)})$ after normalization
3. $\mathbf{cap(T) \le 1}$ for normalized operators.

When can we scale?

Matrix scaling $\Leftrightarrow$ there was no Hall blocker. Analog in this case?

Definition [Gur'05]: $(A_1, \ldots, A_m)$ rank non-decreasing (**RND**) iff for all $V \subseteq \mathbb{C}^n$

$$\boldsymbol{dim}\left(\bigcup_i A_i V\right) \geq \boldsymbol{dim}(V)$$

Theorem [Gur'05]: $\mathrm{T} = (A_1, \ldots, A_m)$ then

$$\boldsymbol{cap}(T) > \mathbf{0} \Leftrightarrow (A_1, \ldots, A_m) \text{ RND}$$

Observation: $(A_1, \ldots, A_m)$ rank decreasing $\Leftrightarrow$ in some

basis they have a common Hall Blocker!

Lower Bound on Capacity

Reminder:

$$T(X) = \sum_i A_i X A_i^\dagger$$

$$cap(T) = inf\{\det(T(X)) : X \succ 0, \det(X) = 1\}.$$

Want to prove that:

$$cap(T) > 0 \Rightarrow cap(T) > e^{-poly(n)}$$

Basic case of **RND**: A_1 is an invertible matrix.

$$T(X) \succcurlyeq A_1 X A_1^\dagger \Rightarrow det(T(X)) \geq det(A_1 X A_1^\dagger)$$

$$det(X) = 1 \Rightarrow det(A_1 X A_1^\dagger) = det(A_1)^2 \geq 1$$

Lower Bound on Capacity

Next basic case: $\boldsymbol{A_1, \dots, A_m}$ span an invertible matrix.

Easy Lemma I: for any unitary matrix $\boldsymbol{B} \in \mathbb{C}^{\boldsymbol{m \times m}}$, let

$\boldsymbol{C_i = \sum_j b_{ij} A_j}$ and $\boldsymbol{T_B(X) = \sum_i C_i X C_i^\dagger}$. Then

$$\boldsymbol{T_B(X) = T(X).}$$

$\boldsymbol{A_1, \dots, A_m}$ span an invertible matrix, then $\exists$ unitary $\boldsymbol{B} \in \mathbb{C}^{\boldsymbol{m \times m}}$

with $b_{1j} \in \mathbb{Q}$ ($b_{1j} = p_j/q$, q small) s.t. $\boldsymbol{C_1 = \sum_j b_{1j} A_j}$ invertible.

$$\boldsymbol{T(X) = T_B(X) \succcurlyeq C_1 X C_1^\dagger \Rightarrow det\big(T(X)\big) \geq det\big(C_1 X C_1^\dagger\big)}$$

$$\boldsymbol{det(X) = 1 \Rightarrow det\big(C_1 X C_1^\dagger\big) = det(C_1)^2 \geq \frac{1}{q^{2n}}}$$

Lower Bound on Capacity

General case: $\boldsymbol{T}(\boldsymbol{X})$ rank non-decreasing

Definition: If $\boldsymbol{T_1}: \boldsymbol{M_{n_1}}(\mathbb{C}) \rightarrow \boldsymbol{M_{n_1}}(\mathbb{C}), \boldsymbol{T_2}: \boldsymbol{M_{n_2}}(\mathbb{C}) \rightarrow \boldsymbol{M_{n_2}}(\mathbb{C})$

given by $\boldsymbol{T_i}(\boldsymbol{X}) = \sum_j \boldsymbol{A_{ij}} \boldsymbol{X} \boldsymbol{A_{ij}^\dagger}$, define

$\boldsymbol{T_{12}} \stackrel{\text{def}}{=} \boldsymbol{T_1} \otimes \boldsymbol{T_2} : \boldsymbol{M_{n_1 n_2}}(\mathbb{C}) \rightarrow \boldsymbol{M_{n_1 n_2}}(\mathbb{C})$ as

$$\boldsymbol{T_{12}}(\boldsymbol{Y}) = \sum \boldsymbol{B_{ij}} \boldsymbol{X} \boldsymbol{B_{ij}^\dagger}$$

Where $\boldsymbol{B_{ij}} = \boldsymbol{A_{1i}} \otimes \boldsymbol{A_{2j}}$.

