

# A Generalization of Permanent Inequalities and Applications in Counting and Optimization

Nima Anari<sup>1</sup> and Shayan Oveis Gharan<sup>\*2</sup>

<sup>1</sup>Stanford University, [anari@stanford.edu](mailto:anari@stanford.edu)

<sup>2</sup>University of Washington, [shayan@cs.washington.edu](mailto:shayan@cs.washington.edu)

## Abstract

A polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  is real stable if it has no roots in the upper-half complex plane. Gurvits's permanent inequality gives a lower bound on the coefficient of the  $z_1 z_2 \dots z_n$  monomial of a real stable polynomial  $p$  with nonnegative coefficients. This fundamental inequality has been used to attack several counting and optimization problems.

Here, we study a more general question: Given a stable multilinear polynomial  $p$  with nonnegative coefficients and a set of monomials  $S$ , we show that if the polynomial obtained by summing up all monomials in  $S$  is real stable, then we can lowerbound the sum of coefficients of monomials of  $p$  that are in  $S$ . We also prove generalizations of this theorem to (real stable) polynomials that are not multilinear. We use our theorem to give a new proof of Schrijver's inequality on the number of perfect matchings of a regular bipartite graph, generalize a recent result of Nikolov and Singh [NS16], and give deterministic polynomial time approximation algorithms for several counting problems.

---

\*This material is based on work supported by the National Science Foundation Career award.

# 1 Introduction

Suppose we are given a polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients. For a vector  $\boldsymbol{\kappa} \in \mathbb{Z}_+^n$  we write  $C_p(\boldsymbol{\kappa})$  to denote the coefficient of the monomial  $\prod_{i=1}^n z_i^{\kappa_i}$  in  $p$ . Given a set  $S$  of  $\boldsymbol{\kappa}$ 's, we want to estimate

$$\sum_{\boldsymbol{\kappa} \in S} C_p(\boldsymbol{\kappa}).$$

Let  $\partial_{z_i}$  be the operator that performs partial differentiation in  $z_i$ , i.e.,  $\partial/\partial z_i$ . We can rewrite the above quantity using a differential operator. Let

$$q(\partial \mathbf{z}) = \sum_{\boldsymbol{\kappa} \in S} \prod_{i=1}^n \partial_{z_i}^{\kappa_i}$$

Then,

$$\sum_{\boldsymbol{\kappa} \in S} \boldsymbol{\kappa}! C_p(\boldsymbol{\kappa}) = q(\partial \mathbf{z}) p(\mathbf{z})|_{\mathbf{z}=0}.$$

More generally, given two polynomials  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients we analytically and algorithmically study the quantity

$$q(\partial \mathbf{z}) p(\mathbf{z})|_{\mathbf{z}=0} = p(\partial \mathbf{z}) q(\mathbf{z})|_{\mathbf{z}=0} = \sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} \boldsymbol{\kappa}! C_p(\boldsymbol{\kappa}) C_q(\boldsymbol{\kappa}). \quad (1)$$

for a large class of polynomials known as real stable polynomials.

Before stating our results, let us motivate this question.

**Permanent:** For a matrix  $A \in \mathbb{R}_+^{n \times n}$  recall that the permanent of  $A$  is

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

Then,  $\text{per}(A)$  is  $C_p(\mathbf{1})$  of the polynomial  $p$ ,

$$p(z_1, \dots, z_n) = \prod_{i=1}^n \sum_{j=1}^n A_{i,j} z_j.$$

**Matroid Intersection:** Let  $\mathcal{M}_1([n], I_1)$  and  $\mathcal{M}_2([n], I_2)$  be two matroids on elements  $[n] = \{1, 2, \dots, n\}$  with set of bases  $\mathcal{B}_1, \mathcal{B}_2$  respectively. Let

$$\begin{aligned} p(\mathbf{z}) &= \sum_{B \in \mathcal{B}_1} \prod_{i \in B} z_i, \\ q(\partial \mathbf{z}) &= \sum_{B \in \mathcal{B}_2} \prod_{i \in B} \partial_{z_i}. \end{aligned}$$

Then,  $p(\mathbf{z}) q(\partial \mathbf{z})|_{\mathbf{z}=0}$  is the number of bases in the intersection of  $\mathcal{M}_1$  and  $\mathcal{M}_2$ .

We say a polynomial  $p(z_1, \dots, z_n) \in \mathbb{R}[z_1, \dots, z_n]$  is *real stable* if it has no roots in the upper half complex plane. Theory of stable polynomials recently had many applications in multiple

areas of mathematics and computer science [Gur06, MSS13, AO15a, AO15b]. Gurvits used this theory to give a new proof of the van der Waerden conjecture [Gur06]. Most notably he proved that for any  $n$ -homogeneous real stable polynomial  $p(z_1, \dots, z_n)$  with nonnegative coefficients and  $q = \partial z_1 \dots \partial z_n$ ,

$$e^{-n} \inf_{\mathbf{z} > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n} \leq q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0} \leq \inf_{\mathbf{z} > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}. \quad (2)$$

In other words, the coefficient of the monomial  $z_1 \dots z_n$  is at least  $e^{-n} \inf_{\mathbf{z} > 0} \frac{p(z_1, \dots, z_n)}{z_1 \dots z_n}$ . One can use this inequality to give a deterministic  $e^{-n}$  approximation algorithm for the permanent of a nonnegative matrix.

In this paper we prove bounds analogous to (2) for the quantity  $p(\mathbf{z})q(\partial \mathbf{z})$  when  $p, q$  are real stable polynomials. We discuss several applications in counting and optimization. We expect to see many more applications in the future.

## 1.1 Our Contributions

For two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  we write  $\mathbf{x}\mathbf{y}$  to denote the  $n$  dimensional vector where the  $i$ -th coordinate is  $x_i y_i$ . For vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  we write  $\mathbf{x}^{\mathbf{y}} := \prod_{i=1}^d x_i^{y_i}$ . We say a polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  is multilinear if  $C_p(\boldsymbol{\kappa}) = 0$  for  $\boldsymbol{\kappa} \notin \{0, 1\}^n$ . For a vector  $\boldsymbol{\alpha} \in \mathbb{R}^n$ , we write  $\boldsymbol{\alpha} \geq 0$  to denote  $\boldsymbol{\alpha} \in \mathbb{R}_+^n$ .

**Theorem 1.1.** *For any two real stable polynomials  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients,*

$$\sup_{\boldsymbol{\alpha} \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} e^{-\boldsymbol{\alpha}} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^{\boldsymbol{\alpha}}} \leq q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0} \leq \sup_{\boldsymbol{\alpha} \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^{\boldsymbol{\alpha}}}. \quad (3)$$

Furthermore, if  $p, q$  are multilinear then

$$\sup_{\boldsymbol{\alpha} \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} (1 - \boldsymbol{\alpha})^{1-\boldsymbol{\alpha}} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^{\boldsymbol{\alpha}}} \leq q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0}. \quad (4)$$

Note that when  $\|\boldsymbol{\alpha}\|_{\infty} > 1$ , the term  $(1 - \boldsymbol{\alpha})^{1-\boldsymbol{\alpha}}$  is ill-defined in the above theorem. However in that case we can replace it by any number, since by Fact 2.18, the LHS will be zero.

Before describing our main applications of the above theorem let us derive Gurvits's inequality (2). Let  $q(\mathbf{z}) = z_1 z_2 \dots z_n$ . Then, it follows that the supremum is obtained by  $\boldsymbol{\alpha} = \mathbf{1}$ . For that  $\boldsymbol{\alpha}$  we have

$$\inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^{\boldsymbol{\alpha}}} = \inf_{\mathbf{y} > 0} \frac{p(\mathbf{y})}{y_1 \dots y_n}.$$

In Section 3 we discuss several applications of the above theorem. The following two theorems are the main consequences that we prove. Firstly, we give an algorithm to approximate  $q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0}$  for two given stable polynomials  $p, q$  with nonnegative coefficients. We use this to give deterministic approximate counting algorithms. In our results, we use convex programs involving certain polynomials. We make sure these convex programs can be approximately solved in time polynomial in the complexity of the involved polynomials. For a polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients, we define its complexity as  $n + \deg p + |\log \min_{\boldsymbol{\kappa}: C_p(\boldsymbol{\kappa}) \neq 0} C_p(\boldsymbol{\kappa})| + |\log \max_{\boldsymbol{\kappa}: C_p(\boldsymbol{\kappa}) \neq 0} C_p(\boldsymbol{\kappa})|$ , and we denote this by  $\langle p \rangle$ .

**Theorem 1.2.** *Given two real stable polynomials  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients and an oracle that for any  $\mathbf{z} \in \mathbb{R}^n$ , returns  $p(\mathbf{z}), q(\mathbf{z})$ , there is a polynomial time algorithm that returns an  $e^{\min\{\deg p, \deg q\} + \epsilon}$  approximation of  $q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0}$  that runs in time  $\text{poly}(\langle p \rangle + \langle q \rangle + \log(1/\epsilon))$ .*

Our next theorem gives a polynomial time algorithm to approximate  $\max_{\kappa \in \{0,1\}^n} C_p(\kappa)C_q(\kappa)$  for two given multilinear stable polynomials  $p, q$  with nonnegative coefficients. We use this to generalize a recent results of Nikolov and Singh [NS16].

**Theorem 1.3.** *Given two real stable multilinear  $D$ -homogeneous polynomials  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients and an oracle that for any  $\mathbf{z} \in \mathbb{R}^n$  returns  $p(\mathbf{z}), q(\mathbf{z})$ , there is a polynomial time algorithm that outputs an  $e^{2D+\epsilon}$  approximation to*

$$\max_{\kappa \in \{0,1\}^n} C_p(\kappa)C_q(\kappa).$$

that runs in time  $\text{poly}(\langle p \rangle + \langle q \rangle + \log(1/\epsilon))$ .

Independent of our work, Straszak and Vishnoi [SV16] studied several variants of our problem. They show that if  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  are multilinear real stable polynomials with nonnegative coefficients, there is a convex relaxation to the quantity  $q(\partial \mathbf{z})p(\mathbf{z})$  with an integrality gap at most  $e^n$ . In contrast, our approximation factor depends only on  $\min\{\deg p, \deg q\}$  which can be significantly smaller than  $n$ .

**Structure of the paper.** In Section 2 we describe many properties of real stable polynomials that we use throughout the paper. In Section 3 we describe several applications of Theorem 1.1 in counting and optimization. Finally, in sections 4 and 5 we prove Theorem 1.1. In particular, in Section 4 we lower bound  $q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0}$ , i.e., we prove the (4) and the LHS of (3). And, in Section 5 we prove upper bound this quantity, i.e., the RHS of (3).

