Chapter 18

Real-Stability and Log-Concavity

In this chapter, we study the pioneer work of Gurvits on using real stable polynomials in combinatorial problems that inspired the recent developments, and we see some applications of real-stable polynomials in designing approximation algorithms for combinatorial optimization problems. This concludes the second part of the course about real stable polynomials, and connects to the third part of the course that involves log-concave polynomials. Our presentation follows closely that of the course notes by Oveis Gharan [Ove20].

18.1 Gurvits' Capacity Inequality

An influential concept defined by Gurvits is the capacity function [Gur04, Gur06].

Definition 18.1 (Capacity of a Polynomial). *Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a polynomial. The capacity of p is defined as*

$$\operatorname{cap}(p) := \inf_{x>0} \frac{p(x_1, \ldots, x_n)}{x_1 \cdots x_n}.$$

The main theorem in [Gur06] is to use the capacity of a real-stable polynomial p to estimate the coefficient of the monomial $x_1 \cdots x_n$ in p.

Theorem 18.2 (Gurvits [Gur06]). *Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a real-stable polynomial with non-negative coefficients. Then*

$$e^{-n} \cdot \operatorname{cap}(p) \leq \partial_{x_1} \cdots \partial_{x_n} p\big|_{x=0} \leq \operatorname{cap}(p).$$

The proof of the second inequality is easy and holds for any non-negative polynomial. The proof of the first inequality is deferred to the next section after we introduced log-concavity.

The optimization problem in the capacity function can be formulated as a convex program, which can be solved in polynomial time using only a value oracle.

Proposition 18.3 (Computing Capacity). *Given a real stable polynomial p with non-negative coefficients, and an oracle that for any $x \in \mathbb{R}^n$ returns the value $p(x)$, there is an algorithm to compute the capacity of p up to $(1+\epsilon)$ factor in time* $\operatorname{poly}\left(\langle p \rangle, 1/\epsilon\right)$*, where $\langle p \rangle := n + \deg(p) + |\log c_{\max}| + |\log c_{\min}|$ and $c_{\max}, c_{\min}$ are defined as the maximum and minimum coefficients in p.*

Proof Sketch: The idea is to do a change of variables to replace x_i by e^{y_i}, which is valid since $x > 0$. Then consider the logarithm of the objective function

$$\log \operatorname{cap}(p) = \inf_{y \in \mathbb{R}^n} \Big\{ \log\Big(p\big(e^{y_1}, \ldots, e^{y_n}\big)\Big) - \sum_{i=1}^{n} y_i \Big\}.$$

Since all the coefficients of p are non-negative, the term $\log\big(p\big(e^{y_1}, \ldots, e^{y_n}\big)\big)$ can be written as $\log \sum_i a_i e^{\langle b_i, y\rangle}$ where $a_i \geq 0$ and $b_i \in \mathbb{R}^n$ for all i (one term for each monomial). This is known as the log-sum-exponential function, which is a convex function in y (often used as a soft-max function in convex optimization).

Using the ellipsoid method to compute $\log \operatorname{cap}(p)$, one can implement the separation oracle using only a value oracle and analyze the time complexity by bounding the volumes of the outer ellipsoid and the inner ellipsoid. See [AO17] for the details. □

Permanent

Gurvits [Gur06] used Theorem 18.2 to approximate the permanent of a matrix and to give a beautiful proof of the Van der Waerden's conjecture.

Definition 18.4 (Permanent of a Matrix). *The permanent of a matrix $A \in \mathbb{R}^{n \times n}$ is defined as*

$$\operatorname{per}(A) = \sum_{\sigma \in S^n} \prod_{i=1}^{n} A_{i,\sigma(i)},$$

where the summation runs over the set of all permutations of n elements.

Gurvits' idea is to read the permanent of a matrix from the following real-stable polynomial.