Easy Lemma II:

$$\boldsymbol{cap}(\boldsymbol{T_{12}}) \leq \boldsymbol{cap}(\boldsymbol{T_1})^{n_2} \boldsymbol{cap}(\boldsymbol{T_2})^{n_1}.$$

To get good lower bound on capacity, it is enough to find an operator $\boldsymbol{T'}: \boldsymbol{M_d}(\mathbb{C}) \rightarrow \boldsymbol{M_d}(\mathbb{C})$ with $\boldsymbol{d} = \boldsymbol{e^{poly(n)}}$ such that $\boldsymbol{T} \otimes \boldsymbol{T'}$ has an invertible matrix in their span.

Invariant Theory for analysis

Invariant Theory:

Group $G = \mathbb{SL}_n(\mathbb{C})^2$ acts on $(A_1, \dots, A_m)$ by L-R multiplication:

$$(A_1, \dots, A_m) \to (LA_1R, \dots, LA_mR)$$

Null-cone Problem: given $(A_1, \dots, A_m)$, is there sequence of scalings (L_t, R_t) such that

$$\lim_{t\to\infty} (L_tA_1R_t, \dots, L_tA_mR_t) = (0, \dots, 0)?$$

Invariant Theory [DW'00, DZ'01, SdB'01, ANS'10]:

$(A_1, \dots, A_m)$ in Null Cone $\Leftrightarrow$ $(A_1, \dots, A_m)$ RND

$$\Leftrightarrow det(\textstyle\sum_i A_i \otimes B_i) = 0 \;\forall\, B_i \in \mathcal{M}_d(\mathbb{C}), \forall\, d$$

[Derksen'01]: Enough to take $d \le 2^{n^2}$.

Pulling things together (in a nutshell)

$$\boldsymbol{cap(T) = 0} \Leftrightarrow (\boldsymbol{A_1, \dots, A_m}) \text{ RND}$$
$$\Leftrightarrow (\boldsymbol{A_1, \dots, A_m}) \text{ in Nullcone}$$
$$\Leftrightarrow \boldsymbol{det(\textstyle\sum_i A_i \otimes B_i) = 0} \ \forall\, \boldsymbol{B_i} \in \mathcal{M}_{\boldsymbol{d}}(\mathbb{C}), \boldsymbol{d \le 2^{n^2}}$$

Lemma 1: $\boldsymbol{T_1}$ given by $(\boldsymbol{A_1, \dots, A_m}), \boldsymbol{T_2}$ given by $(\boldsymbol{B_1, \dots, B_m})$ and $\boldsymbol{T}$ given by $(\boldsymbol{A_1 \otimes B_1, A_1 \otimes B_2, \dots, A_m \otimes B_m})$ then

$$\boldsymbol{cap(T) \le cap(T_1)^{d_2} cap(T_2)^{d_1}}$$

Lemma 2: $\mathbf{T}$ given by $(\boldsymbol{C_1, \dots, C_m})$ s.t. $(\boldsymbol{C_1, \dots, C_m})$ span invertible matrix then

$$\boldsymbol{cap(T) \ge 2^{-n \cdot polylog(n)}}$$

Theorem: $\mathbf{T}$ given by $(\boldsymbol{A_1, \dots, A_m})$ s.t. $\boldsymbol{cap(T) > 0}$ then

$$\boldsymbol{cap(T) \ge 2^{-n^2 \cdot polylog(n)}}$$

Approximating Capacity

Algorithm G can easily be modified to approximate Capacity within $(\mathbf{1}+\boldsymbol{\epsilon})$-multiplicative factor.

