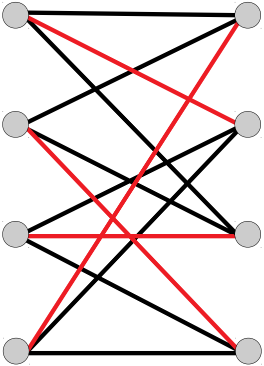
Approximate Counting via Lorentzian Polynomials and Entropy Optimization



$$\prod_{i=1}^{4} \sum_{j=1}^{4} \delta_{i \to j} x_j =$$

$$(x_1 + x_2 + x_3) \cdot (x_1 + x_3 + x_4) \cdot (x_2 + x_3 + x_4) \cdot (x_1 + x_2 + x_4)$$

Jonathan Leake

April 14th, 2022

The Van der Waerden conjecture

Question: Given matrix M with **non-negative entries** and specified **row and column sums** $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$, how small can $\text{per}(M)$ be?

	β_1	β_2	$\cdots$	β_n
α_1	m_{11}	m_{12}	$\cdots$	m_{1n}
α_2	m_{21}	m_{22}	$\cdots$	m_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_n	m_{n1}	m_{n2}	$\cdots$	m_{nn}

$$\Longrightarrow \qquad \text{per}(M) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} m_{i,\sigma(i)} \geq ?$$

Conjecture [vdW '26]**:** When $\boldsymbol{\alpha} = \boldsymbol{\beta} = \mathbf{1}$, we have $\text{per}(M) \geq \frac{n!}{n^n}$.

Theorem: Conjecture is true. [Egorychev, Falikman '80]

Theorem: Conjecture is true, basically via calculus. [Gurvits '04]

Bonus: If $\boldsymbol{\alpha} = \mathbf{1}$, some bound possible if and only if $\|\mathbf{1} - \boldsymbol{\beta}\|_1 < 2$. [Gurvits-L '21]

Question: How is this possible after 80 years?

Proof sketch of Gurvits' method

1. Convert the matrix M into a **polynomial** p:

	1	1	$\cdots$	1
1	m_{11}	m_{12}	$\cdots$	m_{1n}
1	m_{21}	m_{22}	$\cdots$	m_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
1	m_{n1}	m_{n2}	$\cdots$	m_{nn}

$$\implies \quad p(\boldsymbol{x}) := \prod_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_j.$$

2. Permanent of M is a **coefficient** of p: (∂_x = partial derivative)

$$\mathrm{per}(M) = \langle x_1 x_2 \cdots x_n \rangle p = \partial_{x_1}\big|_{x_1=0} \, \partial_{x_2}\big|_{x_2=0} \cdots \partial_{x_n}\big|_{x_n=0} \, p.$$

3. Use **induction** to bound $\partial_{x_1}\big|_{x_1=0} \, \partial_{x_2}\big|_{x_2=0} \cdots \partial_{x_n}\big|_{x_n=0} \, p$.

Problem: Derivatives don't preserve "matrix-ness" of p.

Dealing with derivatives

$$p(\boldsymbol{x}) := \prod_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_j \quad \Longrightarrow \quad \text{per}(M) = \partial_{x_1}\big|_{x_1=0} \cdots \partial_{x_n}\big|_{x_n=0}\, p$$

Problem: Derivatives don't preserve "matrix-ness" of p.

Solution: Derivatives do preserve **real stability**. How?

- **Definition:** $p \in \mathbb{R}[z_1, \ldots, z_n]$ and $p(\boldsymbol{z}) \neq 0$ when $\mathrm{Im}(z_i) > 0, \forall i$.
- **Gauss-Lucas:** $\{\text{roots of } \frac{df}{dt}\} \subset \text{hull}\{\text{roots of } f\}$ for $f \in \mathbb{C}[t]$.
- Apply Gauss-Lucas to $f(t) = p(t + z_1, z_2, \ldots, z_n)$ for $\mathrm{Im}(z_i) > 0, \forall i$: $\partial_{x_1} p(z_1, \ldots, z_n) = \partial_t\big|_{t=0}\, p(t + z_1, z_2, \ldots, z_n) = f'(0) \neq 0$.

