Polynomial Capacity and Gurvits' Theorem

Polynomial Capacity: Theory, Applications, Generalizations

Jonathan Leake

Technische Universität Berlin

December 10th, 2020

Notation

Polynomial notation:

- $\mathbb{R}, \mathbb{R}_+, \mathbb{Z}_+ :=$ reals, non-negative reals, non-negative integers.
- $\boldsymbol{x}^{\boldsymbol{\mu}} := \prod_i x_i^{\mu_i}$ and $\boldsymbol{\mu} \leq \boldsymbol{\lambda}$ is entrywise.
- $\mathbb{R}[\boldsymbol{x}] :=$ v.s. of real polynomials in n variables.
- $\mathbb{R}_+[\boldsymbol{x}] :=$ v.s. of real polynomials with non-negative coefficients.
- $\mathbb{R}^{\boldsymbol{\lambda}}[\boldsymbol{x}] :=$ v.s. of polynomials of degree at most λ_i in x_i.
- For $p \in \mathbb{R}[\boldsymbol{x}]$, we write $p(\boldsymbol{x}) = \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}}$.
- For d-homogeneous $p \in \mathbb{R}[\boldsymbol{x}]$, we write $p(\boldsymbol{x}) = \sum_{|\boldsymbol{\mu}|=d} p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}}$.
- $\frac{d}{dx} = \frac{\partial}{\partial x} = \partial_x :=$ derivative with respect to x, and $\partial_{\boldsymbol{x}}^{\boldsymbol{\mu}} := \prod_i \partial_{x_i}^{\mu_i}$.
- $\text{supp}(p) =$ **support** of $p =$ the set of $\boldsymbol{\mu} \in \mathbb{Z}_+^n$ for which $p_{\boldsymbol{\mu}} \neq 0$.
- $\text{Newt}(p) =$ **Newton polytope** of $p =$ convex hull of the support of p as a subset of $\mathbb{R}^n$.

Recall: The big three

The **geometry of polynomials** is generally an investigation of the connections between the various properties of polynomials:

- **Algebraic**, via the roots/zeros of the polynomial.
- **Combinatorial**, via the coefficients of the polynomial.
- **Analytic**, via the evaluations of the polynomial.

Why do we care? We use features of the interplay between these three to prove facts about mathematical objects which a priori have nothing to do with polynomials.

Typical method:

1. Encode some object as a polynomial which has some nice properties.
2. Apply operations to that polynomial which preserve those properties.
3. **Extract information** at the end which relates back to the object.

Outline

1. Motivation
   - The permanent of a matrix with non-negative entries
   - The Van der Waerden "conjecture"
   - Doubly stochastic polynomials
2. Polynomial capacity
   - Definition
   - Properties of capacity: Newton polytope
   - Properties of capacity: Marginal probabilities
3. Gurvits' theorem
   - Statement of the theorem
   - The Van der Waerden bound
   - Beyond doubly stochastic matrices/polynomials
   - "Proof" of Gurvits' theorem

Outline

1. Motivation
   - The permanent of a matrix with non-negative entries
   - The Van der Waerden “conjecture”
   - Doubly stochastic polynomials
2. Polynomial capacity
   - Definition
   - Properties of capacity: Newton polytope
   - Properties of capacity: Marginal probabilities
3. Gurvits’ theorem
   - Statement of the theorem
   - The Van der Waerden bound
   - Beyond doubly stochastic matrices/polynomials
   - “Proof” of Gurvits’ theorem

The permanent

Permanent of a matrix $A = (a_{ij})_{i,j=1}^n$:

$$\mathrm{per}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^n a_{i,\sigma(i)}.$$

Barvinok: "Like the determinant, only simpler."

Determinant: Compute exactly in poly(n) time.
Permanent: #P-hard to compute exactly, **even for 0-1 matrices**.

The 0-1 matrix case is equivalent to counting perfect matchings of a bipartite graph. (In a few slides.)

Searching for a perfect matching can be done in polynomial time, unlike **counting** which is #P-hard.