Exercise 18.5 (Permanent Polynomial). *Given a non-negative matrix $A \in \mathbb{R}^{n \times n}_{\geq 0}$, the permanent polynomial is defined as*

$$p_A(x_1, \ldots, x_n) := \prod_{i=1}^{n} \sum_{j=1}^{n} A_{i,j} \cdot x_j.$$

Show that p_A is a homogeneous real-stable polynomial with non-negative coefficients and $\operatorname{per}(A) = \partial_{x_1} \cdots \partial_{x_n} p_A(x_1, \ldots, x_n)\big|_{x=0}$.

It follow from Theorem 18.2 and Exercise 18.5 that there is a determininstic e^n-approximation algorithm for the permanent of a non-negative matrix. The best known deterministic approximation ratio is $(\sqrt{2})^n$ by Anari and Rezaei [AR19]. A major breakthrough in approximate counting is a randomized $(1+\epsilon)$-approximation algorithm for the permanent of a non-negative matrix by Jerrum, Sinclair and Vigoda [JSV04] with running time polynomial in n and $1/\epsilon$. Their method is to use Markov chains to sample a uniform random perfect matching of the bipartite graph of the input matrix. It has been a long standing open question to match this result with a deterministic algorithm.

Van der Waerden conjectured that the permanent of an $n \times n$ non-negative doubly stochastic matrix is at least e^{-n}. This conjecture was proven in the 80s by Gyires and by Egorychev and Falikman. Gurvits provided a simple and elegant proof using Theorem 18.2 and the AM-GM inequality (which follows from the concavity of the logarithmic function).

Exercise 18.6 (Weighted AM-GM Inequality). *Let $a_1, \ldots, a_n \geq 0$ and $\mu_1, \ldots, \mu_n \geq 0$ with $\sum_{i=1}^n \mu_i = 1$. Then*

$$\sum_{i=1}^n \mu_i a_i \geq \prod_{i=1}^n a_i^{\mu_i}.$$

Theorem 18.7 (Van der Waerden's Conjecture). *The permanent of any non-negative doubly stochastic matrix $A \in \mathbb{R}^{n \times n}$ is at least e^{-n}, where a matrix is called doubly stochastic if every row sum and every column sum is equal to 1.*

Proof. Let p_A be the permanent polynomial in Exercise 18.5. By Exercise 18.5 and Theorem 18.2,

$$\text{per}(A) = \partial_{x_1} \cdots \partial_{x_n} p_A(x_1, \ldots, x_n)\big|_{x=0} \geq e^{-n} \cdot \text{cap}(p_A),$$

and so the statement would follow if we could prove that $\text{cap}(p_A) \geq 1$ for any non-negative doubly stochastic matrix A. For any $x \in \mathbb{R}_+^n$,

$$p_A(x_1, \ldots, x_n) = \prod_{i=1}^n \sum_{j=1}^n A_{i,j} x_j \geq \prod_{i=1}^n \prod_{j=1}^n x_j^{A_{i,j}} = \prod_{i=1}^n x_i^{\sum_{j=1}^n A_{i,j}} = \prod_{i=1}^n x_i,$$

where the inequality is by the weighted AM-GM inequality in Exercise 18.6 and the assumption that every row sum is equal to one, and the last equality is by the assumption that every column sum is equal to one. This implies that $\text{cap}(p_A) \geq 1$ and completes the proof. □

Gurvits' result can also be used to give a simple proof of the following bound by Schrijver, whose original proof is combinatorial and highly complicated.

Problem 18.8 (Schrijver's Bound). *Let G be a k-regular bipartite graph with n vertices. Prove that the number of perfect matchings in G is at least $\left(\frac{k}{e}\right)^n$.*

18.2 Log-Concavity

We follow the proof of Theorem 18.2 by Oveis Gharan [Ove20] that highlights the role of log-concavity. The following simple lemma about univariate polynomials will be used in the base case.