Lorentzian is a generalization of real stable (more later).

Next problem: Proof by "induction"...on what?

- Coefficient of $x_1 x_2 \cdots x_n$ preserved by derivatives.
- How can we actually obtain a bound $(\geq \frac{n!}{n^n})$ from this?

Gurvits' main idea

Idea: Keep track of some "coefficient-like" quantity.

Polynomial capacity: For $p \in \mathbb{R}_{\geq 0}[\boldsymbol{x}]$,

$$\mathrm{Cap}_{\mathbf{1}}(p) := \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\mathbf{1}}} = \inf_{x_1,\ldots,x_n>0} \frac{\sum_{\boldsymbol{\kappa}} c_{\boldsymbol{\kappa}} x_1^{\kappa_1} \cdots x_n^{\kappa_n}}{x_1 \cdots x_n}$$

Theorem [Gurvits '04]**:** For n-homog. **real stable** $p \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$,

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

Proof: Reduce to univariate $\implies$ bound on **linear coefficient** via calculus

---

For $q_k := \partial_{x_{k+1}}\big|_{x_{k+1}=0} \cdots \partial_{x_n}\big|_{x_n=0} p$, we have

$$\mathrm{Cap}_{\mathbf{1}}(q_{k-1}) \geq \left(\frac{k-1}{k}\right)^{k-1} \mathrm{Cap}_{\mathbf{1}}(q_k)$$

By induction: $\mathrm{Cap}_{\mathbf{1}}(q_0) \geq \prod_{k=1}^{n} \left(\frac{k-1}{k}\right)^{k-1} \mathrm{Cap}_{\mathbf{1}}(q_n) = \frac{n!}{n^n} \mathrm{Cap}_{\mathbf{1}}(p)$

Proof of the van der Waerden bound

	1	1	$\cdots$	1
1	m_{11}	m_{12}	$\cdots$	m_{1n}
1	m_{21}	m_{22}	$\cdots$	m_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
1	m_{n1}	m_{n2}	$\cdots$	m_{nn}

$$\implies \quad p(\boldsymbol{x}) := \prod_{i=1}^{n} \sum_{j=1}^{n} m_{ij} x_j$$

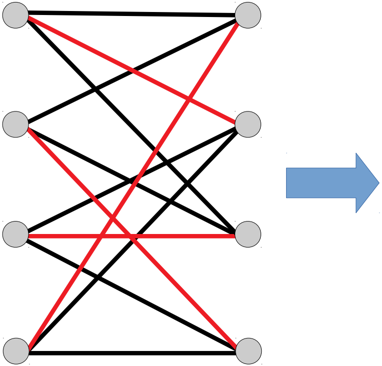
1. Apply **Gurvits' theorem** to real stable p:

$$\mathrm{Cap}_{\boldsymbol{1}} \left(\partial_{x_1}|_{x_1=0} \cdots \partial_{x_n}|_{x_n=0}\, p \right) \geq \frac{n!}{n^n} \mathrm{Cap}_{\boldsymbol{1}}(p) \quad \left[= \frac{n!}{n^n} \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\boldsymbol{1}}} \right]$$

2. **Degree 0:** $\mathrm{Cap}_{\boldsymbol{1}} \left(\partial_{x_1}|_{x_1=0} \cdots \partial_{x_n}|_{x_n=0}\, p \right) = c_{\boldsymbol{1}} = \mathrm{per}(M)$
3. Row/column sums $\implies$ $p(\boldsymbol{1}) = 1$ and $\nabla p(\boldsymbol{1}) = \boldsymbol{1}$ $\implies$ $\mathrm{Cap}_{\boldsymbol{1}}(p) = 1$

Question: What actually is polynomial capacity?