Polynomial coefficients

Connect to polynomials? Given A with $\mathbb{R}_+$ entries, define:

$$p_A(\boldsymbol{x}) := \prod_{i=1}^n \sum_{j=1}^n a_{ij} x_j = \prod_{i=1}^n (a_{i1}x_1 + \cdots + a_{in}x_n) \in \mathbb{R}_+[x_1, \ldots, x_n].$$

Properties:

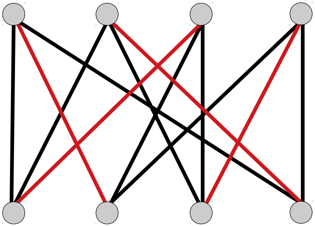
- Homogeneous polynomial of degree n in n variables.
- p_A is a product of non-negative linear forms $\implies$ p_A is real stable.
- Coefficient of $x_1 x_2 \cdots x_n$ is equal to $\mathrm{per}(A)$. **Why?** This coefficient corresponds to summing over all ways to choose distinct variables from each linear form.
- $\mathrm{per}(A) = \partial_{x_1}\partial_{x_2}\cdots\partial_{x_n} p_A = \partial_{x_1}|_{x_1=0}\, \partial_{x_2}|_{x_2=0} \cdots \partial_{x_n}|_{x_n=0}\, p_A$.

Last one: Gives us a way to induct on n-homogeneous polynomials in n variables, even though $\partial_{x_n}|_{x_n=0}\, p_A$ is **not** equal to p_B for some B.

Compare: Contraction of a matroid $\iff$ corresponding operation on A?

Bipartite perfect matchings

Let G be a bipartite graph on $2n$ vertices and consider:



Bipartite adjacency matrix, A:

$$\begin{bmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

perfect matchings = permanent

d-regular $\iff$ rows = cols = d

perfect matchings of G = permanent of $A = \partial_{x_1}|_{x_1=0} \cdots \partial_{x_n}|_{x_n=0}\, p_A$.

Existence of perfect matching of G $\iff$ Permanent of A is > 0.

Regular graphs and doubly stochastic matrices

Special case: d-regular graph $G \iff$ degree of every vertex is d.

Adjacency matrix of regular graph: All row/column sums equal to d.

Doubly stochastic (DS) matrix: Non-negative entries, and all row and column sums equal to 1.

Birkhoff polytope: Set of all doubly stochastic matrices. The extreme points are the permutation matrices.

$$\begin{aligned} \exists \text{ perfect matching of } G &\iff \operatorname{per}(A) > 0 \\ &\iff \max[A \cdot X] > n - 1 \text{ over Birkhoff polytope.} \end{aligned}$$

Two questions:

1. What can we say about the permanent/# perfect matchings for doubly stochastic A or d-regular G?
2. What can we say about the polynomial p_A for doubly stochastic A?

Van der Waerden "conjecture"

"Conjecture" [Falikman, Egorychev '81]: If A is doubly stochastic, then

$$\text{per}(A) \geq \frac{n!}{n^n} \approx \sqrt{2\pi n} \cdot e^{-n}.$$

Worst-case: $A = \frac{1}{n} J_n$, where J_n is the all-ones matrix $\implies$ $\text{per}(A) = \frac{n!}{n^n}$.

Corollary: $\#\text{pm}(G) \geq d^n \cdot \frac{n!}{n^n}$ for d-regular G.

Original proofs: Relied on Alexandrov-Fenchel inequalities (good sign for us), but pretty complicated. **Better/more illuminating proof?**

Algorithmic implications:

1. Also know that $\text{per}(A) \leq 1$ for any DS A (exercise).
2. Therefore, $\frac{n!}{n^n}$ is an e^n-approximation to $\text{per}(A)$ for **any** DS A.
3. Iterative procedure to "scale" a non-negative entry matrix to DS: e^n-approximation to $\text{per}(A)$ [Linial-Samorodnitsky-Wigderson '00].