Lemma 18.9 (Log-Concavity of Non-Negative Real-Rooted Polynomial). *Let $f \in \mathbb{R}[x]$ be a real-rooted polynomial with non-negative coefficients then $\log f$ is a concave function on $\mathbb{R}_{\geq 0}$.*

Proof. Let $\alpha_1, \ldots, \alpha_n$ be the roots of f. As all coefficients of f are non-negative, it follows that $f(x) > 0$ for all $x > 0$ as long as $f \not\equiv 0$, and so all the roots of f must be non-positive. Assume without loss that the leading coefficient of f is one, then

$$\log f = \log\left(\prod_{i=1}^n (x - \alpha_i)\right) = \sum_{i=1}^n \log(x - \alpha_i).$$

Since $\alpha_i \leq 0$ for $1 \leq i \leq n$, we conclude that $\log f$ is concave on $\mathbb{R}_{\geq 0}$ as each $\log(x - \alpha_i)$ is well-defined and concave on $\mathbb{R}_{\geq 0}$ and the sum of concave functions is a concave function. □

The following is Theorem 18.2 in the univariate case.

Lemma 18.10 (Univariate Case of Theorem 18.2)**.** *For any real-rooted polynomial $f \in \mathbb{R}[x]$ with non-negative coefficients,*

$$f'(0) \geq \frac{1}{e} \inf_{x>0} \frac{f(x)}{x}.$$

Proof. If $f(0) = 0$, then $f'(0) = \inf_{x>0} \frac{f(x)}{x}$ as $f(x)$ is a convex function in x, and so the inequality holds trivially. Henceforth we assume that $f(0) > 0$. By log-concavity of f from Lemma 18.9, for any $x \geq 0$,

$$\log f(x) \leq \log f(0) + x(\log f(0))' \quad \Longrightarrow \quad \log \frac{f(x)}{x} \leq \log f(0) + x\frac{f'(0)}{f(0)} - \log x$$

The RHS is minimized when $x = f(0)/f'(0)$, and this implies that

$$\inf_{x>0} \log \frac{f(x)}{x} \leq \log f(0) + 1 - \log \frac{f(0)}{f'(0)} \quad \Longrightarrow \quad 1 + \log f'(0) \geq \inf_{x>0} \log \frac{f(x)}{x},$$

which implies the lemma. □

We are ready to prove Theorem 18.2.

Proof of Theorem 18.2. The proof is by induction on the number of variables n. Let

$$q(x_1, \ldots, x_{n-1}) := \partial_{x_n} p|_{x_n=0}.$$

Note that q is real-stable as differentiation and substituting real number preserve real stability by Exercise 13.17 and Proposition 13.13, and also q has non-negative coefficients as p has.

For any $x_1, \ldots, x_{n-1} > 0$, consider the univariate polynomial $f(x_n) := p(x_1, \ldots, x_{n-1}, x_n)$. Note that f is real-stable and thus real-rooted, and also $f'(0) = q(x_1, \ldots, x_{n-1})$. Applying Lemma 18.10 on f,

$$q(x_1, \ldots, x_{n-1}) = f'(0) \geq \frac{1}{e} \inf_{x_n>0} \frac{f(x_n)}{x_n} = \frac{1}{e} \inf_{x_n>0} \frac{p(x_1, \ldots, x_n)}{x_n}.$$

Using this and applying the induction hypothesis on q, we conclude that

$$\begin{aligned} \partial_{x_1} \cdots \partial_{x_n} p\big|_{x_1=\ldots=x_n=0} &= \partial_{x_1} \cdots \partial_{x_{n-1}} q\big|_{x_1=\ldots=x_{n-1}=0} \\ &\geq e^{-n-1} \inf_{x_1,\ldots,x_{n-1}>0} \frac{q(x_1, \ldots, x_{n-1})}{x_1 \cdots x_{n-1}} \geq e^{-n} \inf_{x_1,\ldots,x_n>0} \frac{p(x_1, \ldots, x_n)}{x_1 \cdots x_n}. \end{aligned}$$

□

Finally, we prove the following generalization of Lemma 18.9, which can be seen as a generalization of the well-known fact that $\det(X)$ is log-concave over the space of positive semidefinite matrices. This result will be useful in the next section for designing approximation algorithms.