## 2 Preliminaries

Throughout the paper, we use bold letters to denote vectors. For a vector  $\mathbf{y}$ , we write  $\mathbf{y} \leq 1$  to denote that all coordinates of  $\mathbf{y}$  are at most 1. For two vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  we define  $\mathbf{x}\mathbf{y} = (x_1y_1, \dots, x_ny_n)$ . Similarly, we define  $\mathbf{x}/\mathbf{y} = (x_1/y_1, \dots, x_n/y_n)$ . For a vector  $\mathbf{y} \in \mathbb{R}^n$ , we define  $\exp(\mathbf{x}) := (e^{x_1}, \dots, e^{x_n})$ .

For vectors  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  we write

$$\mathbf{x}^{\mathbf{y}} := \prod_{i=1}^n x_i^{y_i}.$$

As a special case, for a real number  $c \in \mathbb{R}$  we write  $c^{\mathbf{x}}$  to denote  $\prod_{i=1}^n c^{x_i}$ .

We use  $\mathbb{R}_{++} = \{x : x > 0\}$  to denote the set of positive real numbers. For an integer  $n \geq 1$  we use  $[n]$  to denote the set of numbers  $\{1, 2, \dots, n\}$ . For any  $m, n$ , we let  $\binom{[m]}{n}$  denote the collection of subsets of  $[m]$  of size  $n$ . Throughout the paper all logs are in base  $e$ .

For a vector  $\mathbf{x} \in \mathbb{R}^n$  and an integer  $1 \leq i \leq n$  we use  $\mathbf{x}_{-i}$  to denote the  $n-1$  dimensional vector obtained from  $\mathbf{x}$  by dropping the  $i$ -th coordinate of  $\mathbf{x}$ . We use  $\mathbb{Z}_+$  to denote the set of nonnegative integers.

We say a matrix  $A \in \mathbb{R}^{n \times n}$  is *doubly stochastic* if all entries of  $A$  are nonnegative and all row sums and columns sums are equal to 1,

$$\begin{aligned} \sum_j A_{i,j} &= 1 \quad \forall i, \\ \sum_i A_{i,j} &= 1 \quad \forall j. \end{aligned}$$

## 2.1 Stable Polynomials

Stable polynomials are natural multivariate generalizations of real-rooted univariate polynomials. For a complex number  $z$ , let  $\text{Im}(z)$  denote the imaginary part of  $z$ . We say a polynomial  $p(z_1, \dots, z_n) \in \mathbb{C}[z_1, \dots, z_n]$  is *stable* if whenever  $\text{Im}(z_i) > 0$  for all  $1 \leq i \leq n$ ,  $p(z_1, \dots, z_n) \neq 0$ . As the only exception, we also call the zero polynomial stable. We say  $p(\cdot)$  is real stable, if it is stable and all of its coefficients are real. It is easy to see that any univariate polynomial is real stable if and only if it is real rooted.

We say a polynomial  $p(z_1, \dots, z_n)$  is degree  $k$ -homogeneous, or  $k$ -homogeneous, if every monomial of  $p$  has total degree exactly  $k$ . Equivalently,  $p$  is  $k$ -homogeneous if for all  $a \in \mathbb{R}$ , we have

$$p(a \cdot z_1, \dots, a \cdot z_n) = a^k p(z_1, \dots, z_n).$$

We say a monomial  $z_1^{\alpha_1} \dots z_n^{\alpha_n}$  is *square-free* if  $\alpha_1, \dots, \alpha_n \in \{0, 1\}$ . We say a polynomial is multilinear if all of its monomials are square-free. For a polynomial  $p$ , we write  $\deg p$  to denote the maximum degree of all monomials of  $p$ . One of the famous examples of real stable polynomials is the set of elementary symmetric polynomials. For an integer  $k \geq 0$ , the  $k$ -th elementary symmetric polynomial in  $z_1, \dots, z_n$ ,  $e_k(z_1, \dots, z_n)$  is defined as

$$e_k(z_1, \dots, z_n) := \sum_{S \in \binom{[n]}{k}} \prod_{i \in S} z_i.$$

Note that the above polynomial is  $k$ -homogeneous and multilinear.

Let  $\mu$  be a probability distribution on subsets of  $[n]$ . The *generating polynomial* of  $\mu$  is defined as follows:

$$g_\mu(z_1, \dots, z_n) = \sum_S \mu(S) \prod_{i \in S} z_i.$$

We say  $\mu$  is a *strongly Rayleigh* (SR) distribution if  $g_\mu$  is a real stable polynomial. Borcea, Brändén, and Liggett [BBL09] defined (SR) distributions and proved numerous properties of them. Many interesting family of probability distributions are special cases of SR including product distributions, random spanning trees, and determinantal point processes. Note that  $\mu(S) \geq 0$  for all  $S \subseteq [n]$ ; so  $g_\mu(S)$  has nonnegative coefficients. Many of theorems that we prove here have natural probabilistic interpretations if the underlying real stable polynomial is a generating polynomial of a SR distribution.

We say a polynomial  $p(y_1, \dots, y_n, z_1, \dots, z_n)$  is *bistable* if  $p(y_1, \dots, y_n, -z_1, \dots, -z_n)$  is a real stable polynomial.

**Fact 2.1.** For any set of PSD matrices  $A_1, A_2, \dots, A_n \in \mathbb{R}^{d \times d}$  and a symmetric matrix  $B \in \mathbb{R}^{d \times d}$

$$p(z_1, \dots, z_n) = \det(B + z_1 A_1 + \dots + z_n A_n)$$

is real stable.

**Fact 2.2.** If  $p(\mathbf{z})$  and  $q(\mathbf{z})$  are real stable, then  $p(\mathbf{z})q(\mathbf{z})$  is real stable.

**Fact 2.3.** If  $p(\mathbf{y})$  and  $q(\mathbf{z})$  are real stable, then  $p(\mathbf{y})q(\mathbf{z})$  is bistable.

**Fact 2.4.** For any stable polynomial  $p(z_1, \dots, z_n)$ , an integer  $1 \leq i \leq n$ , and any number  $c$  with nonnegative imaginary part,  $\text{Im}(c) \geq 0$ ,  $p(z_1, \dots, z_{i-1}, c, z_{i+1}, \dots, z_n)$  is stable.

**Fact 2.5.** *If  $p(\mathbf{y})$  is real stable, then for any nonnegative vector  $\boldsymbol{\lambda} \in \mathbb{R}_+^n$ ,  $p(\boldsymbol{\lambda}\mathbf{y})$  is real stable.*

**Lemma 2.6.** *For any real stable polynomial  $p(z_1, z_2, \dots, z_n)$  and any  $1 \leq i, j \leq n$ ,  $(1 - \partial_{z_i} \partial_{z_j})p$  is stable.*

*Proof.* We say an operator  $T$  is stability preserver if for any real stable polynomial  $p(\mathbf{z})$ ,  $Tp$  is real stable. So, to prove the lemma we need to show that  $1 - \partial_{z_i} \partial_{z_j}$  is a stability preserver operator. Borcea and Brändén [BB10] showed that for any differential operator  $T = \sum_{\boldsymbol{\kappa}} c_{\boldsymbol{\kappa}} \partial_{\mathbf{z}}^{\boldsymbol{\kappa}}$ ,  $T$  is stability preserver if the polynomial

$$\sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} c_{\boldsymbol{\kappa}} (-\mathbf{z})^{\boldsymbol{\kappa}}$$

is real stable. So, to prove the lemma it is enough to show that  $1 - z_i z_j$  is real stable. Obviously, if  $\text{Im}(z_i), \text{Im}(z_j) > 0$ , then  $z_i z_j \neq 1$ . So,  $1 - z_i z_j$  is real stable.  $\square$

Brändén [Brä07] proved the following characterization of multilinear stable polynomials.

**Theorem 2.7** (Brändén [Brä07, Theorem 5.6]). *For any multilinear polynomial  $p(y_1, \dots, y_m)$ ,  $p$  is stable if and only if for all  $1 \leq i < j \leq m$ ,*

$$\partial_{y_i} p|_{y=x} \cdot \partial_{y_j} p|_{y=x} \geq p(x) \cdot \partial_{y_i} \partial_{y_j} p|_{y=x},$$

for all  $x \in \mathbb{R}^m$ .

## 2.2 Polarization

Let  $p \in \mathbb{C}[z_1, \dots, z_n]$  be a stable polynomial of degree  $d_i$  in  $z_i$ ,  $1 \leq i \leq n$ . The polarization is an operator that turns  $p$  into a multilinear “symmetric” polynomial.

For any  $m \geq \max_i d_i$ , The polarization  $\pi_m(p)$  of  $p$  is a polynomial in  $m \cdot n$  variables  $z_{i,j}$  for  $1 \leq i \leq n, 1 \leq j \leq m$  defined as follows: For all  $1 \leq i \leq n$ , and  $1 \leq j \leq d_i$  substitute any occurrence of  $z_i^k$  in  $p$  with  $\frac{1}{\binom{m}{k}} e_k(z_{i,1}, \dots, z_{i,m})$ .

For example, if  $p(z_1, z_2) = z_1^2 + z_1 z_2$ , then

$$\pi_2(p) = z_{1,1} z_{1,2} + \frac{1}{2}(z_{1,1} + z_{1,2}) \cdot \frac{1}{2}(z_{2,1} + z_{2,2}).$$

Note that for any  $m$  and any polynomial  $p$ ,  $\pi_m(p)$  has the following properties:

- $\pi_m(p)$  is multilinear;
- $\pi_m(p)$  is symmetric in the variables  $z_{i,1}, \dots, z_{i,m}$  for all  $1 \leq i \leq n$ ;
- if we substitute  $z_{i,j} = z_i$  for all  $i, j$  in  $\pi_m(p)$  we recover  $p$ .

Borcea, Brändén, and Liggett used Grace-Walsh-Szego’s coincidence theorem [Gra02, Sze22, Wal22] to prove that stability is preserved under polarization [BBL09, Thm 4.7].

**Theorem 2.8** (Borcea, Brändén, Liggett [BBL09]). *If  $p \in \mathbb{C}[z_1, \dots, z_n]$  is stable then  $\pi_m(p)$  is stable for all  $m$ .*

### 2.3 Stability and Convexity

Polynomial optimization problems involving real stable polynomials with nonnegative coefficients can often be turned into concave/convex programs. Such polynomials are log-concave in the positive orthant:

**Theorem 2.9** ([Gül97]). *For any real stable homogeneous polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  such that  $p(\mathbf{z}) > 0$  for all  $\mathbf{z} \in \mathbb{R}_{++}^n$ ,  $\log p(\mathbf{z})$  is concave in the positive orthant.*

*Proof.* Firstly, note that  $\log p(\mathbf{z})$  is well-defined for any  $\mathbf{z} \in \mathbb{R}_{++}^n$ . It is enough to show that the function is concave along any interval in the positive orthant. Let  $\mathbf{a} \in \mathbb{R}_{++}^n$ ,  $\mathbf{b} \in \mathbb{R}^n$ , and consider the line  $\mathbf{a} + t\mathbf{b}$  where for any  $t \in [0, 1]$ ,  $\mathbf{a} + t\mathbf{b} \in \mathbb{R}_{++}^n$ . We show that  $\log p(\mathbf{a} + t\mathbf{b})$  is concave. Say  $p$  is  $k$ -homogeneous, then

$$p(\mathbf{a} + t\mathbf{b}) = p(t(\mathbf{a}/t + \mathbf{b})) = t^k p(\mathbf{a}/t + \mathbf{b}).$$

We claim that  $p(\mathbf{a}/t + \mathbf{b})$  is real rooted. Firstly, since  $\mathbf{a} \in \mathbb{R}_{++}^n$ , and  $p(\cdot)$  is stable,  $p(\mathbf{a}t + \mathbf{b})$  is real rooted. But for any root  $\lambda$  of  $p(\mathbf{a}t + \mathbf{b})$ ,  $1/\lambda$  is a root of  $p(\mathbf{a}/t + \mathbf{b})$ . Note that  $p(\mathbf{a}t + \mathbf{b})$  has no roots at 0, because  $\mathbf{a}t + \mathbf{b} \in \mathbb{R}_{++}^n$  for  $t \in [0, 1]$ . So,  $p(\mathbf{a}/t + \mathbf{b})$  is real rooted.