Doubly stochastic polynomials

Gurvits: Let's get the stable polynomial p_A in the mix:

$$p_A(\boldsymbol{x}) := \prod_{i=1}^{n} (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) = \prod_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_j.$$

Properties of p_A for doubly stochastic A:

- Row sums $= 1 \implies p_A(\mathbf{1}) = 1$.
- Row/column sums $= 1 \implies \nabla p_A(\mathbf{1}) = \mathbf{1}$. **Why?** Product rule:

$$\partial_{x_1} p_A(\mathbf{1}) = \sum_{k=1}^{n} a_{k1} \prod_{i \neq k} (a_{i1} \cdot 1 + a_{i2} \cdot 1 + \cdots + a_{in} \cdot 1) = \sum_{k=1}^{n} a_{k1} = 1.$$

- **Already know:** Coefficient of $x_1 x_2 \cdots x_n$ equals $\mathrm{per}(A)$.

Doubly stochastic polynomial: $p(\mathbf{1}) = 1$ and $\nabla p(\mathbf{1}) = \mathbf{1}$.

New conjecture: All-ones coefficient of **stable** DS polynomial $\geq \frac{n!}{n^n}$.

Polynomial is **easy** to evaluate, but it's **#P-hard** to compute coefficients.

Outline

1. Motivation
   - The permanent of a matrix with non-negative entries
   - The Van der Waerden “conjecture”
   - Doubly stochastic polynomials
2. Polynomial capacity
   - Definition
   - Properties of capacity: Newton polytope
   - Properties of capacity: Marginal probabilities
3. Gurvits’ theorem
   - Statement of the theorem
   - The Van der Waerden bound
   - Beyond doubly stochastic matrices/polynomials
   - “Proof” of Gurvits’ theorem

Polynomial capacity

Definition: Given $p \in \mathbb{R}_+[\boldsymbol{x}]$ and $\boldsymbol{\alpha} \in \mathbb{R}_+^n$ (always from now on), define:

$$\mathrm{Cap}_{\boldsymbol{\alpha}}(p) := \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\alpha}}} = \inf_{x_1,\ldots,x_n>0} \frac{p(x_1,\ldots,x_n)}{x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}}.$$

First: For $\boldsymbol{\mu} \in \mathrm{supp}(p)$, we have

$$\mathrm{Cap}_{\boldsymbol{\mu}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\mu}}} \geq \inf_{\boldsymbol{x}>0} \frac{p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}}}{\boldsymbol{x}^{\boldsymbol{\mu}}} = p_{\boldsymbol{\mu}}.$$

Maybe: $\mathrm{Cap}_{\mathbf{1}}(p)$ is a good approximation to $p_{\mathbf{1}}$ (all-ones coefficient).

So what? $\mathrm{Cap}_{\mathbf{1}}(p)$ looks hard to compute in any case. **Actually:**

$$\log \mathrm{Cap}_{\boldsymbol{\alpha}}(p) = \inf_{\boldsymbol{y}\in\mathbb{R}^n} \left[-\langle \boldsymbol{y}, \boldsymbol{\alpha} \rangle + \log \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} e^{\langle \boldsymbol{y}, \boldsymbol{\mu} \rangle} \right] \quad \text{via} \quad \boldsymbol{x} \to e^{\boldsymbol{y}}.$$

Since $p_{\boldsymbol{\mu}} > 0$, this is a **convex program**. (Or: **geometric programming**.)

Compute via: Ellipsoid method, interior point method, etc.

Properties of polynomial capacity: Newton polytope

Last slide: "Something like" the coefficients of p; efficiently computable.

Crucial connection: Polynomial capacity and AM-GM inequality,

$$\omega_1 x_1 + \omega_2 x_n + \cdots + \omega_n x_n \geq x_1^{\omega_1} x_2^{\omega_2} \cdots x_n^{\omega_n},$$

where $\boldsymbol{x}, \boldsymbol{\omega} \in \mathbb{R}_+^n$ and $\sum_i \omega_i = 1$.