$$cap(T) = inf\left\{\frac{det(T(X))}{det(X)} : X \succ 0\right\}$$

- Keep track of scalings
- $\mathbf{ds(T)} \le \epsilon \Rightarrow \mathbf{1} \ge cap(T) \ge \left(\mathbf{1} - \sqrt{n\epsilon}\right)^{n}$
- $cap(T) = \prod(det.of\ scalings) \cdot cap(T_0)$

BL inequalities – [BL’76, Lieb’90]

- **BL Datum:**
  - Matrices $\boldsymbol{B_i} : \mathbb{R}^{\boldsymbol{n}} \rightarrow \mathbb{R}^{\boldsymbol{n_i}}$
  - Numbers $\boldsymbol{p_1, p_2, \ldots, p_m > 0}$

- **Functional Inequality:** for all integrable functions f_i : $\mathbb{R}^{n_i} \rightarrow \mathbb{R}_{\geq 0}$

$$\int_{x \in \mathbb{R}^n} \prod_{i=1}^{m} f_i(B_i(x))\, dx \leq C \cdot \prod_{i=1}^{m} \|f_i\|_{\frac{1}{p_i}}$$

For which constant $\boldsymbol{C}$ does this inequality hold, if at all?

I.e., how do we prove inequalities?

Example: Cauchy-Schwarz Inequality

- For all integrable functions $f_i \colon \mathbb{R}^n \to \mathbb{R}_{\geq 0}$,

$$\int f_1(x) f_2(x)\, dx \leq ||f_1||_2\ ||f_2||_2$$

$$||f||_2 = \left(\int f(x)^2\, dx \right)^{\frac{1}{2}}$$

Example: Hölder's Inequality

- If $\sum_{i=1}^{m} p_i = 1$.

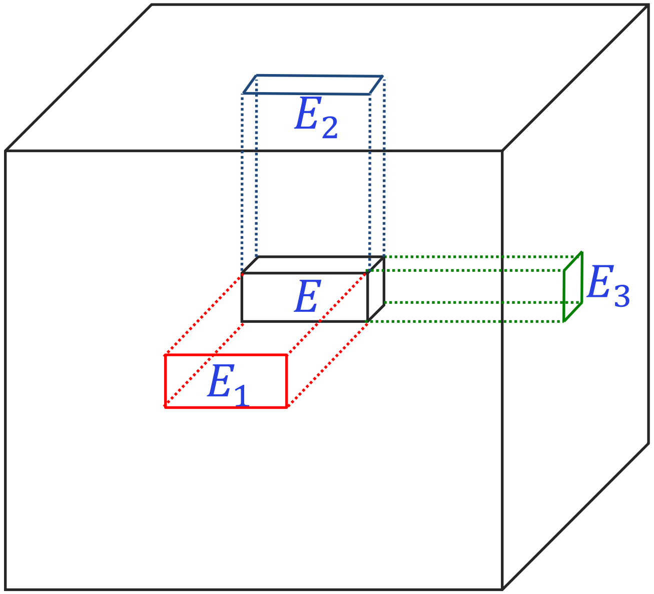
$$\int \prod_{i=1}^{m} f_i(x)\, dx \le \prod_{i=1}^{m} ||f_i||_{\frac{1}{p_i}}$$

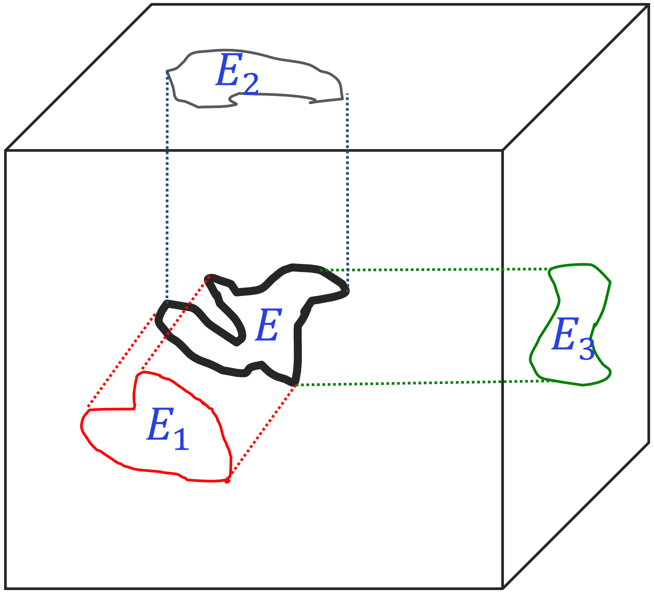
$$||f_i||_{\frac{1}{p_i}} = \left(\int f_i(x)^{\frac{1}{p_i}}\, dx \right)^{p_i}$$