Let  $\lambda_1, \dots, \lambda_k$  be the roots of  $p(\mathbf{a}/t + \mathbf{b})$ . We have

$$p(\mathbf{a} + t\mathbf{b}) = t^k p(\mathbf{a}/t + \mathbf{b}) = t^k p(\mathbf{a}) \prod_{i=1}^k (1/t - \lambda_i) = p(\mathbf{a}) \prod_{i=1}^k (1 - t\lambda_i).$$

We claim that for all  $1 \leq i \leq k$ ,  $\lambda_i < 1$ . Otherwise, for some  $t \in [0, 1]$ ,  $p(\mathbf{a} + t\mathbf{b}) = 0$ , but since  $\mathbf{a} + t\mathbf{b} \in \mathbb{R}_{++}^n$ ,  $p(\mathbf{a} + t\mathbf{b}) > 0$  which is a contradiction. Therefore,

$$\log p(\mathbf{a} + t\mathbf{b}) = \log p(\mathbf{a}) + \sum_{i=1}^k \log(1 - t\lambda_i).$$

The theorem follows by the fact that  $\log(1 - t\lambda)$  is a concave function of  $t$  for  $t \in [0, 1]$  when  $\lambda < 1$ .  $\square$

**Corollary 2.10.** *For any real stable polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients,  $\log p(\mathbf{z})$  is concave in positive orthant.*

*Proof.* The homogenization of a polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  is defined as the polynomial  $p_H \in \mathbb{R}[z_1, \dots, z_n, z_{n+1}]$  given by

$$p_H(z_1, \dots, z_{n+1}) = z_{n+1}^{\deg p} p(z_1/z_{n+1}, \dots, z_n/z_{n+1}).$$

Borcea, Brändén and Liggett [BBL09, Thm 4.5] show that for a polynomial  $p$  with nonnegative coefficients,  $p$  is real stable if and only if  $p_H$  is real stable. Since  $p_H$  is real stable, by [Theorem 2.9](#),  $\log p_H(z_1, \dots, z_{n+1})$  is concave in the positive orthant. It follows that  $p_H(z_1, \dots, z_n, 1)$  is also concave in the positive orthant. But,  $p(z_1, \dots, z_n) = p_H(z_1, \dots, z_n, 1)$ . So,  $p(\mathbf{z})$  is concave in the positive orthant.  $\square$

It is also an immediate corollary of Hölder's inequality that a polynomial with nonnegative coefficients is log-convex in terms of the log of its variables (for more details on log-convex functions see [BV06]).

**Fact 2.11.** For any polynomial  $p(y_1, \dots, y_n)$  with nonnegative coefficients, the expression

$$\log p(\exp(\mathbf{y})) = \log p(e^{y_1}, \dots, e^{y_n})$$

is convex as a function of  $\mathbf{y}$ .

**Lemma 2.12.** For any polynomial  $p$  with nonnegative coefficients, and any  $\alpha \geq 0$ ,

$$\inf_{\mathbf{z} > 0} \frac{p(\mathbf{z})}{(\mathbf{z}/\alpha)^\alpha} = \inf_{\mathbf{z} > 0} \frac{p(\alpha\mathbf{z})}{\mathbf{z}^\alpha}.$$

*Proof.* If  $\alpha > 0$ , the two sides become equal by a change of variable  $\mathbf{z} \leftrightarrow \alpha\mathbf{z}$ . If  $\alpha_i = 0$  for some  $i$ , since  $p$  has nonnegative coefficients, it is obvious that one should let  $z_i \rightarrow 0$  on the LHS, since the numerator decreases monotonically as  $z_i \rightarrow 0$ , and the denominator does not change. In the limit the LHS is the same as what we have on the RHS.  $\square$

## 2.4 Jump Systems and Newton Polytope

**Definition 2.13.** For a polynomial  $p(y_1, \dots, y_n)$ , we define  $\text{supp}(p)$  as

$$\text{supp}(p) := \{\boldsymbol{\kappa} \in \mathbb{Z}_+^n : C_p(\boldsymbol{\kappa}) \neq 0\}.$$

Brändén [Brä07] proved that if  $p$  is a real stable polynomial then  $\text{supp}(p)$  forms what is known as a *jump system*.

**Definition 2.14.** For two points  $\mathbf{x}, \mathbf{y} \in \mathbb{Z}^n$  define an  $(\mathbf{x}, \mathbf{y})$ -step to be any  $\boldsymbol{\delta} \in \mathbb{Z}^n$  such that  $\|\boldsymbol{\delta}\|_1 = 1$  and  $\|\mathbf{y} - (\mathbf{x} + \boldsymbol{\delta})\|_1 = \|\mathbf{y} - \mathbf{x}\|_1 - 1$ . A set  $\mathcal{F} \subseteq \mathbb{Z}^n$  is called a *jump system* iff for every  $\mathbf{x}, \mathbf{y} \in \mathcal{F}$  and  $(\mathbf{x}, \mathbf{y})$ -step  $\boldsymbol{\delta}$ , either  $\mathbf{x} + \boldsymbol{\delta} \in \mathcal{F}$  or there exists an  $(\mathbf{x} + \boldsymbol{\delta}, \mathbf{y})$ -step  $\boldsymbol{\delta}'$  such that  $\mathbf{x} + \boldsymbol{\delta} + \boldsymbol{\delta}' \in \mathcal{F}$ .

Brändén [Brä07] proved that for any real stable polynomial  $p$ ,  $\text{supp}(p)$  is a jump system. We use the following immediate corollary which follows from the fact that for any polynomial  $p \in \mathbb{R}[y_1, \dots, y_n, z_1, \dots, z_n]$ ,  $\text{supp}(p(\mathbf{y}, \mathbf{z})) = \text{supp}(p(\mathbf{y}, -\mathbf{z}))$ .

**Fact 2.15.** If the polynomial  $p(\mathbf{y}, \mathbf{z})$  is bistable, then  $\text{supp}(p)$  is a jump system.

Lovász showed that the natural generalization of the greedy algorithm for maximizing linear functions over matroids, also works for jump systems [Lov97]. Given a finite jump system  $\mathcal{F} \subseteq \mathbb{Z}^n$ , and  $\mathbf{w} \in \mathbb{R}^n$ , the following algorithm finds  $\max_{\mathbf{x} \in \mathcal{F}} \langle \mathbf{w}, \mathbf{x} \rangle$ : Sort the coordinates in such a way that  $|w_1| \geq |w_2| \geq \dots \geq |w_n|$ . Fix the coordinates of  $\mathbf{x}$ , one by one; for  $i = 1, \dots, n$ , we will fix  $x_i$  as follows: among all members of  $\mathcal{F}$  whose first  $i-1$  coordinates match  $x_1, \dots, x_{i-1}$ , let  $x_i$  be largest  $i$ -th coordinate if  $w_i > 0$  and let  $x_i$  be the smallest otherwise.

**Theorem 2.16** ([Lov97]). For any finite jump system  $\mathcal{F} \subseteq \mathbb{Z}^n$  and  $\mathbf{w} \in \mathbb{R}^n$ , the above greedy algorithm correctly finds a point  $\mathbf{x}$  maximizing  $\langle \mathbf{w}, \mathbf{x} \rangle$ .

We will be working with the convex hull of  $\text{supp}(p)$ , often known as the Newton polytope.

**Definition 2.17.** For a polynomial  $p(y_1, \dots, y_n)$ , define the *Newton polytope* of  $p$  as

$$\text{newt}(p) := \text{conv}(\text{supp}(p)) = \text{conv}\{\boldsymbol{\kappa} \in \mathbb{Z}_+^n : C_p(\boldsymbol{\kappa}) \neq 0\},$$

where  $\text{conv}$  denotes convex hull.



We use the following fact in multiple places in the proofs.

**Fact 2.18.** For a polynomial  $p(\mathbf{y}) \in \mathbb{R}[y_1, \dots, y_n]$  with nonnegative coefficients, and  $\boldsymbol{\alpha} \in \mathbb{R}_+^n$  we have

$$\inf_{\mathbf{y} > 0} \frac{p(\mathbf{y})}{\mathbf{y}^\alpha} > 0 \iff \boldsymbol{\alpha} \in \text{newt}(p).$$

*Proof.* Let  $\boldsymbol{\alpha} \in \text{newt}(p)$ ; then we can write  $\boldsymbol{\alpha}$  as a convex combination of the vertices:  $\boldsymbol{\alpha} = \sum_i \lambda_i \boldsymbol{\kappa}_i$ , where  $\lambda_i$ 's are nonnegative and sum up to 1, and for all  $i$ ,  $C_p(\boldsymbol{\kappa}_i) > 0$ . Therefore,

$$p(\mathbf{y}) \geq \sum_i C_p(\boldsymbol{\kappa}_i) \mathbf{y}^{\boldsymbol{\kappa}_i} = \sum_i \lambda_i \frac{C_p(\boldsymbol{\kappa}_i) \mathbf{y}^{\boldsymbol{\kappa}_i}}{\lambda_i} \geq \prod_i \left( \frac{C_p(\boldsymbol{\kappa}_i) \mathbf{y}^{\boldsymbol{\kappa}_i}}{\lambda_i} \right)^{\lambda_i} = \mathbf{y}^\alpha \prod_i \left( \frac{C_p(\boldsymbol{\kappa}_i)}{\lambda_i} \right)^{\lambda_i},$$

where for the last inequality we used weighted AM-GM. This implies that

$$\inf_{\mathbf{y} > 0} \frac{p(\mathbf{y})}{\mathbf{y}^\alpha} \geq \prod_i \left( \frac{C_p(\boldsymbol{\kappa}_i)}{\lambda_i} \right)^{\lambda_i} > 0.$$

Conversely, suppose  $\boldsymbol{\alpha} \notin \text{newt}(p)$ ; we show that  $\inf_{\mathbf{y} > 0} p(\mathbf{y})/\mathbf{y}^\alpha = 0$ . Since  $\text{newt}(p)$  is convex, by separating hyperplane theorem, there exists  $\mathbf{c} \in \mathbb{R}^n$ , such that  $\langle \mathbf{c}, \boldsymbol{\alpha} \rangle > \langle \mathbf{c}, \boldsymbol{\kappa} \rangle$  for every  $\boldsymbol{\kappa} \in \text{newt}(p)$ . Now let  $\mathbf{y} = \exp(t\mathbf{c}) = (e^{tc_1}, \dots, e^{tc_n})$ , where  $t > 0$  will be fixed later. Then we have

$$\frac{p(\mathbf{y})}{\mathbf{y}^\alpha} = \frac{\sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} C_p(\boldsymbol{\kappa}) \mathbf{y}^\boldsymbol{\kappa}}{\mathbf{y}^\alpha} \leq \exp \left( -t \left( \langle \mathbf{c}, \boldsymbol{\alpha} \rangle - \max_{\boldsymbol{\kappa} \in \text{newt}(p)} \langle \mathbf{c}, \boldsymbol{\kappa} \rangle \right) \right) \cdot \sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} C_p(\boldsymbol{\kappa}).$$

It is easy to see that as  $t \rightarrow \infty$ , the RHS of the above inequality becomes zero.  $\square$

The convex hull of jump systems has been shown to be bisubmodular polyhedra in [BC95]. The following is a restatement of this result.