Fact: $0 < \mathrm{Cap}_{\boldsymbol{\alpha}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\alpha}}}$ if and only if $\boldsymbol{\alpha} \in \mathrm{Newt}(p)$.
Proof of ($\Longleftarrow$): Since $\boldsymbol{\alpha} \in \mathrm{Newt}(p)$, there is a convex combination:

$$\boldsymbol{\alpha} = \sum_{\boldsymbol{\mu} \in S} c_{\boldsymbol{\mu}} \boldsymbol{\mu} \quad \text{with} \quad S \subset \mathrm{supp}(p) \quad \text{and} \quad c_{\boldsymbol{\mu}} > 0 \quad \text{and} \quad \sum_{\boldsymbol{\mu} \in S} c_{\boldsymbol{\mu}} = 1.$$

For any $\boldsymbol{x} > 0$, we now apply AM-GM:

$$p(\boldsymbol{x}) \geq \sum_{\boldsymbol{\mu} \in S} p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}} = \sum_{\boldsymbol{\mu} \in S} c_{\boldsymbol{\mu}} \left(\frac{p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}}}{c_{\boldsymbol{\mu}}} \right) \geq \prod_{\boldsymbol{\mu} \in S} \left(\frac{p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}}}{c_{\boldsymbol{\mu}}} \right)^{c_{\boldsymbol{\mu}}} \geq \left(\min_{\boldsymbol{\mu} \in S} \frac{p_{\boldsymbol{\mu}}}{c_{\boldsymbol{\mu}}} \right) \boldsymbol{x}^{\boldsymbol{\alpha}}.$$

Properties of polynomial capacity: Newton polytope

Fact: $0 < \text{Cap}_{\boldsymbol{\alpha}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\alpha}}}$ if and only if $\boldsymbol{\alpha} \in \text{Newt}(p)$.

Proof of ($\Longrightarrow$): We prove the contrapositive. Let's rewrite capacity as

$$\text{Cap}_{\boldsymbol{\alpha}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\alpha}}} = \inf_{\boldsymbol{y}\in\mathbb{R}^n} \sum_{\boldsymbol{\mu}\in\text{supp}(p)} p_{\boldsymbol{\mu}} e^{\langle \boldsymbol{y}, \boldsymbol{\mu}-\boldsymbol{\alpha}\rangle}.$$

$\boldsymbol{\alpha} \notin \text{Newt}(p) \iff \boldsymbol{\mu} - \boldsymbol{\alpha}$ in the same open half-space for all $\boldsymbol{\mu} \iff$ existence of $\boldsymbol{y}$ such that $\langle \boldsymbol{y}, \boldsymbol{\mu} - \boldsymbol{\alpha}\rangle < 0$ for all $\boldsymbol{\mu} \in \text{supp}(p)$.

So: Scale this $\boldsymbol{y}$ larger and larger to limit to 0.

Key idea: Separating hyperplane for Newton polytope.

Conceptual conclusion: Capacity "picks out" a polytope, and is related to algorithmic/optimization questions concerning the polytope.

Properties of polynomial capacity: Marginals

Fix $p \in \mathbb{R}_+[\boldsymbol{x}]$ such that $p(\mathbf{1}) = 1$. Define a probability distribution $\boldsymbol{\nu}_p$:

$$\mathbb{P}[\boldsymbol{w} = \boldsymbol{\mu}] = p_{\boldsymbol{\mu}} \quad \text{where} \quad \boldsymbol{w} \sim \boldsymbol{\nu}_p.$$

That is: $\boldsymbol{\nu}_p$ is distributed on $\mathbb{Z}^n$ according to the coefficients/support of p.

Marginal probabilities ("marginals") are coordinates of the expectation:

$$\mathbb{E}[\boldsymbol{w}] = \sum_{\boldsymbol{\mu} \in \text{supp}(p)} p_{\boldsymbol{\mu}} \cdot \boldsymbol{\mu} \quad \text{where} \quad \boldsymbol{w} \sim \boldsymbol{\nu}_p.$$

In fact, the i^{th} marginal of $\boldsymbol{\nu}_p$ can be written as

$$\mathbb{E}[w_i] = \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} \mu_i = \partial_{x_i} \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} \boldsymbol{x}^{\boldsymbol{\mu}} \bigg|_{\boldsymbol{x}=\mathbf{1}} = \partial_{x_i} p(\mathbf{1}).$$

That is: Marginals of $\boldsymbol{\nu}_p = \nabla p(\mathbf{1})$.