Approximate Counting via Lorentzian Polynomials and Entropy Optimization



$$\prod_{i=1}^{4}\sum_{j=1}^{4}\delta_{i\to j}x_j =$$

$$(x_1 + \color{red}{x_2} + x_3)\cdot$$
$$(x_1 + x_3 + \color{red}{x_4})\cdot$$
$$(x_2 + \color{red}{x_3} + x_4)\cdot$$
$$(\color{red}{x_1} + x_2 + x_4)$$

Jonathan Leake

April 14th, 2022

At the heart of Lorentzian polynomials

Question: Given real symmetric matrix A with $B(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{x}^\top A \boldsymbol{y}$,

- Eigenvalues of A all positive $\implies B(\boldsymbol{x}, \boldsymbol{y})^2 \leq B(\boldsymbol{x}, \boldsymbol{x}) B(\boldsymbol{y}, \boldsymbol{y})$ (**Cauchy-Schwarz inequality**)
- Eigenvalues of A $\underline{\quad ??? \quad}$ $\implies B(\boldsymbol{x}, \boldsymbol{y})^2 \geq B(\boldsymbol{x}, \boldsymbol{x}) B(\boldsymbol{y}, \boldsymbol{y})$ (**Alexandrov-Fenchel inequality**)

Answer: Exactly one positive eigenvalue, or **Lorentzian signature**

So what?

- The AF inequality is a **log-concavity** type of inequality:

$$B(\boldsymbol{x}, \boldsymbol{y})^2 \geq B(\boldsymbol{x}, \boldsymbol{x}) B(\boldsymbol{y}, \boldsymbol{y}) \quad \sim \quad c_k^2 \geq c_{k-1} c_{k+1}$$

- At the heart of **real stable**, **hyperbolic**, **Lorentzian** polynomials

---

Question: What are these **log-concave polynomials** like?

Lorentzian, real stable, and hyperbolic polynomials

Lorentzian polynomial: Homogeneous $p(\boldsymbol{x}) \in \mathbb{R}_{\geq 0}[x_1, \ldots, x_n]$ s.t. Hessians of all $\mathbb{R}^n_{>0}$-derivatives have **Lorentzian signature**

- Bivariate Lorentzian $\sum_{k=0}^{d} \binom{d}{k} c_k t^k s^{d-k} \iff c_k^2 \geq c_{k-1} c_{k+1}$
- Preserved by **linear restrictions** $p \mapsto p(t\boldsymbol{x} + s\boldsymbol{y})$ for $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n_{\geq 0}$

Examples of Lorentzian polynomials:

- **Real stable** polynomials ($p(\boldsymbol{x}) \neq 0$ whenever $\mathrm{Im}(\boldsymbol{x}) \in \mathbb{R}^n_{>0}$): spanning tree generating polynomials, $\det(\sum_i A_i x_i)$ for PSD A_i, etc.
- **Convex geometry:** $\mathrm{vol}(\sum_i K_i x_i)$ for compact convex K_i
- **Matroids:** Basis and independent set generating polynomials
- **Denormalized Lorentzian:** Schur and Schubert (conjectured) polynomials, contingency tables generating polynomials
- **Other cones:** Consider C-directional derivatives for convex cone C (see [Brändén-L '21], generalizes **hyperbolic polynomials** in cone C)

Question: What can we do with **log-concave polynomials**?

Some applications of log-concave polynomials

Inequalities, bounds, and approximation:

- Kadison-Singer conjecture [Marcus-Spielman-Srivastava '13]
- Optimization and counting on matroids [Straszak-Vishnoi '16], [Anari-Oveis Gharan '17], [Anari-Liu-Oveis Gharan-Vinzant '18]
- Permanent, matchings, mixed discriminant, mixed volume [Gurvits '00s]
- Monotone column permanent conjecture [Brändén-Haglund-Visontai-Wagner '10]
- Integer points of polytopes [Barvinok '00s], [Barvinok-Hartigan '09], [Gurvits '15], [Gurvits-L '18], [Csikvári-Schweitzer '20], [Brändén-L-Pak '21]
- Metric TSP approximation [Karlin-Klein-Oveis Gharan '21]

Log-concave (integer) sequences:

- Graph polynomials (matching, independence, etc.) [Heilmann-Lieb '72], [Choe-Oxley-Sokal-Wagner '02], [Chudnovsky-Seymour '07], [Borcea-Brändén '09]
- Schur, Schubert (conjectures), and Tutte polynomials [Sokal '05], [Huh-Matherne-Mészáros-St. Dizier '19], [Berget-Eur-Spink-Tseng '21]
- Mason's conjectures (size-k independent sets of a matroid) [Wagner '06], [Anari-Liu-Oveis Gharan-Vinzant '18], [Brändén-Huh '18]
- Heron-Rota-Welsh conjecture (characteristic polynomial of a matroid) [Adiprasito-Huh-Katz '15], [Backman-Eur-Simpson '18], [Brändén-L '21]

Some applications of log-concave polynomials

Combinatorial and algebraic structures:

- Matroids, delta matroids, jump systems, etc. [Choe-Oxley-Sokal-Wagner '02], [Brändén '07], [Anari-Liu-Oveis Gharan-Vinzant '18], [Brändén-Huh '18]
- Totally non-negative Grassmannian [Purbhoo '18]

Convex optimization:

- Hyperbolic programming [Gårding '59], [Renegar '06], [Renegar-Sondjaja '14]
- Interior-point methods [Güler '97], [Myklebust-Tunçel '14], [Nesterov-Tunçel '16]
- Lax conjecture (hyperbolic = SDP) [Helton-Vinnikov '07], [Scheiderer '16]

Statistical physics:

- General partition function approximation [Barvinok '15], [Patel-Regts '17]
- Hard-core lattice gas model (multivariate independence polynomial) [Patel-Regts '17], [Harvey-Srivastava-Vondrák '17, Bencs-Csikvári '18]
- Ising model [Lee-Yang '52], [Borcea-Brändén '09], [Liu-Sinclair-Srivastava '18]
- Also related to CLT talk from two weeks ago [Michelen-Sahasrabudhe '19], [Pemantle '17], [Borcea-Brändén-Liggett '07]

This talk: Lower bounds and approximate counting

Question: How can we **use log-concave polynomials to count**?

One thought: Consider generating functions $\sum_{\boldsymbol{\kappa} \in \mathbb{Z}^n_{\geq 0}} c_{\boldsymbol{\kappa}} x_1^{\kappa_1} \cdots x_n^{\kappa_n}$

- Coefficients of generating functions count things
- Lorentzian property implies log-concavity inequalities

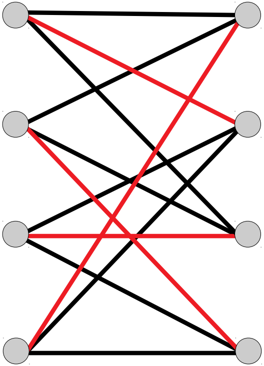
Cheap idea: Lorentzian $p(t,s) = \sum_{k=0}^{d} \binom{d}{k} c_k t^k s^{d-k}$

- Log-concavity of c_k implies $c_k \geq c_0^{1-\frac{k}{d}} c_d^{\frac{k}{d}} = p(0,1)^{1-\frac{k}{d}} p(1,0)^{\frac{k}{d}}$
- **Upshot:** Coefficients can be bounded by evaluations

Questions: Can we use this? Can we do this more delicately?

Example: Perfect matchings of a bipartite graph

Bipartite graph $G \implies$ **Lorentzian (real stable)** polynomial p_G:



$$\prod_{i=1}^{4} \sum_{j=1}^{4} \delta_{i \to j} x_j =$$

$$(x_1 + x_2 + x_3) \cdot (x_1 + x_3 + x_4) \cdot (x_2 + x_3 + x_4) \cdot (x_1 + x_2 + x_4)$$

Fact: The coefficient of $\mathbf{x}^{\mathbf{1}} = x_1 \cdots x_n$ counts perfect matchings

- Counting matchings $\iff$ computing permanent $\implies$ **#P-hard**
- Evaluation of p_G is **easy**, but computing coefficient of $\mathbf{x}^{\mathbf{1}}$ is **hard**

Gurvits' approach via polynomial capacity

Bipartite graph $G \implies$ **Lorentzian** $p_G(\boldsymbol{x}) = \prod_{i=1}^n \sum_{j=1}^n \delta_{i \to j} x_j$