**Theorem 2.19** (Restatement of [BC95]). For a finite jump system  $\mathcal{F} \subseteq \mathbb{Z}^n$ , define the function  $f : \{-1, 0, +1\}^n \rightarrow \mathbb{R}$  as

$$f(\mathbf{c}) := \max_{\boldsymbol{\alpha} \in \mathcal{F}} \langle \mathbf{c}, \boldsymbol{\alpha} \rangle.$$

Then

$$\text{conv}(\mathcal{F}) = \bigcap_{\mathbf{c} \in \{-1, 0, +1\}^n} \{ \mathbf{x} \in \mathbb{R}^n : \langle \mathbf{c}, \mathbf{x} \rangle \leq f(\mathbf{c}) \}.$$

Furthermore,  $f$  is a bisubmodular function.

For the definition of bisubmodular functions, see [BC95]. We will only use the following well-known fact about bisubmodular functions. For more details see [MF08].

**Fact 2.20.** Suppose that  $\mathcal{F}$  is a jump system and  $f$  is the bisubmodular function associated with  $\text{conv}(\mathcal{F})$ . Given an oracle that returns  $f(\mathbf{c})$  for any  $\mathbf{c} \in \{-1, 0, +1\}^n$ , we can construct a separation oracle for  $\text{conv}(\mathcal{F})$ . Each call to the separation oracle uses at most  $\text{poly}(n)$  evaluations of  $f$ .

*Proof.* To construct a separation oracle, it is enough to design a procedure that given a point  $\mathbf{x} \in \mathbb{R}^n$ , finds a  $\mathbf{c} \in \{-1, 0, +1\}^n$  such that  $\langle \mathbf{c}, \mathbf{x} \rangle > f(\mathbf{c})$  or declares that there is no such  $\mathbf{c}$ .

The function  $\mathbf{c} \mapsto f(\mathbf{c}) - \langle \mathbf{c}, \mathbf{x} \rangle$  is bisubmodular, since  $f$  is bisubmodular and  $\mathbf{c} \mapsto \langle \mathbf{c}, \mathbf{x} \rangle$  is modular. Bisubmodular functions can be minimized in strongly polynomial time, see e.g. [MF08]. Therefore one can solve the following optimization problem using polynomially many evaluations of  $f$ .

$$\min_{\mathbf{c} \in \{-1, 0, +1\}^n} f(\mathbf{c}) - \langle \mathbf{c}, \mathbf{x} \rangle.$$

If the answer is nonnegative, then  $x \in \text{conv}(\mathcal{F})$ . Otherwise, the optimizing  $\mathbf{c}$  produces the separating hyperplane.  $\square$

## 3 Applications

### 3.1 Applications in Counting

In this section we prove [Theorem 1.2](#) and we discuss several applications of our results for counting problems.

We start with Schrijver's inequality. The following theorem is the main technical result of [Sch98].

**Theorem 3.1** (Schrijver [Sch98]). *Let  $A \in \mathbb{R}^{n \times n}$  be a doubly stochastic matrix, and let  $\tilde{A}$  be defined as follows:*

$$\tilde{A}_{i,j} = A_{i,j}(1 - A_{i,j})$$

*Then,*

$$\text{per}(\tilde{A}) \geq \prod_{i,j} (1 - A_{i,j}).$$

Schrijver [Sch98] used the above inequality to prove that any  $d$ -regular bipartite graph with  $2n$  vertices has  $\left(\frac{(d-1)^{d-1}}{d^{d-2}}\right)^n$  matchings. Gurvits and [Gur11, GS14] used the above theorem to design a deterministic  $2^n$ -approximation algorithm for the permanent. We note that Gurvits used the machinery of stable polynomials to prove several lower bounds on permanent, but to the best of our knowledge his machinery could not capture Schrijver's theorem. Very recently, Lelarge [Lel17] used the machinery developed by Csikvári [Csi14] to reprove Schrijver's theorem.

Here, we give a simple proof of this theorem using [Theorem 1.1](#).

*Proof.* Let  $\mathbf{z} = \{z_{i,j}\}_{1 \leq i,j \leq n}$  be a vector of  $n^2$  variables. Define,

$$\begin{aligned} p(\mathbf{z}) &= \prod_{i=1}^n \sum_{j=1}^n A_{i,j} z_{i,j}, \\ q(\mathbf{z}) &= \prod_{j=1}^n \sum_{i=1}^n (1 - A_{i,j}) z_{i,j}. \end{aligned}$$

Then, it is not hard to see that

$$q(\partial \mathbf{z}) p(\mathbf{z}) = \text{per}(\tilde{A}).$$

This follows because each monomial of  $p$  is a product  $\prod_{i=1}^n z_{i,\sigma(i)}$  for some mapping  $\sigma : [n] \rightarrow [n]$  and each monomial of  $q$  is a product  $\prod_{j=1}^n z_{\sigma(j),j}$  for some  $\sigma : [n] \rightarrow [n]$ .

Let  $\alpha_{i,j} = A_{i,j}$  for all  $1 \leq i, j \leq n$ . Since  $p, q$  are multilinear, by [Theorem 1.1](#),

$$\text{per}(\tilde{A}) \geq (1 - \alpha)^{1-\alpha} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}. \quad (5)$$

Firstly, observe that by weighted AM-GM inequality

$$p(\mathbf{y}) = \prod_{i=1}^n \sum_{j=1}^n A_{i,j} y_{i,j} \geq \prod_{i,j} y_{i,j}^{A_{i,j}} = \mathbf{y}^\alpha,$$

where in the inequality we used that  $A$  is doubly stochastic. Secondly, by another application of AM-GM inequality we have

$$q(\mathbf{z}) = \prod_{j=1}^n \sum_{i=1}^n A_{i,j} \frac{(1 - A_{i,j})z_{i,j}}{A_{i,j}} \geq \prod_{i,j} \left( \frac{(1 - A_{i,j})z_{i,j}}{A_{i,j}} \right)^{A_{i,j}} = ((1 - \alpha)\mathbf{z}/\alpha)^\alpha,$$

where in the inequality we used that  $A$  is doubly stochastic. Plugging the above two inequalities in (5) we get

$$\text{per}(\tilde{A}) \geq (1 - \alpha)^{1-\alpha} (1 - \alpha)^\alpha = \prod_{i,j} (1 - A_{i,j})$$

as desired.  $\square$

**Theorem 1.2.** *Given two real stable polynomials  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients and an oracle that for any  $\mathbf{z} \in \mathbb{R}^n$ , returns  $p(\mathbf{z}), q(\mathbf{z})$ , there is a polynomial time algorithm that returns an  $e^{\min\{\deg p, \deg q\} + \epsilon}$  approximation of  $q(\partial\mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0}$  that runs in time  $\text{poly}(\langle p \rangle + \langle q \rangle + \log(1/\epsilon))$ .*

*Proof.* Firstly, we show that

$$e^{-\min\{\deg p, \deg q\}} \sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \leq q(\partial\mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0} \leq \sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \quad (6)$$

Without loss of generality assume that  $\deg p \leq \deg q$ . By [Theorem 1.1](#), it is enough to show that for any  $\alpha \geq 0$ ,

$$e^{-\deg p} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \leq \inf_{\mathbf{y}, \mathbf{z} > 0} e^{-\alpha} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}.$$

To prove the above, it is enough to show that if  $\sum_i \alpha_i > \deg p$ , then both sides of the above equal 0. Suppose  $\langle \alpha, \mathbf{1} \rangle > \deg p$ . Let  $\mathbf{z} = \mathbf{1}$  and  $\mathbf{y} = t\mathbf{1}$  for some parameter  $t$ . Then, as  $t \rightarrow \infty$ ,  $p(\mathbf{y})/\mathbf{y}^\alpha \rightarrow 0$ . This is because the degree of every monomial of  $p(\mathbf{y})$  is less than  $\mathbf{y}^\alpha$ . This was a special case of [Fact 2.18](#). In general if  $\alpha \notin \text{newt}(p) \cap \text{newt}(q)$ , then both sides of the above are zero.

Using (6), to prove the theorem it is enough to design a polynomial time algorithm that computes  $\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}$ . Firstly, by [Lemma 2.12](#) it is enough to solve  $\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\alpha\mathbf{z})}{(\mathbf{y}\mathbf{z})^\alpha}$ . Now, let us do a change of variable  $\mathbf{y} \leftrightarrow \exp(\mathbf{y})$  and  $\mathbf{z} \leftrightarrow \exp(\mathbf{z})$ . By [Fact 2.11](#) and [Corollary 2.10](#),

$$\log \frac{p(\exp(\mathbf{y}))q(\alpha \exp(\mathbf{z}))}{(\exp(\mathbf{y}) \exp(\mathbf{z}))^\alpha}$$

is concave in  $\alpha$  and convex in  $\mathbf{y}, \mathbf{z}$ . Since the infimum of concave functions is also concave,  $\inf_{\mathbf{y}, \mathbf{z}} \log \frac{p(\exp(\mathbf{y}))q(\alpha \exp(\mathbf{z}))}{(\exp(\mathbf{y}) \exp(\mathbf{z}))^\alpha}$  is a concave function of  $\alpha$ . Therefore,

$$\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\exp(\mathbf{y}))q(\exp(\mathbf{z}))}{(\exp(\mathbf{y}) \exp(\mathbf{z}))^\alpha}.$$

is a convex program, or more precisely a concave-convex saddle point problem. In [Subsection 3.3](#) we show how this convex program can be approximately solved in polynomial time. This gives an  $e^{\deg p}$  approximation of  $q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0}$  as desired.  $\square$

Let us state some applications of the above result. Our first application is in the context of recent results on Nash welfare maximization (NWM) problems. Recently the authors together with Saberi and Singh proved a generalization of Gurvits's inequality (2) to design an  $e^n$  approximation for NWM problem [[AOSS17](#)]. Given a real stable polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients and an integer  $1 \leq D \leq n$ , the main technical contribution of [[AOSS17](#)] is a polynomial time algorithm that gives an  $e^D$  approximation to

$$\left( \sum_{S \in \binom{[n]}{D}} \prod_{i \in S} \partial_{z_i} \right) p(\mathbf{z})|_{\mathbf{z}=0}.$$

This result immediately follows from [Theorem 1.2](#). Let  $q$  be the  $D$ -th elementary symmetric polynomial on  $z_1, \dots, z_n$ , i.e.,  $q(\mathbf{z}) = \sum_{S \in \binom{[n]}{D}} \prod_{i \in S} \partial_{z_i}$ . Recall that all elementary symmetric polynomials are real stable. Since  $\deg q = D$ , we can approximate the above quantity up to an  $e^D$  factor by [Theorem 1.2](#).