Fact: A polynomial p is doubly stochastic iff the marginals of $\boldsymbol{\nu}_p$ are $\mathbf{1}$.

Properties of polynomial capacity: Marginals

Easy: If $p(\mathbf{1}) = 1$, then $0 \leq \mathrm{Cap}_{\boldsymbol{\alpha}}(p) \leq 1$ for all $\boldsymbol{\alpha}$.

Proof: $\mathrm{Cap}_{\boldsymbol{\alpha}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\alpha}}} \leq \frac{p(\mathbf{1})}{\mathbf{1}^{\boldsymbol{\alpha}}} = 1.$

Recall: $0 < \mathrm{Cap}_{\boldsymbol{\alpha}}(p)$ iff $\boldsymbol{\alpha} \in \mathrm{Newt}(p)$.

Fact: If $p(\mathbf{1}) = 1$, then $\mathrm{Cap}_{\boldsymbol{\alpha}}(p) = 1$ iff $\boldsymbol{\alpha}$ are the marginals of p (of $\boldsymbol{\nu}_p$).

Proof: Recall that we have

$$0 \geq \log \mathrm{Cap}_{\boldsymbol{\alpha}}(p) = \inf_{\boldsymbol{y}\in\mathbb{R}^n} \left[-\langle \boldsymbol{y}, \boldsymbol{\alpha} \rangle + \log \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} e^{\langle \boldsymbol{y}, \boldsymbol{\mu} \rangle} \right] =: \inf_{y\in\mathbb{R}^n} F(\boldsymbol{y}).$$

Now, compute the gradient of the **convex** objective at $\boldsymbol{y} = \mathbf{0}$:

$$\partial_{y_i} F(\boldsymbol{y})\big|_{\boldsymbol{y}=\mathbf{0}} = -\alpha_i + \left.\frac{\partial_{y_i} \sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} e^{\langle \boldsymbol{y}, \boldsymbol{\mu} \rangle}}{\sum_{\boldsymbol{\mu}} p_{\boldsymbol{\mu}} e^{\langle \boldsymbol{y}, \boldsymbol{\mu} \rangle}}\right|_{\boldsymbol{y}=\mathbf{0}} = -\alpha_i + \frac{\partial_{x_i} p(\mathbf{1})}{p(\mathbf{1})}.$$

So marginals $= \boldsymbol{\alpha}$ iff $\nabla F(\boldsymbol{y})\big|_{\boldsymbol{y}=\mathbf{0}} = \mathbf{0}$ iff infimum attained at $F(\mathbf{0}) = 0$.

Recap of capacity properties

For $p \in \mathbb{R}_+[\boldsymbol{x}] = \mathbb{R}_+[x_1, \ldots .x_n]$ and $\boldsymbol{\alpha} \in \mathbb{R}_+^n$, we write:

$$\mathrm{Cap}_{\boldsymbol{\alpha}}(p) := \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{\alpha}}}.$$

Properties for $p(\mathbf{1}) = 1$:

- $0 \leq \mathrm{Cap}_{\boldsymbol{\alpha}}(p) \leq 1$.
- $0 < \mathrm{Cap}_{\boldsymbol{\alpha}}(p)$ iff $\boldsymbol{\alpha} \in \mathrm{Newt}(p)$.
- $\mathrm{Cap}_{\boldsymbol{\alpha}}(p) = 1$ iff $\boldsymbol{\alpha} = \nabla p(\mathbf{1})$.
- p is doubly stochastic iff $\mathrm{Cap}_{\mathbf{1}}(p) = 1$.
- If row sums are 1, then A is doubly stochastic iff $\mathrm{Cap}_{\mathbf{1}}(p_A) = 1$.
- For $\boldsymbol{\mu} \in \mathrm{supp}(p)$, we have $p_{\boldsymbol{\mu}} \leq \mathrm{Cap}_{\boldsymbol{\mu}}(p)$.