Theorem 18.11 (Log-Concavity of Homogeneous Real-Stable Polynomials)**.** *Let $p \in \mathbb{R}[x_1, \ldots, x_n]$ be a homogeneous real-stable polynomial with non-negative coefficients. Then p is log-concave on $\mathbb{R}^n_{\geq 0}$.*

Proof. To prove the statement, we will prove that $\log p(a+tb)$ is concave along the line $a+tb$, for any $a \in \mathbb{R}_+^n$ and $b \in \mathbb{R}_n$ such that $a+tb \in \mathbb{R}_+^n$ for all $t \in [0,1]$. Let p be homogeneous of degree k. Then

$$p(a+tb) = p\Big(t\Big(\frac{a}{t}+b\Big)\Big) = t^k \cdot p\Big(\frac{a}{t}+b\Big).$$

Since p is real-stable and $a \in \mathbb{R}_+^n$, $p(at+b)$ is real rooted. Let $p(at+b) = c\prod_{i=1}^k (t-\lambda_i)$ where $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ are the roots. Then $p(\frac{a}{t}+b) = c\prod_{i=1}^k(\frac{1}{t}-\lambda_i)$, and so

$$p(a+tb) = t^k \cdot p\Big(\frac{a}{t}+b\Big) = c\prod_{i=1}^k (1-t\lambda_i)$$

Note that $\lambda_i < 1$ for $1 \leq i \leq k$, as otherwise there exists $t \in [0,1]$ such that $p(a+tb) = 0$, contradicting to our assumption that the line $a+tb \in \mathbb{R}_+^n$ for $t \in [0,1]$ and so $p(a+tb) > 0$ as p has non-negative coefficients. Therefore,

$$\log p(a+tb) = \log(c) + \sum_{i=1}^k \log(1-t\lambda_i),$$

which is a concave function as each $\log(1-t\lambda_i)$ is a concave function of t for $t \in [0,1]$ when $\lambda_i < 1$. □

18.3 Determinant Maximization

In this section, we see some applications of Gurvits' capacity inequality in Theorem 18.2 and the log-concavity of real-stable polynomial in designing approximation algorithms for some combinatorial optimization problems.

The determinant maximization problem is closely related to the D-design problem that we have discussed in Section 11.1.

Definition 18.12 (Determinant Maximization Problem). *Given a positive semidefinite matrix $M \in \mathbb{R}^{n\times n}$ and an integer k, the goal is to output a set $S \subseteq [n]$ with $|S| = k$ that maximizes* $\det(M_{S,S})$.

Oveis Gharan [Ove20] give a simple proof of the following result by Nikolov [Nik15] using the theory of real-stable polynomials.

Theorem 18.13 (Nikolov [Nik15]). *There is a polynomial time algorithm that gives a e^{-k} approximation to the determinant maximization problem.*

Proof Sketch: Consider the following mathematical program for the problem:

$$\begin{aligned} \max \quad & \log \sum_{S\subseteq[n]:|S|=k} \det(M_{S,S}) \prod_{i\in S} x_i \\ \text{subject to} \quad & \sum_{i=1}^n x_i = k \\ & x_i \geq 0 \quad \text{for } 1 \leq i \leq n. \end{aligned}$$

First we argue that the program is convex and can be solved in polynomial time. We know from Problem 16.15 that the polynomial $\sum_{S\subseteq[n]:|S|=k}\det(M_{S,S})\prod_{i\in S} x_i$ is real-stable when $M \succcurlyeq 0$. As this polynomial is homogeneous with non-negative coefficients as $M \succcurlyeq 0$, it follows from Theorem 18.11 that the objective function is a concave function, and thus the program is a convex program. We remark that the proof of Problem 16.15 also provides a compact representation of the polynomial in the objective function, so that we can evaluate the objective function in time polynomial in terms of the size of the input M.