Our second application is a partial answer to an open problem of Kulesza and Taskar about determinantal point processes [[KT12](#)]. A determinantal point process (DPP) is a probability distribution  $\mu$  on subsets of  $[n]$  identified by a positive semidefinite ensemble matrix  $L \in \mathbb{R}^{n \times n}$  where for any  $S \subseteq [n]$  we have

$$\mathbb{P}[S] \propto \det(L_{S,S}),$$

where  $L_{S,S}$  is the square submatrix of  $L$  corresponding to rows and columns indexed by  $S$ . We refer interested readers to [[KT12](#)] for applications of DPPs in machine learning. For an integer  $0 \leq k \leq n$ , and a DPP  $\mu$ , the truncation of  $\mu$  to  $k$ ,  $\mu_k$  is called a  $k$ -DPP, i.e.,

$$\mu_k(S) \propto \begin{cases} \mu(S) & \text{if } |S| = k \\ 0 & \text{otherwise.} \end{cases}$$

Kulesza and Taskar left an open problem to estimate the quantity

$$\sum_{S \subseteq [n]} \det(L_{S,S})^p$$

for some  $p > 1$ . Their motivation is to study the Hellinger distance between two DPPs.

Borcea, Brändén and Liggett [[BBL09](#)] showed that DPPs and  $k$ -DPPs are strongly Rayleigh, i.e., if  $\mu$  is a DPP, then  $g_\mu(z_1, \dots, z_n)$  and  $g_{\mu_k}(z_1, \dots, z_n)$  are real stable polynomials (see [Subsection 2.1](#)). Let  $L, L'$  be ensemble matrices of two  $k$ -DPPs  $\mu, \mu'$ . Let  $p = g_{\mu_k}$  and  $q = g_{\mu'_k}$ . Note that  $p, q$  are

$k$ -homogeneous, multilinear, and have nonnegative coefficients. Therefore, by the above theorem we can estimate

$$q(\partial \mathbf{z})p(\mathbf{z})|_{\mathbf{z}=0} = \sum_{S \subseteq [n]} \det(L_S) \det(L'_S)$$

up to a factor of  $e^k$ . To the best of our knowledge, to this date no approximation algorithm for the above quantity was known.

### 3.2 Applications in Optimization

In this section we prove [Theorem 1.3](#). Let us restate the theorem.

**Theorem 1.3.** *Given two real stable multilinear  $D$ -homogeneous polynomials  $p, q \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients and an oracle that for any  $\mathbf{z} \in \mathbb{R}^n$  returns  $p(\mathbf{z}), q(\mathbf{z})$ , there is a polynomial time algorithm that outputs an  $e^{2D+\epsilon}$  approximation to*

$$\max_{\boldsymbol{\kappa} \in \{0,1\}^n} C_p(\boldsymbol{\kappa})C_q(\boldsymbol{\kappa}).$$

that runs in time  $\text{poly}(\langle p \rangle + \langle q \rangle + \log(1/\epsilon))$ .

*Proof.* We solve the following mathematical program and we show that its optimum solution gives an  $e^{2D}$ -approximation of the optimum.

$$\begin{aligned} \sup_{\boldsymbol{\alpha}, \boldsymbol{\lambda} \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\boldsymbol{\lambda} \mathbf{y})q(\mathbf{z})}{(\mathbf{y} \mathbf{z} / \boldsymbol{\alpha})^\alpha} \\ \text{s.t.} \quad \sum_i \lambda_i = D. \end{aligned} \tag{7}$$

Firstly, we show that we can solve the above mathematical program. The idea is very similar to the proof [Theorem 1.2](#). We do a change of variable  $\mathbf{y} \leftrightarrow \exp(\mathbf{y})$ ,  $\mathbf{z} \leftrightarrow \exp(\mathbf{z})$  and we push the  $\alpha$  in the denominator to  $q$ . So, instead we solve the following program:

$$\sup_{\boldsymbol{\alpha}, \boldsymbol{\lambda} \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \log \frac{p(\boldsymbol{\lambda} \exp(\mathbf{y}))q(\boldsymbol{\alpha} \exp(\mathbf{z}))}{(\exp(\mathbf{y}) \exp(\mathbf{z}))^\alpha}$$

As before,  $\log \frac{p(\boldsymbol{\lambda} \exp(\mathbf{y}))q(\boldsymbol{\alpha} \exp(\mathbf{z}))}{(\exp(\mathbf{y}) \exp(\mathbf{z}))^\alpha}$  is convex in  $\mathbf{y}, \mathbf{z}$  and concave in  $\boldsymbol{\alpha}, \boldsymbol{\lambda}$ . So, we can solve the above program in polynomial time; for more details see [Subsection 3.3](#).

Secondly, we show that the above program is a relaxation of the problem. Let

$$\boldsymbol{\kappa}^* = \operatorname{argmax}_{\boldsymbol{\kappa} \in \{0,1\}^n} C_p(\boldsymbol{\kappa})C_q(\boldsymbol{\kappa}).$$

Since every monomial of  $p$  has degree exactly  $D$ ,  $\sum_i \kappa_i^* = D$ , so  $\boldsymbol{\lambda} = \boldsymbol{\kappa}^*$  is a feasible solution. Let  $\boldsymbol{\alpha} = \boldsymbol{\kappa}^*$ . Then,

$$\inf_{\mathbf{y} > 0} \frac{p(\boldsymbol{\lambda} \mathbf{y})}{\mathbf{y}^\alpha} \geq \inf_{\mathbf{y} > 0} \frac{C_p(\boldsymbol{\kappa}^*) \mathbf{y}^{\boldsymbol{\kappa}^*}}{\mathbf{y}^{\boldsymbol{\kappa}^*}} = C_p(\boldsymbol{\kappa}^*)$$

Similarly,

$$\inf_{\mathbf{z} > 0} \frac{q(\mathbf{z})}{(\mathbf{z} / \boldsymbol{\alpha})^\alpha} = \inf_{\mathbf{z} > 0} \frac{q(\boldsymbol{\alpha} \mathbf{z})}{\mathbf{z}^\alpha} = \inf_{\mathbf{z} > 0} \frac{q(\boldsymbol{\kappa}^* \mathbf{z})}{\mathbf{z}^{\boldsymbol{\kappa}^*}} = C_q(\boldsymbol{\kappa}^*),$$

The first equality follows by [Lemma 2.12](#). Using the above two equations it follows that the optimum of (7) is at least  $C_p(\boldsymbol{\kappa}^*)C_q(\boldsymbol{\kappa}^*)$ .

Next, we show that the optimum of (7) is within  $e^{2D}$  factor of  $C_p(\boldsymbol{\kappa}^*)C_q(\boldsymbol{\kappa}^*)$ . Let  $\lambda^*, \alpha^*$  be the arguments achieving optimum value in the sup in (7)<sup>1</sup>. Consider the following probability distribution on  $\boldsymbol{\kappa} \in \{0, 1\}^n$ . We sample  $D$  numbers  $X_1, X_2, \dots, X_D$  independently (with replacement) where for each  $1 \leq i \leq D$  we let  $X_i = j$  with probability  $\frac{\lambda_j^*}{D}$ . We let  $\kappa_i = 1$ , if  $X_j = i$  for some  $i$  and  $\kappa_i = 0$  otherwise. We claim that

$$\mathbb{E}[C_p(\boldsymbol{\kappa})C_q(\boldsymbol{\kappa})] \geq e^{-2D} \text{OPT}. \quad (8)$$

Define

$$\tilde{p}(\mathbf{z}) = p(\boldsymbol{\lambda}^* \mathbf{z}).$$

Then, by [Fact 2.5](#),  $\tilde{p}$  is real stable. Therefore, by [Theorem 1.1](#),

$$\sum_{\boldsymbol{\kappa} \in \{0,1\}^n} C_{\tilde{p}}(\boldsymbol{\kappa})C_q(\boldsymbol{\kappa}) \geq e^{-\alpha^*} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{\tilde{p}(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha^*)^{\alpha^*}} \geq e^{-D} \sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\boldsymbol{\lambda}^* \mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^{\alpha}} \geq e^{-D} \text{OPT}.$$

The second to last inequality follows by (6) and the last inequality follows by the fact that (7) is a relaxation of OPT.

So, to prove (8) it is enough to show that

$$\mathbb{E}[C_p(\boldsymbol{\kappa})C_q(\boldsymbol{\kappa})] \geq e^{-D} \sum_{\boldsymbol{\kappa} \in \{0,1\}^n} C_{\tilde{p}}(\boldsymbol{\kappa})C_q(\boldsymbol{\kappa}). \quad (9)$$

Fix any monomial  $\boldsymbol{\kappa}$  such that  $C_p(\boldsymbol{\kappa}), C_q(\boldsymbol{\kappa}) > 0$ . Then,

$$C_{\tilde{p}}(\boldsymbol{\kappa}) = C_p(\boldsymbol{\kappa}) \prod_{i:\kappa_i=1} \lambda_i^*.$$

On the other hand,

$$\mathbb{P}[\boldsymbol{\kappa}] = D! \prod_{i:\kappa_i=1} \frac{\lambda_i^*}{D} \geq e^{-D} \prod_{i:\kappa_i=1} \lambda_i^*,$$

The  $D!$  factor comes from the fact that we can sample each such  $\boldsymbol{\kappa}$  in  $D!$  possible ways. Plugging these two equations in (9) proves it. This completes the proof of [Theorem 1.3](#).  $\square$

One of the consequences of the above theorem is a recent result of Nikolov and Singh [[NS16](#)]. Given a PSD matrix  $L$  and a partitioning  $P_1, \dots, P_m$  of  $[n]$ , among all sets  $S \subseteq [n]$  where for each  $1 \leq i \leq m$ ,  $|S \cap P_i| = b_i$  we want to choose one that maximizes  $\det(L_{S,S})$  where  $L_{S,S}$  is the square submatrix of  $L$  with rows and columns indexed by  $S$ . We refer interested readers to [[NS16](#)] for applications of this problem.