Can we use these properties (especially the last three) to prove the Van der Waerden lower bound for doubly stochastic matrices/polynomials?

Outline

Gurvits' theorem

Theorem (Gurvits '05)

If $p \in \mathbb{R}_+[\boldsymbol{x}]$ is an n-homogeneous n-variate real stable polynomial, then:

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_n} p\big|_{x_n=0}\right) \geq \left(\frac{n-1}{n}\right)^{n-1} \mathrm{Cap}_{\mathbf{1}}(p).$$

Here the length of $\mathbf{1}$ *depends on the input polynomial, and* $0^0 = 1$.

Note: Length of $\mathbf{1}$ depends on the number of variables. **So:**

$$\mathrm{Cap}_{\mathbf{1}}(p) = \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{x_1 x_2 \cdots x_n} \quad \text{but} \quad \mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_n} p\big|_{x_n=0}\right) = \inf_{\boldsymbol{x}>0} \frac{\partial_{x_n} p(\boldsymbol{x})}{x_1 x_2 \cdots x_{n-1}}.$$

Further: What about when $n = 1$?

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_1} p\big|_{x_1=0}\right) = \inf_{\boldsymbol{x}>0} \frac{p_1}{\prod_{i=1}^{0} x_i} = p_1 = \text{const}.$$

Gurvits' theorem and the Van der Waerden bound

Theorem: For real stable n-homogeneous $p \in \mathbb{R}_+[\boldsymbol{x}]$, we have

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_n} p\big|_{x_n=0}\right) \geq \left(\frac{n-1}{n}\right)^{n-1} \mathrm{Cap}_{\mathbf{1}}(p).$$

Corollary: $p_{\mathbf{1}} \geq \frac{n!}{n^n}$ (VdW bound) holds for doubly stochastic real stable p.
Proof: Since $\partial_{x_n} p\big|_{x_n=0}$ is $(n-1)$-homogeneous $(n-1)$-variate real stable, we can apply the theorem inductively. **E.g.:**

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_{n-1}}\big|_{x_{n-1}=0} \partial_{x_n}\big|_{x_n=0} p\right) \geq \left(\frac{n-2}{n-1}\right)^{n-2} \mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_n} p\big|_{x_n=0}\right).$$

This leads to:

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_1}\big|_{x_1=0} \cdots \partial_{x_n}\big|_{x_n=0} p\right) \geq \left(\frac{1}{2}\right)^{1} \left(\frac{2}{3}\right)^{2} \cdots \left(\frac{n-1}{n}\right)^{n-1} \mathrm{Cap}_{\mathbf{1}}(p).$$

Gurvits' theorem and the Van der Waerden bound

Last slide: p doubly stochastic real stable, and

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_1}\big|_{x_1=0}\cdots\partial_{x_n}\big|_{x_n=0}\,p\right) \geq \left(\frac{1}{2}\right)^1\left(\frac{2}{3}\right)^2\cdots\left(\frac{n-1}{n}\right)^{n-1}\mathrm{Cap}_{\mathbf{1}}(p).$$

First: Since p is doubly stochastic, we have $\mathrm{Cap}_{\mathbf{1}}(p) = 1$.

Next, let's simplify the constant:

$$\left(\frac{1}{2}\right)^1\left(\frac{2}{3}\right)^2\cdots\left(\frac{n-1}{n}\right)^{n-1} = \frac{1\cdot 2\cdot 3\cdots(n-1)}{n^{n-1}} = \frac{n!}{n^n}.$$

Finally, what is $\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_1}\big|_{x_1=0}\cdots\partial_{x_n}\big|_{x_n=0}\,p\right)$?

Recall: Length of $\mathbf{1}$ determined by input, a **constant polynomial**. So:

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_1}\big|_{x_1=0}\cdots\partial_{x_n}\big|_{x_n=0}\,p\right) = \inf_{\boldsymbol{x}>0}\frac{p_{\mathbf{1}}}{\prod_{i=1}^0 x_i^1} = p_{\mathbf{1}}.$$

Therefore: $p_{\mathbf{1}} \geq \frac{n!}{n^n}\,\mathrm{Cap}_{\mathbf{1}}(p) = \frac{n!}{n^n} \implies \mathrm{per}(A) \geq \frac{n!}{n^n}$ for DS A.