Example: Loomis-Whitney Inequality

- Geometric inequality:
- Let $E \subset \mathbb{R}^3$ be a body.
- Let π_j denote the projection onto the coordinates $\{1,2,3\} \backslash \{j\}$ and $E_j = \pi_j(E)$.
- Then $\mathrm{Vol}(E) \leq \sqrt{\mathrm{Vol}(E_1) \cdot \mathrm{Vol}(E_2) \cdot \mathrm{Vol}(E_3)}$.

Example: Loomis-Whitney Inequality





$$\mathrm{Vol}(E) \leq \sqrt{\mathrm{Vol}(E_1) \cdot \mathrm{Vol}(E_2) \cdot \mathrm{Vol}(E_3)}$$

Example: Loomis-Whitney Inequality

- Functional inequality: Let π_j denote the projection onto the coordinates $\{1,2,3\}\backslash\{j\}$.

$$\int_{\mathbb{R}^3} \prod_{i=1}^{3} f_i(\pi_i(x))\ dx \ \leq \prod_{i=1}^{3} ||f_i||_2$$

$$||f||_2 = \left(\int_{\mathbb{R}^2} f(x)^2\, dx \right)^{\frac{1}{2}}$$

Example: Shearer's Lemma

- Let $S_1, \ldots, S_m \subseteq [n]$ s.t. each $i \in [n]$ appears in *exactly* k sets.

$$\int_{\mathbb{R}^n} \prod_{i=1}^{m} f_i(x_{S_i})\ dx \ \leq \prod_{i=1}^{m} ||f_i||_k$$

- Loomis-Whitney special case when $n = 3$ and $k = 2$.
- Equivalent entropy version [CCE'08, LCCV'16] and discrete analogue [CDKSY'15]

BL inequalities

- A BL datum $(\boldsymbol{B}, \boldsymbol{p})$ will be called feasible if the BL inequality holds with a finite constant.
- The optimal constant in the inequality will be called BL constant and denoted by $\mathrm{BL}(\boldsymbol{B}, \boldsymbol{p})$.

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

Lieb's Theorem [Lieb'90]

- Maximizers are Gaussians

$$\int_{\mathbb{R}^{n_i}} \exp(-\pi y^T A_i y)\ dy = \frac{1}{\sqrt{\det(A_i)}}.$$

- Hence

$$\mathrm{BL}(\boldsymbol{B},\boldsymbol{p})^2 = \sup_{\substack{A_1,\dots,A_m \\ A_i \succ 0\ n_i \times n_i}} \frac{\prod_{i=1}^m \det(A_i)^{p_i}}{\det\left(\sum_{i=1}^m p_i B_i^T A_i B_i\right)}$$

- Looks an awful lot like capacity

BL Polytope [BCCT'05]

- $\mathrm{BL}(\boldsymbol{B},\boldsymbol{p}) < \infty$ iff the following hold:

1. $\mathbf{n} = \sum_{\mathbf{i}=1}^{\mathbf{m}} \mathbf{p_i n_i}$.
2. For all subspaces $V \subseteq \mathbb{R}^n$,

$$\mathbf{dim}(\boldsymbol{V}) \leq \sum_{i=1}^{m} \boldsymbol{p_i}\, \boldsymbol{dim}(\boldsymbol{B_i}(\boldsymbol{V}))$$

- Fix $\boldsymbol{B}$. Let $P_{\boldsymbol{B}}$ set of $\boldsymbol{p}'s$ that satisfy above conditions.
- Finitely many constraints, so $P_{\boldsymbol{B}}$ is a polytope.