Let  $k = \sum_i b_i$ . Let  $v_1, \dots, v_n$  be the vectors of the Cholesky decomposition of  $L$ , i.e., for all  $i, j$ ,  $L_{i,j} = \langle v_i, v_j \rangle$ . Let

$$p(y_1, \dots, y_n) = \partial_x^{n-k} \det(xI + y_1 v_1 v_1^T + \dots + y_n v_n v_n^T)|_{x=0}.$$

---

<sup>1</sup>Note that since we can assume the domain of  $\boldsymbol{\alpha}, \boldsymbol{\lambda}$  is compact, the supremum is attained at some point of the domain

By [Fact 2.1](#) and [Fact 2.4](#) the above polynomial is real-stable. Furthermore, it is not hard to see that it is  $k$ -homogeneous and has nonnegative coefficients. In fact for any set  $S \subseteq [n]$  of size  $k$ , the coefficient of the monomial  $\mathbf{z}^{\mathbf{1}_S}$  is  $\det(L_{S,S})$ , where  $\mathbf{1}_S$  is the indicator vector of the set  $S$ . Now, let

$$q(z_1, \dots, z_n) = \prod_{i=1}^m e_{b_i}(\{z_j\}_{j \in P_i})$$

be the generating polynomial corresponding to the partition matroid  $P_1, \dots, P_k$ . Again, by closure of stable polynomials under product, the above polynomial is real stable, multilinear,  $k$ -homogeneous, and has nonnegative coefficients. It follows by the above theorem that we can approximate

$$\max_{S: |S \cap P_i| = b_i \forall i} \det(L_{S,S})$$

within an  $e^{2k}$  factor. Note that our approximation factor is slightly worse but our theorem is significantly more general. In particular, we can replace  $p, q$  in the above argument with any real stable multilinear polynomial.

For example, suppose we are given a graph  $G = (V, E)$  and we have assigned a vector  $v_e$  to any edge  $e$ . We can let  $q$  be the generating polynomial of the uniform spanning tree distribution of  $G$ . Then, using the above theorem we can approximate

$$\max_T \det \left( \sum_{e \in T} v_e v_e^T \right)$$

within an  $e^{2n}$  factor.

### 3.3 Solving the Convex Program

In this section we show how the convex programs from the previous two sections can be approximately solved in polynomial time. We measure the complexity of our algorithm in terms of the complexity of the involved polynomials. Recall the following definition.

**Definition 3.2.** For a polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients, we define its complexity as  $n + \deg p + |\log \min_{\kappa: C_p(\kappa) \neq 0} C_p(\kappa)| + |\log \max_{\kappa: C_p(\kappa) \neq 0} C_p(\kappa)|$ , and we denote this by  $\langle p \rangle$ .

The main result of this section is the following.

**Theorem 3.3.** Given a bistable polynomial  $p \in \mathbb{R}[y_1, \dots, y_n, z_1, \dots, z_n]$  with nonnegative coefficients and an oracle that evaluates  $p$  at any requested point  $(\mathbf{y}, \mathbf{z}) \in \mathbb{R}_+^{2n}$ , one can find a  $1 + \epsilon$  approximation of

$$\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}, \tag{10}$$

in time  $\text{poly}(\langle p \rangle + \log(1/\epsilon))$ .

In the rest of this section, we prove this theorem. We will use the ellipsoid method to compute the inner inf and also the outer sup. The main difficulty for computing the inf is that the variables  $\mathbf{y}, \mathbf{z}$  are unbounded. In other words, we do not have a bounded outer ellipse for the ellipsoid method. In the first step, we use the fact that  $\text{supp}(p)$  is a jump system ([Fact 2.15](#)) to show that we can

work with bounded  $\mathbf{y}, \mathbf{z}$  and only lose a small amount in the objective. The rest of the argument consists of constructing separation oracles and lower bounding the volume of inner ellipses for the ellipsoid methods.

**Lemma 3.4.** *Given a polynomial  $p \in \mathbb{R}[z_1, \dots, z_n]$  with nonnegative coefficients with the guarantee that  $\text{supp}(p)$  is a jump system and  $\boldsymbol{\alpha} \in \text{newt}(p)$  and  $\epsilon > 0$ , we have*

$$\inf_{\mathbf{z}: \|\log(\mathbf{z})\|_\infty < M} \frac{p(\mathbf{z})}{\mathbf{z}^\alpha} \leq (1 + \epsilon) \inf_{\mathbf{z} > 0} \frac{p(\mathbf{z})}{\mathbf{z}^\alpha},$$

for  $M = \text{poly}(\langle p \rangle + \log(1/\epsilon))$ .

*Proof.* It is enough to prove that for any  $\mathbf{z} > 0$ , we can find  $\mathbf{z}^*$  with  $\|\log(\mathbf{z}^*)\|_\infty < M$  such that

$$\frac{p(\mathbf{z}^*)}{(\mathbf{z}^*)^\alpha} < (1 + \epsilon) \frac{p(\mathbf{z})}{\mathbf{z}^\alpha}.$$

For the convenience of notation, assume that  $z_{n+1} = 1$ . Without loss of generality, after renaming coordinates, assume that

$$|\log(z_1)| \geq |\log(z_2)| \geq \dots \geq |\log(z_n)| \geq |\log(z_{n+1})| = 0.$$

Let  $\Delta$  be a value bounded by  $\text{poly}(\langle p \rangle + \log(1/\epsilon))$  that will be fixed later. Let  $i$  be the coordinate for which  $|\log(z_i)| - |\log(z_{i+1})|$  is maximized. If this maximum value is bounded by  $\Delta$ , we are done. Otherwise we change  $\mathbf{z}$  by subtracting a constant from  $|\log(z_1)|, \dots, |\log(z_i)|$  to make sure that  $|\log(z_i)| - |\log(z_{i+1})|$  is reduced to  $\Delta$  and  $p(\mathbf{z})/\mathbf{z}^\alpha$  does not grow by more than a  $1 + \epsilon/2n$  factor. We obtain  $\mathbf{z}^*$  by repeating this “gap-closing” procedure for no more than  $n$  times.

Define the vector  $\mathbf{c} \in \{-1, 0, +1\}^n$  in the following way:

$$c_i := \begin{cases} +1 & \text{if } j \leq i \text{ and } z_j > 1, \\ -1 & \text{if } j \leq i \text{ and } z_j < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let  $s = |\log(z_i)| - |\log(z_{i+1})| - \Delta$ . Let  $\tilde{\mathbf{z}} = \exp(-s\mathbf{c})\mathbf{z}$  (i.e.,  $\tilde{z}_i = e^{-sc_i} z_i$ ) be the result of the “gap-closing” operation. We just need to set  $\Delta$  in such a way that

$$\frac{p(\tilde{\mathbf{z}})}{\tilde{\mathbf{z}}^\alpha} \leq (1 + \epsilon/2n) \frac{p(\mathbf{z})}{\mathbf{z}^\alpha}. \quad (11)$$

Let us partition  $\mathcal{F} = \text{supp}(p)$  into  $\mathcal{F}_1 \cup \mathcal{F}_2$ ; let

$$\mathcal{F}_1 := \{\boldsymbol{\kappa} \in \mathcal{F} : \langle \mathbf{c}, \boldsymbol{\kappa} \rangle \geq \langle \mathbf{c}, \boldsymbol{\alpha} \rangle\},$$

and let  $\mathcal{F}_2$  be the rest. Note that  $\mathcal{F}_1 \neq \emptyset$  because  $\boldsymbol{\alpha} \in \text{newt}(p) = \text{conv}(\mathcal{F})$ . We can decompose  $p = p_1 + p_2$ :

$$p(\mathbf{z}) = \underbrace{\sum_{\boldsymbol{\kappa} \in \mathcal{F}_1} C_p(\boldsymbol{\kappa}) \mathbf{z}^\boldsymbol{\kappa}}_{p_1(\mathbf{z})} + \underbrace{\sum_{\boldsymbol{\kappa} \in \mathcal{F}_2} C_p(\boldsymbol{\kappa}) \mathbf{z}^\boldsymbol{\kappa}}_{p_2(\mathbf{z})}.$$



To prove (11), it is enough to show the following two inequalities

$$\frac{p_1(\tilde{\mathbf{z}})}{\tilde{\mathbf{z}}^\alpha} \leq \frac{p_1(\mathbf{z})}{\mathbf{z}^\alpha}, \quad (12)$$

$$p_2(\tilde{\mathbf{z}}) \leq \frac{\epsilon}{2n} p_1(\tilde{\mathbf{z}}). \quad (13)$$

First we prove (12). For any  $\boldsymbol{\kappa} \in \mathcal{F}_1$ , we have

$$\frac{\tilde{\mathbf{z}}^\boldsymbol{\kappa}}{\tilde{\mathbf{z}}^\alpha} = \tilde{\mathbf{z}}^{\boldsymbol{\kappa}-\alpha} = (\exp(-s\mathbf{c})\mathbf{z})^{\boldsymbol{\kappa}-\alpha} = \exp(-s\langle \mathbf{c}, \boldsymbol{\kappa} - \alpha \rangle) \frac{\mathbf{z}^\boldsymbol{\kappa}}{\mathbf{z}^\alpha} \leq \frac{\mathbf{z}^\boldsymbol{\kappa}}{\mathbf{z}^\alpha},$$

where in the inequality we used  $\boldsymbol{\kappa} \in \mathcal{F}_1$  and  $s \geq 0$ . Multiplying the above inequality with  $C_p(\boldsymbol{\kappa})$  and summing over all  $\boldsymbol{\kappa} \in \mathcal{F}_1$  we get (12).

It remains to prove (13). We will use the following claim to prove (13).

**Claim 3.5.** *For any  $\boldsymbol{\kappa} \in \mathcal{F}_2$ , there exists  $\boldsymbol{\kappa}' \in \mathcal{F}_1$  such that*

$$\tilde{\mathbf{z}}^\boldsymbol{\kappa} \leq e^{-\Delta} \tilde{\mathbf{z}}^{\boldsymbol{\kappa}'}.$$

*Proof.* Let  $\boldsymbol{\kappa}'$  be the output of the greedy algorithm (with the ordering of coordinates being  $1, \dots, n$ ) from Theorem 2.16 for the jump system  $\mathcal{F}$  and linear function  $\mathbf{x} \mapsto \langle \mathbf{c}, \mathbf{x} \rangle$ . We start from  $\boldsymbol{\kappa}$  and move towards  $\boldsymbol{\kappa}'$ , ensuring that  $\boldsymbol{\kappa}$  remains in  $\mathcal{F}$  and that  $\tilde{\mathbf{z}}^\boldsymbol{\kappa}$  never decreases and in at least one step it increases by a factor of  $e^\Delta$ .