Gurvits' capacity optimization problem: For $p \in \mathbb{R}_{\geq 0}[\boldsymbol{x}]$,

$$\mathrm{Cap}_{\mathbf{1}}(p) := \inf_{\boldsymbol{x}>0} \frac{p(\boldsymbol{x})}{\boldsymbol{x}^{\mathbf{1}}} = \inf_{x_1,\ldots,x_n>0} \frac{\sum_{\boldsymbol{\kappa}} c_{\boldsymbol{\kappa}} x_1^{\kappa_1} \cdots x_n^{\kappa_n}}{x_1 \cdots x_n}$$

- Easy **upper bound:** $\mathrm{Cap}_{\mathbf{1}}(p) \geq c_{\mathbf{1}}$ $\quad (c_{\mathbf{1}} = \#\text{pm for } p_G)$
- Up to log and exp, $\mathrm{Cap}_{\mathbf{1}}$ is a **convex program:**

$$\log \mathrm{Cap}_{\mathbf{1}}(p) = \inf_{\boldsymbol{y} \in \mathbb{R}^n} [\log p(e^{\boldsymbol{y}}) - \langle \boldsymbol{y}, \mathbf{1} \rangle]$$

Lower bound [Gurvits '04]**:** If p is **Lorentzian**, then we have

$$c_{\mathbf{1}} \;\geq\; \frac{n!}{n^n} \mathrm{Cap}_{\mathbf{1}}(p) \;\geq\; e^{-n} \mathrm{Cap}_{\mathbf{1}}(p)$$

- **Corollary:** $\mathrm{Cap}_{\mathbf{1}}(p_G)$ approximately counts perfect matchings of G
- **Bonus:** When G is regular, $\mathrm{Cap}_{\mathbf{1}}(p_G)$ can be computed exactly

Polynomials, probability, and entropy

Connection between probability and polynomials: (for $p(\mathbf{1}) = 1$)

finite distribution μ on $\mathbb{Z}_{\geq 0}^n$	$\iff$	polynomial p with ≥ 0 coeff.
strong Rayleigh distribution	$\iff$	real stable polynomial
$\mathbb{P}[\mu = \boldsymbol{\kappa}] = c_{\boldsymbol{\kappa}}$	$\iff$	$p(\boldsymbol{x}) = \sum_{\boldsymbol{\kappa} \in \mathbb{Z}_{\geq 0}^n} c_{\boldsymbol{\kappa}} \boldsymbol{x}^{\boldsymbol{\kappa}}$
$\mathbb{E}[\mu]$	$\iff$	$\nabla p(\mathbf{1})$

From polynomial capacity to entropy optimization:

primal program	$\iff$	dual program
$\inf_{\substack{\mathbb{E}[\nu]=\boldsymbol{\alpha} \\ \mathrm{supp}(\nu)\subseteq\mathrm{supp}(\mu)}} \mathrm{D}_{\mathrm{KL}}(\nu \\| \mu)$	$\iff$	$-\log \mathrm{Cap}_{\boldsymbol{\alpha}}(p)$

- D_{KL} = Kullback-Leibler divergence $\approx$ negative entropy
- **Strong duality:** Capacity solves the entropy optimization problem
- **Intuition:** Min. D_{KL} $\iff$ max. entropy $\iff$ no extra information

Another example: Counting contingency tables

Contingency table: Matrix $M \in \mathbb{Z}_{\geq 0}^{m \times n}$ with fixed row/col sums:

	β_1	β_2	$\cdots$	β_n
α_1	m_{11}	m_{12}	$\cdots$	m_{1n}
α_2	m_{21}	m_{22}	$\cdots$	m_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
α_m	m_{m1}	m_{m2}	$\cdots$	m_{mn}

Examples:

- **Permutation matrices** for $\boldsymbol{\alpha} = \boldsymbol{\beta} = \mathbf{1}$
- **Bipartite multigraph** adjacency matrices for fixed degrees $\boldsymbol{\alpha}, \boldsymbol{\beta}$
- Integer points of **transportation polytopes** for any $\boldsymbol{\alpha}, \boldsymbol{\beta}$

Motivation: (see [De Loera-Kim '14])

- **Combinatorial optimization:** Assignment problem
- **Statistics:** Dependence structure between random variables
- **Convex geometry:** Volume of transportation polytopes hard to compute

---

Question: Can we apply the entropy optimization method?