Gurvits' theorem and non-DS matrices

Last slides: Simple proof of the VdW bound with Gurvits' theorem.

Interesting: Proof utilized optimization problem $\text{Cap}_{\mathbf{1}}(p)$, but no computation was actually needed.

Next: Remove DS condition. Proof still goes through, but $\text{Cap}_{\mathbf{1}}(p) \neq 1$.

Corollary: If $p \in \mathbb{R}_+[x_1, \ldots, x_n]$ is n-homogeneous real stable, then

$$p_{\mathbf{1}} \geq \frac{n!}{n^n} \text{Cap}_{\mathbf{1}}(p) \geq e^{-n} \cdot \text{Cap}_{\mathbf{1}}(p).$$

Recall: $\text{Cap}_{\mathbf{1}}(p) \geq p_{\mathbf{1}}$ in general.

Corollary: "Convex" program $\text{Cap}_{\mathbf{1}}(p)$ gives an e^n-approximation to $p_{\mathbf{1}}$, given an evaluation oracle for p.

Corollary: Convex program for e^n-approximating the permanent of a matrix with $\mathbb{R}_+$ entries.

Proof of Gurvits' theorem

Theorem (Gurvits '05)

If $p \in \mathbb{R}_+[\boldsymbol{x}]$ is an n-homogeneous n-variate real stable polynomial, then:

$$\mathrm{Cap}_{\mathbf{1}}\left(\partial_{x_n} p\big|_{x_n=0}\right) \geq \left(\frac{n-1}{n}\right)^{n-1} \mathrm{Cap}_{\mathbf{1}}(p).$$

Here the length of $\mathbf{1}$ depends on the input polynomial, and $0^0 = 1$.

Proof idea: Want to show

$$\inf_{\boldsymbol{x}>0} \frac{\partial_{x_n} p(x_1,\ldots,x_{n-1},0)}{x_1 x_2 \cdots x_{n-1}} \geq \left(\frac{n-1}{n}\right)^{n-1} \cdot \inf_{\boldsymbol{x}>0} \frac{p(x_1,\ldots,x_{n-1},x_n)}{x_1 x_2 \cdots x_{n-1} x_n}.$$

Enough: For all $x_1,\ldots,x_{n-1} > 0$ and $q(t) := p(x_1,\ldots,x_{n-1},t)$, we have

$$q_1 = \partial_t q(0) \geq \left(\frac{n-1}{n}\right)^{n-1} \cdot \inf_{t>0} \frac{q(t)}{t} = \left(\frac{n-1}{n}\right)^{n-1} \mathrm{Cap}_1(q).$$

Boils down to a univariate coefficient bound for real-rooted $q \in \mathbb{R}_+^n[t]$.

Foreshadowing: The univariate coefficient bound

Recall we want: $q_1 \geq \left(\frac{n-1}{n}\right)^{n-1} \mathrm{Cap}_1(q)$.

Lemma (Brändén-L-Pak '20)

Let $q, w \in \mathbb{R}_+^n[t]$ have all positive coefficients such that $\left(\frac{q_k}{w_k}\right)_{k=0}^n$ forms a log-concave sequence. Then for all $k \in \{0, \ldots, n\}$, we have

$$\frac{q_k}{\mathrm{Cap}_k(q)} \geq \frac{w_k}{\mathrm{Cap}_k(w)}.$$

We will prove this next time. **How can we use it?** Setting $w_k = \binom{n}{k}$ translates to q_k being ultra log-concave. For $k = 1$, we obtain

$$q_1 \geq \frac{n}{\mathrm{Cap}_1(w)} \cdot \mathrm{Cap}_1(q) \quad \text{and} \quad \mathrm{Cap}_1(w) = \inf_{t>0} \frac{(t+1)^n}{t} = n\left(\frac{n}{n-1}\right)^{n-1}$$

where the last equality is a basic calculus exercise.