While  $\boldsymbol{\kappa} \neq \boldsymbol{\kappa}'$ , choose the smallest  $j$  for which  $\kappa_j \neq \kappa'_j$ . If  $j > i$ , stop. Otherwise, by the greedy construction of  $\boldsymbol{\kappa}'$ , we know that  $c_j(\kappa'_j - \kappa_j) > 0$ . Let  $\boldsymbol{\delta} \in \mathbb{Z}^n$  be such that  $\delta_j = c_j$  and  $\delta_k = 0$  for  $k \neq j$ . Note that  $\boldsymbol{\delta}$  is a  $(\boldsymbol{\kappa}, \boldsymbol{\kappa}')$ -step, see Definition 2.14. Therefore either  $\boldsymbol{\kappa} + \boldsymbol{\delta} \in \mathcal{F}$  or there exists a  $(\boldsymbol{\kappa} + \boldsymbol{\delta}, \boldsymbol{\kappa}')$ -step  $\boldsymbol{\delta}'$  such that  $\boldsymbol{\kappa} + \boldsymbol{\delta} + \boldsymbol{\delta}' \in \mathcal{F}$ . In the former case, we set  $\boldsymbol{\kappa}$  to  $\boldsymbol{\kappa} + \boldsymbol{\delta}$  and in the latter case we set  $\boldsymbol{\kappa}$  to  $\boldsymbol{\kappa} + \boldsymbol{\delta} + \boldsymbol{\delta}'$ . We show that after the move,  $\tilde{\mathbf{z}}^\boldsymbol{\kappa}$  does not decrease.

**Case 1** ( $\boldsymbol{\kappa} + \boldsymbol{\delta} \in \mathcal{F}$ ). Since  $c_j \log(\tilde{z}_j) \geq c_i \log(\tilde{z}_i) \geq \Delta > 0$ , we get

$$\tilde{\mathbf{z}}^{\boldsymbol{\kappa}+\boldsymbol{\delta}} = z_j^{c_j} \tilde{\mathbf{z}}^\boldsymbol{\kappa} \geq e^\Delta \tilde{\mathbf{z}}^\boldsymbol{\kappa} \geq \tilde{\mathbf{z}}^\boldsymbol{\kappa}.$$

**Case 2** ( $\boldsymbol{\kappa} + \boldsymbol{\delta} + \boldsymbol{\delta}' \in \mathcal{F}$ ). Recall that  $\boldsymbol{\delta}'$  has only one nonzero coordinate and that coordinate is  $\pm 1$ . Let that be the  $k$ -th coordinate. Since the first  $j-1$  coordinates of  $\boldsymbol{\kappa}$  and  $\boldsymbol{\kappa}'$  are the same, it must be that  $k \geq j$ . Therefore

$$\tilde{\mathbf{z}}^{\boldsymbol{\kappa}+\boldsymbol{\delta}+\boldsymbol{\delta}'} = z_j^{c_j} z_k^{\delta'_k} \tilde{\mathbf{z}}^\boldsymbol{\kappa} \geq z_j^{c_j} z_k^{-c_k} \tilde{\mathbf{z}}^\boldsymbol{\kappa} \geq \tilde{\mathbf{z}}^\boldsymbol{\kappa}.$$

It remains to show that in one of the steps,  $\tilde{\mathbf{z}}^\boldsymbol{\kappa}$  grows by at least  $e^\Delta$ . Look at the quantity  $\langle \mathbf{c}, \boldsymbol{\kappa} \rangle$  and how it changes over the course of the above algorithm. At the beginning, since  $\boldsymbol{\kappa} \in \mathcal{F}_2$

$$\langle \mathbf{c}, \boldsymbol{\kappa} \rangle < \langle \mathbf{c}, \boldsymbol{\alpha} \rangle.$$

When the algorithm finishes  $\langle \mathbf{c}, \boldsymbol{\kappa} \rangle = \langle \mathbf{c}, \boldsymbol{\kappa}' \rangle$ ; this is because  $\kappa_j = \kappa'_j$  for  $j \leq i$  and  $c_j = 0$  for  $j > i$ . Note that  $\langle \mathbf{c}, \boldsymbol{\kappa}' \rangle \geq \langle \mathbf{c}, \boldsymbol{\alpha} \rangle$  because  $\boldsymbol{\kappa}' \in \mathcal{F}_1$ . Therefore, at one point the quantity  $\langle \mathbf{c}, \boldsymbol{\kappa} \rangle$  must increase. It is easy to see that when  $\langle \mathbf{c}, \boldsymbol{\kappa} \rangle$  increases,  $\tilde{\mathbf{z}}^\boldsymbol{\kappa}$  grows by at least a factor of  $e^\Delta$ . In case 1,  $\tilde{\mathbf{z}}^\boldsymbol{\kappa}$  always increases by a factor of  $e^\Delta$ . In case 2,  $\langle \mathbf{c}, \boldsymbol{\kappa} \rangle$  increases exactly when either  $\delta'_k = c_k$  or  $k > i$ , and in both situations  $\tilde{\mathbf{z}}^\boldsymbol{\kappa}$  also increases by a factor of  $e^\Delta$ .  $\square$

Now let us finish the proof of (13). We have

$$p_2(\tilde{\mathbf{z}}) = \sum_{\kappa \in \mathcal{F}_2} C_p(\kappa) \tilde{\mathbf{z}}^\kappa \leq (\deg p + 1)^n \cdot \max_{\kappa \in \mathcal{F}_2} C_p(\kappa) \cdot \max_{\kappa \in \mathcal{F}_2} \tilde{\mathbf{z}}^\kappa,$$

where we used the fact that the number of terms in  $p$  is bounded by  $(\deg p + 1)^n$ . Similarly for  $p_1$ , we have

$$p_1(\tilde{\mathbf{z}}) = \sum_{\kappa \in \mathcal{F}_1} C_p(\kappa) \tilde{\mathbf{z}}^\kappa \geq \min_{\kappa \in \mathcal{F}_1} C_p(\kappa) \cdot \max_{\kappa \in \mathcal{F}_1} \tilde{\mathbf{z}}^\kappa.$$

Putting the above two inequalities and using Claim 3.5, we get

$$\frac{p_2(\tilde{\mathbf{z}})}{p_1(\tilde{\mathbf{z}})} \leq \frac{(\deg p + 1)^n \cdot \max_{\kappa \in \mathcal{F}_2} C_p(\kappa)}{\min_{\kappa \in \mathcal{F}_1} C_p(\kappa)} \cdot \frac{\max_{\kappa \in \mathcal{F}_2} \tilde{\mathbf{z}}^\kappa}{\max_{\kappa \in \mathcal{F}_1} \tilde{\mathbf{z}}^\kappa} \leq e^{\text{poly}(\langle p \rangle) - \Delta}.$$

It is easy to see that by making  $\Delta = \text{poly}(\langle p \rangle) + \log(2n/\epsilon) = \text{poly}(\langle p \rangle) + \log(1/\epsilon)$ , we get the desired inequality (13).  $\square$

In the remaining part of this section, we first show that we can construct separation oracles, and then we bound the volume of the inner ellipses for the two ellipsoid methods. Let us construct the separation oracles for the two ellipsoid methods. For the ellipsoid computing  $\inf_{\mathbf{y}, \mathbf{z} > 0} p(\mathbf{y}, \mathbf{z}) / (\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^\alpha$ , the separation oracle just needs to compute the gradient of the (log of the) objective with respect to  $\log(\mathbf{y}), \log(\mathbf{z})$ . It is easy to see that this gradient can be computed from the partial derivatives of  $p$  with respect to  $\mathbf{y}, \mathbf{z}$ . The latter can be computed using the following lemma.

**Lemma 3.6.** *Given oracle access to a polynomial  $p \in \mathbb{R}[y_1, \dots, y_n]$ , for any  $i$ , the quantity  $\partial_{y_i} p$  can be computed at any given point using  $O(\deg p)$  oracle calls.*

*Proof.* Given a point  $\mathbf{y}$ , consider the polynomial

$$q(t) = p(y_1, \dots, y_{i-1}, y_i + t, y_{i+1}, \dots, y_n).$$

This is a univariate polynomial of degree at most  $\deg p$ . Therefore we can compute all of its coefficients by evaluating  $p$  at  $\deg p + 1$  many points. Then we can simply output the coefficient of  $t$ .  $\square$

Now we construct the separation oracle for the ellipsoid method computing

$$\sup_{\boldsymbol{\alpha} \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^\alpha}.$$

Recall that if  $\boldsymbol{\alpha} \notin \text{newt}(p)$ , then  $\inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^\alpha} = 0$ , see Fact 2.18. Given some  $\boldsymbol{\alpha} \geq 0$ , we first check whether  $\boldsymbol{\alpha} \in \text{newt}(p)$  using Fact 2.20, and in case it is not, we return the promised separating hyperplane. So let us assume that  $\boldsymbol{\alpha} \in \text{newt}(p)$ .

In its current form, the expression  $\frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^\alpha}$  is not concave in  $\boldsymbol{\alpha}$ . So we need to use the same trick as in Lemma 2.12 to transform it into a concave form. We do this transformation *after* finding an approximate optimum for the inf. We compute an approximate optimum of  $\inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\alpha})^\alpha}$  by restricting the domain of the variables  $\mathbf{y}, \mathbf{z}$  using Lemma 3.4 and running the inner ellipsoid method to desired accuracy.

Next, we find a separating hyperplane at  $\alpha$ , by calculating the gradient of

$$\frac{p(\beta(\mathbf{y}/\alpha), \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\beta}$$

with respect to  $\beta$  evaluated at  $\beta = \alpha$ . This can again be done using [Lemma 3.6](#).

It remains to lower bound the volume of the inner ellipses in the ellipsoid methods. First consider the ellipsoid method computing the inf. It is enough to lower bound the volume of the following set for any given  $\epsilon > 0$ ,  $\mathbf{y}^*, \mathbf{z}^* > 0$

$$\left\{ (\log(\mathbf{y}), \log(\mathbf{z})) \mid \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \leq (1 + \epsilon) \frac{p(\mathbf{y}^*, \mathbf{z}^*)}{(\mathbf{y}^*\mathbf{z}^*/\alpha)^\alpha} \right\}.$$

To see this, it is enough to note that  $\log(\frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha})$  is Lipschitz with respect to  $\log(\mathbf{y}), \log(\mathbf{z})$  with Lipschitz constant bounded by  $\text{poly}(\deg p, n) = \text{poly}(\langle p \rangle)$ .

Now consider the ellipsoid method for computing the sup. To bound the volume of the inner ellipse we are going to use a similar idea as before, except we prove a slightly weaker form of smoothness compared to being Lipschitz. We only consider the set of  $\alpha$ 's in  $\text{newt}(p)$ . We are going to assume  $\text{newt}(p)$  is full-dimensional. If  $\text{newt}(p)$  is lower-dimensional we can use subspace-identification methods to restrict the ellipsoid method to the affine subspace spanned by  $\text{newt}(p)$ , and proceed similarly to the full-dimensional case.