Geometric BL Datum [Ball'89, Barthe'98]

- $(\boldsymbol{B}, \boldsymbol{p})$ is called geometric if it satisfies the following normalization conditions:

1. Projection: $B_i B_i^T = I_{n_i}$ for all $\boldsymbol{i}$.
2. Isotropy: $\sum_{i=1}^{m} p_i B_i^T B_i = I_n$.

- If $(\boldsymbol{B}, \boldsymbol{p})$ geometric, then $\mathrm{BL}(\boldsymbol{B}, \boldsymbol{p}) = 1$.
- Can we convert (efficiently) any feasible BL datum to the geometric case?

Scaling Algorithm

- Fixing projection: $B_i \leftarrow \left(B_i B_i^T\right)^{-1/2} B_i$.
- Fixing isotropy: $B_i \leftarrow B_i \left(\sum_{i=1}^{m} p_i B_i^T B_i\right)^{-1/2}$.
- Can we fix both? Fixing one might disturb the other.
- Keep fixing both alternately for a few steps. This works!

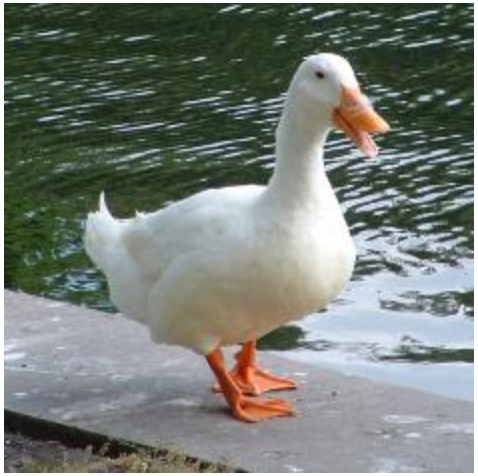
Repeat for $t = \mathrm{poly}(n, b, d, 1/\epsilon)$ steps:

1. Fix projection;
2. Fix isotropy;
3. Output *feasible* if get close to geometric position

Output not feasible

- How to analyze it?

Looks like a duck…



- Analysis by reduction to operator scaling!

Computing Approx. of BL const. [GGOW'16]

- Reduction to operator scaling
  - Matrices $\boldsymbol{B_i} : \mathbb{R}^{n} \rightarrow \mathbb{R}^{n_i}$
  - Numbers $\boldsymbol{p_i} = \frac{c_i}{d}, \quad \boldsymbol{c_i}, \boldsymbol{d} \in \mathbb{N}$

See board.

- Approx. BL const. reduces to approx. capacity!

Open Questions

- More applications of scaling problems?
- Can we obtain new inequalities that generalize capacity for non-abelian group actions?
- Van der Waerden for Operator scaling capacity? For general group actions?
- More BL-type inequalities for other quivers?

Advertisement

Amazing workshop at the IAS!
Videos & materials online
https://www.math.ias.edu/ocit2018

Survey on all of this on arxiv & on
EATCS complexity column!
(link on my webpage)

Landscape

- **[IQS'15]** Algebraic algorithm for operator scaling
- **[F'18]** Generalized operator scaling to arbitrary marginals
- **[GGOW'17]** Computing Brascamp-Lieb constant
- **[ALOW'17, AGLOW'18]** Faster algorithms for matrix & operator scaling
- **[BDWY'12, DGOS'18]** Generalizations of Sylvester-Gallai thms
- **[BGOWW'18]** Generalization to tensor scaling
- **[BFGOWW'18]** Tensor scaling for arbitrary marginals
  - More representation theory and invariant theory
- **[KLLR'18, HM'18]** Solution to Paulsen Problem