Note that the convex program is a relaxation of the determinant maximization problem, and so its objective value $\sum_{S \subseteq [n]:|S|=k} \det(M_{S,S}) \prod_{i \in S} x_i$ is at least the optimal value opt. Given an optimal solution x to the convex program, we consider the following simple randomized rounding algorithm. Let μ be the distribution on $[n]$ where $\mu(i) = x_i/k$. Choose k indexes $i_1, i_2, \ldots, i_k \in [n]$, where each index i_j is sampled from μ independently. If $i_1, \ldots, i_k$ are all distinct, then we output $S = \{i_1, \ldots, i_k\}$, otherwise we output "failed". Then, the expected objective value of the output is

$$\mathbb{E}\left[\mathsf{alg}\right] = \sum_{S \in \binom{[n]}{k}} \Pr[S \text{ is sampled}] \cdot \det(M_{S,S}) = \sum_{S \in \binom{[n]}{k}} \left(k! \cdot \prod_{i \in S} \frac{x_i}{k}\right) \cdot \det(M_{S,S}) \geq e^{-k} \cdot \mathsf{opt},$$

where the second equality is because there are $k!$ permutations to choose the same subset S, each permutation with probability $\prod_{i \in S} \frac{x_i}{k}$. This bounds the integrality gap of the convex program, but note that the randomized rounding algorithm is not a polynomial time algorithm (see Exercise 18.14). Nikolov [Nik15] derandomized this analysis using conditional expectation to give a deterministic polynomial time algorithm with the same guarantee. □

Exercise 18.14 (Exponential Running Time)**.** *Show an example where the randomized rounding algorithm in the proof of Theorem 18.13 runs in time* $\Omega(e^k)$.

Problem 18.15 (Sampling by Volume)**.** *Suppose there is a polynomial time algorithm that outputs a random size-k subset S with probability proportional to* $\det(M_{S,S})$*. Show that this algorithm can be used to give a randomized polynomial time e^{-k}-approximation algorithm for the determinant maximization problem.*

Determinant Maximization with Partition Constraints

Nikolov and Singh [NS16] considered the generalization of the determinant maximization problem with partition constraints. The following is a simple version of their problem.

Definition 18.16 (Determinant Maximization with Partition Constraints)**.** *Given a positive semidefinite matrix $M \in \mathbb{R}^{n \times n}$, an integer k and a partition of the ground set $[n]$ into $P_1 \cup P_2 \cup \ldots \cup P_k$, the goal is to output a set S with $|S \cap P_i| = 1$ for $1 \leq i \leq k$ that maximizes* $\det(M_{S,S})$.

We briefly discuss the main ingredients in [NS16]. The natural relaxation in the proof of Theorem 18.13 (with suitable modification) has unbounded integrality gap. The key contribution by Nikolov and Singh is to come up with a very interesting convex relaxation for the problem. Let $\mathcal{B} := \{S \subseteq [n] \mid |S \cap P_i| = 1\ \forall 1 \leq i \leq k\}$ be the set of subsets that satisfy the partition constraints.

Write $M = V^T V$ and let v_i be the i-th column of V. The relaxation in [NS16] is

$$\begin{aligned} \mathsf{opt} \quad &= \quad \sup_x \inf_y \det\Big(\sum_{i=1}^n x_i y_i v_i v_i^T\Big) \\ \text{subject to} \quad & \quad \sum_{j \in P_i} x_j = 1 \qquad \forall 1 \le i \le k \\ & \quad 0 \le x_j \le 1 \qquad \forall 1 \le j \le n, \\ & \quad \prod_{i \in S} y_i = 1 \qquad \forall S \in \mathcal{B}. \end{aligned}$$

Nikolov and Singh showed that it is indeed a relaxation and it can be solved in polynomial time. They analyzed the simple rounding algorithm where we choose one vector in P_i with probability distribution $\{x_j\}_{j \in P_i}$. The analysis uses the real-stability of the polynomial $p(y) := \det(\sum_{i=1}^n x_i y_i v_i v_i^T)$. They reduced the problem of bounding the expected value of the output to bounding the coefficient of the monomial $z_1 \ldots z_m$ of a related real-stable polynomial, and then they applied Gurvits' inequality in Theorem 18.2 to prove that the convex program gives a e^{-k}-approximation to the problem.