Applying the entropy optimization (capacity) method

Generating function: If $\mathrm{CT}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ counts contingency tables, then:

$$g(\boldsymbol{x}, \boldsymbol{y}) := \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{1}{1 - x_i y_j} = \prod_{i=1}^{m} \prod_{j=1}^{n} \sum_{k=0}^{\infty} (x_i y_j)^k = \sum_{\boldsymbol{\alpha}, \boldsymbol{\beta}} \mathrm{CT}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \cdot \boldsymbol{x}^{\boldsymbol{\alpha}} \boldsymbol{y}^{\boldsymbol{\beta}}$$

- Evaluation of $g(\boldsymbol{x}, \boldsymbol{y})$ is **easy**, but computing $\mathrm{CT}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is **hard**
- **Problem:** $g(\boldsymbol{x}, \boldsymbol{y})$ is **rational** instead of a homogeneous polynomial

---

Solution: Truncate to d and "twist", and then limit $d \to \infty$:

$$\tilde{g}_d(\boldsymbol{x}, \boldsymbol{y}) := \prod_{i=1}^{m} \prod_{j=1}^{n} \sum_{k=0}^{d} x_i^k y_j^{d-k} = \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{x_i^{d+1} - y_j^{d+1}}{x_i - y_j}$$

$\tilde{g}_d$ **is denormalized Lorentizian:** use bounds from [Brändén-L-Pak '21]

- $\mathrm{Cap}_{\boldsymbol{\alpha}, \boldsymbol{\beta}}(g) \geq \mathrm{CT}(\boldsymbol{\alpha}, \boldsymbol{\beta}) \geq e^{-m-n} \prod_i \frac{1}{\alpha_i + 1} \prod_j \frac{1}{\beta_j + 1} \mathrm{Cap}_{\boldsymbol{\alpha}, \boldsymbol{\beta}}(g)$
- **Upshot:** Technique can apply to rational generating functions
- Bound **volume of transportation polytopes** via $c\boldsymbol{\alpha}, c\boldsymbol{\beta}$ limiting $c \to \infty$

Other applications of the entropy optimization method

Via polynomials:

- Non-perfect bipartite matchings [Gurvits-L '18]
- Inner products [Anari-Oveis Gharan '17], [Gurvits-L '18] (e.g., counting matroid intersection via basis generating polys. [Anari-Liu-Oveis Gharan-Vinzant '18])
- Metric TSP approximation [Karlin-Klein-Oveis Gharan '21]

Combinatorial asymptotics:

- Asymptotic volume/points of polytopes [Barvinok '00s], [Barvinok-Hartigan '09]
- Asymptotic volume of spectrahedra [L-Ravichandran '22+]

Generalizations:

- **Entropy optimization on manifolds:** entropic rounding, private low-rank approximation, etc. [L-Vishnoi '20], [L-McSwiggen-Vishnoi '21]
- **Invariant theory:** matrix/operator scaling, orbit intersection problems, scaling problems, non-commutative optimization, etc. (**many** combos of Allen-Zhu, Bürgisser, Franks, Garg, Gurvits, Li, Oliveira, Reichenbach, Walter, Wigderson)
- **Statistics:** MLE, Gaussian graphical models, etc. [Améndola-Kohn-Reichenbach-Seigal '21], [Makam-Reichenbach-Seigal '21]

Open Questions

1. **Other generating functions** to which the techniques can be applied?

2. **Metric TSP** approximation improvement? (see [Gurvits-L '21])

3. Combinatorial applications of **entropy optimization on manifolds**?

4. New bounds via applying entropy optimization method to **Schur and Schubert polynomials**?

Thanks

Thanks!