Generalized Permanent Inequality

Extending the convex program in [NS16], Anari and Oveis Gharan [AO17] obtained an elegant generalization of Gurvits' permanent inequality. The following is a simpler version of their main theorem.

Theorem 18.17 (Generalized Permanent Inequality [AO17])**.** *For any two multi-linear real-stable polynomial $p, q \in \mathbb{R}[x_1, \ldots, x_n]$ with non-negative coefficients,*

$$\sup_{\alpha \ge 0} \inf_{x,y>0} e^{-\alpha} \frac{p(x)q(y)}{(xy/\alpha)^\alpha} \le \sum_\kappa c_p(\kappa) \cdot c_q(\kappa) \le \sup_{\alpha \ge 0} \inf_{x,y>0} \frac{p(x)q(y)}{(xy/\alpha)^\alpha},$$

where $\alpha, x, y \in \mathbb{R}^n$ are vectors, and κ is over the set of monomials of p and q with coefficients $c_p(\kappa)$ and $c_q(\kappa)$ respectively. For two vectors $a, b \in \mathbb{R}^n$, $ab \in \mathbb{R}^n$ denotes the vector with the i-th entry being $a_i b_i$, and $a^b \in \mathbb{R}$ denotes the number $\prod_{i=1}^n a_i^{b_i}$.

They showed that this result generalizes Gurvits' inequality and Nikolov-Singh's result, and provides deterministic polynomial time algorithms for several counting problems.

Concluding Remark

Besides the applications we discussed, Gurvits' concept of capacity has found important applications in analyzing scaling problems including matrix scaling, frame scaling and operator scaling, and there are various applications of these scaling problems (see e.g. [GGOW20]).

18.4 References

[AO17] Nima Anari and Shayan Oveis Gharan. A generalization of permanent inequalities and applications in counting and optimization. In *Proceedings of the 49th Annual ACM*

SIGACT Symposium on Theory of Computing, STOC 2017, Montreal, QC, Canada, June 19-23, 2017, pages 384–396. ACM, 2017. 172, 177

[AR19] Nima Anari and Alireza Rezaei. A tight analysis of bethe approximation for permanent. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 1434–1445. IEEE Computer Society, 2019. 172

[GGOW20] Ankit Garg, Leonid Gurvits, Rafael Oliveira, and Avi Wigderson. Operator scaling: theory and applications. *Foundations of Computational Mathematics*, 20(2):223–290, 2020. 177

[Gur04] Leonid Gurvits. Classical complexity and quantum entanglement. *Journal of Computer and System Sciences*, 69(3):448–484, 2004. 171

[Gur06] Leonid Gurvits. Hyperbolic polynomials approach to Van der Waerden/Schrijver-Valiant like conjectures: sharper bounds, simpler proofs and algorithmic applications. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 417–426, 2006. 171, 172

[JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, 2004. 53, 172

[Nik15] Aleksandar Nikolov. Randomized rounding for the largest simplex problem. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 861–870. ACM, 2015. 175, 176

[NS16] Aleksandar Nikolov and Mohit Singh. Maximizing determinants under partition constraints. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2016, Cambridge, MA, USA, June 18-21, 2016*, pages 192–201. ACM, 2016. 176, 177

[Ove20] Shayan Oveis Gharan. Course notes of "Polynomial paradigm in algorithm design", lecture 3,6. 2020. https://homes.cs.washington.edu/~shayan/courses/polynomials/. 128, 130, 153, 158, 171, 173, 175