Let us start by considering the function

$$g_{\mathbf{y}, \mathbf{z}}(\alpha) = \log\left(\frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z})^\alpha}\right).$$

It is easy to see that for any  $\mathbf{y}, \mathbf{z} > 0$ ,  $g_{\mathbf{y}, \mathbf{z}}$  is Lipschitz with respect to  $\alpha$  with Lipschitz constant bounded by  $\text{poly}(\log(\mathbf{y}), \log(\mathbf{z}), n)$ . So the function

$$g_M(\alpha) = \inf_{\mathbf{y}, \mathbf{z}: \|\mathbf{y}\|_\infty, \|\mathbf{z}\|_\infty < M} g_{\mathbf{y}, \mathbf{z}}(\alpha)$$

is also Lipschitz with respect to  $\alpha$  with Lipschitz constant bounded by  $\text{poly}(\log(M), n)$ . Now consider the function  $h(\alpha) = g_M(\alpha) + \alpha \log(\alpha)$ . This function is not Lipschitz in  $\alpha$ , but it satisfies a weaker form of smoothness. In particular if  $\alpha$  is perturbed by  $\epsilon^2$ , one can show that  $h(\alpha)$  changes by at most  $L\epsilon$  for some  $L = \text{poly}(\log(M), n)$ . The function  $h$  provides a uniform approximation of  $\inf_{\mathbf{y}, \mathbf{z} > 0} \log(\frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha})$  over  $\text{newt}(p)$ . Using this fact and that  $\text{newt}(p)$  has only integral vertices of bounded norm, one can construct the desired inner ellipse.

## 4 The Lower Bound

In this section we prove the LHS of (3) and (4). However, instead of lower bounding the quantity  $\sum_{\kappa \in \mathbb{Z}_+^{2n}} C_p(\kappa) C_q(\kappa)$  for two real stable polynomials  $p(\mathbf{y}), q(\mathbf{z})$ , we treat  $p(\mathbf{y}) \cdot q(\mathbf{z})$  as a new bistable polynomial on  $2n$  variables  $\mathbf{y}, \mathbf{z}$ . Recall that by [Fact 2.3](#) the product of any two stable polynomials is a bistable polynomial. This allows us to prove the lower bound using an inductive argument with a stronger statement. The following theorem implies the LHS of (3).

**Theorem 4.1.** For any bistable polynomial  $p(y_1, \dots, y_n, z_1, \dots, z_n)$  with nonnegative coefficients and any vector  $\alpha \in \mathbb{R}_+^n$ ,

$$\sum_{\kappa \in \mathbb{Z}_+^n} \kappa! \cdot C_p(\kappa, \kappa) \geq e^{-\alpha} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}. \quad (14)$$

Instead of directly proving the above theorem, we reduce the general case to the multilinear case. This reduction allows us to exploit Brändén characterization of real stable multilinear polynomials (see [Theorem 2.7](#)).

**Theorem 4.2.** For any bistable multilinear polynomial  $p(y_1, \dots, y_n, z_1, \dots, z_n)$  with nonnegative coefficients and any vector  $\alpha \in \mathbb{R}_+^n$ ,

$$\sum_{\kappa \in \{0,1\}^n} C_p(\kappa, \kappa) \geq (1 - \alpha)^{1-\alpha} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}$$

Note that not only do we use the above theorem to prove [Theorem 4.1](#), but we also use it to prove (4). In the rest of this section, we prove [Theorem 4.1](#) using the above theorem; then in [Subsection 4.1](#) we prove the above theorem. Observe that if  $p$  is a multilinear bistable polynomial, then [Theorem 4.1](#) follows from the above theorem (this is because  $e^{-x} \leq (1-x)^{1-x}$  for any  $0 \leq x \leq 1$ ). So, we just need to reduce the case of a general  $p$  to the multilinear case.

Let  $m$  be a very large integer that we fix later in the proof. For now, we assume  $m$  is larger than the degree of all variables in  $p$ . Throughout the proof we use  $\mathbf{y}, \mathbf{z}$  to denote  $n$ -dimensional (positive) vectors and we use  $\tilde{\mathbf{y}}, \tilde{\mathbf{z}} \in \mathbb{R}^{n \times m}$  to denote the variables of the polarization of  $p$ . Let  $q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}) := \pi_m(p)$  and  $r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}) := q(\tilde{\mathbf{y}}, m\tilde{\mathbf{z}})$ . Also, let  $\tilde{\alpha} \in \mathbb{R}_+^{n \times m}$  where for all  $1 \leq i \leq n$  and  $1 \leq j \leq m$ ,

$$\tilde{\alpha}_{i,j} = \frac{1}{m} \alpha_i.$$

By applying [Theorem 4.2](#) to  $r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})$  and  $\tilde{\alpha}$  we get

$$\sum_{\kappa \in \{0,1\}^{n \times m}} C_r(\kappa, \kappa) \geq (1 - \tilde{\alpha})^{1-\tilde{\alpha}} \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(\tilde{\mathbf{y}}\tilde{\mathbf{z}}/\tilde{\alpha})^{\tilde{\alpha}}}. \quad (15)$$

First, we show that the RHS of the above is at least the RHS of (14) for any  $m$ . Then, we show that as  $m \rightarrow \infty$  the LHS of the above converges to the LHS of (14).

**Lemma 4.3.**

$$(1 - \tilde{\alpha})^{1-\tilde{\alpha}} \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(\tilde{\mathbf{y}}\tilde{\mathbf{z}}/\tilde{\alpha})^{\tilde{\alpha}}} \geq e^{-\alpha} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}.$$

*Proof.* Firstly, note that by the definition of  $\tilde{\alpha}$ ,  $\tilde{\alpha}^{\tilde{\alpha}} = (\alpha/m)^\alpha$ . Also, using the inequality  $(1-x)^{1-x} \geq e^{-x}$ , we have

$$(1 - \tilde{\alpha})^{1-\tilde{\alpha}} \geq e^{-\tilde{\alpha}} = e^{-\alpha}.$$

Therefore, to prove the lemma it is enough to show that

$$\inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(m\tilde{\mathbf{y}}\tilde{\mathbf{z}})^{\tilde{\alpha}}} \geq \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z})^\alpha}. \quad (16)$$

It follows by the change of variable  $\tilde{\mathbf{z}} \leftrightarrow \tilde{\mathbf{z}}/m$  that

$$\inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(m\tilde{\mathbf{y}}\tilde{\mathbf{z}})^{\tilde{\alpha}}} = \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}/m)}{(\tilde{\mathbf{y}}\tilde{\mathbf{z}})^{\tilde{\alpha}}} = \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(\tilde{\mathbf{y}}\tilde{\mathbf{z}})^{\tilde{\alpha}}},$$

where in the last identity we used  $r(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}/m) = q(\tilde{\mathbf{y}}, m\tilde{\mathbf{z}}/m) = q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})$ . So, to prove (16), it is enough to show that

$$\inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} > 0} \frac{q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(\tilde{\mathbf{y}}\tilde{\mathbf{z}})^{\tilde{\alpha}}} \geq \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z})^{\alpha}}. \quad (17)$$

We use the AM-GM inequality to prove this. Fix an arbitrary  $\tilde{\mathbf{y}}, \tilde{\mathbf{z}} \in \mathbb{R}^{n \times m}$ , let  $y_i = (\prod_j \tilde{y}_{i,j})^{1/m}$  and  $z_i = (\prod_j \tilde{z}_{i,j})^{1/m}$  for all  $i$ . We show that  $\frac{q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})}{(\tilde{\mathbf{y}}\tilde{\mathbf{z}})^{\tilde{\alpha}}} \geq \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z})^{\alpha}}$  which proves (17).

Firstly,  $\tilde{\mathbf{y}}^{\tilde{\alpha}} = \mathbf{y}^{\alpha}$  and similarly  $\tilde{\mathbf{z}}^{\tilde{\alpha}} = \mathbf{z}^{\alpha}$ . Therefore, all we need to show is that  $p(\mathbf{y}, \mathbf{z}) \leq q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})$ . Recall that  $q = \pi_m(p)$ ; this means that we have substituted each term  $y_i^k$  in  $p$  with  $\frac{1}{\binom{m}{k}} e_k(y_{i,1}, \dots, y_{i,m})$ . Since all coefficients of  $p$  are nonnegative, to show  $p(\mathbf{y}, \mathbf{z}) \leq q(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})$  it is enough to show that for all  $i, k, m$

$$\frac{1}{\binom{m}{k}} e_k(\tilde{y}_{i,1}, \dots, \tilde{y}_{i,m}) \leq y_i^k,$$

or equivalently,

$$\sum_{S \in \binom{[m]}{k}} \frac{1}{\binom{m}{k}} \prod_{j \in S} \tilde{y}_{i,j} \geq \left( \prod_{j=1}^m \tilde{y}_{i,j} \right)^{1/m}.$$

The above inequality is just an application of AM-GM. This proves (17) and completes the proof of Lemma 4.3.  $\square$

To finish the proof of Theorem 4.1 it remains to show that the LHS of (15) converges to the LHS of (14) as  $m \rightarrow \infty$ . We prove this statement term by term.

Fix  $\boldsymbol{\kappa} \in \mathbb{Z}_+^n$  and consider the monomial  $C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \prod_{i=1}^n y_i^{\kappa_i} z_i^{\kappa_i}$  of  $p$ . The contribution of this monomial to the LHS of (14) is  $\boldsymbol{\kappa}! \cdot C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa})$ . This monomial is substituted with

$$\begin{aligned} & C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \prod_{i=1}^n \frac{e_{\kappa_i}(y_{i,1}, \dots, y_{i,m}) e_{\kappa_i}(mz_{i,1}, \dots, mz_{i,m})}{\binom{m}{\kappa_i}^2} \\ &= C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \prod_{i=1}^n \frac{m^{\kappa_i} e_{\kappa_i}(y_{i,1}, \dots, y_{i,m}) e_{\kappa_i}(z_{i,1}, \dots, z_{i,m})}{\binom{m}{\kappa_i}^2}. \end{aligned}$$

in  $r$ . The contribution of this to the LHS of (15) is

$$C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \prod_{i=1}^n \frac{m^{\kappa_i}}{\binom{m}{\kappa_i}^2} \cdot \binom{m}{\kappa_i} = \boldsymbol{\kappa}! \cdot C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \cdot \prod_{i=1}^n \frac{m^{\kappa_i}}{m(m-1)\dots(m-\kappa_i+1)}.$$

As  $m \rightarrow \infty$  the ratio in the RHS converges to 1, so we get the same contribution. This completes the proof of Theorem 4.1.

## 4.1 Multilinear Case

In this section we prove [Theorem 4.2](#). Our first observation is the following useful identity:

$$\sum_{\boldsymbol{\kappa} \in \{0,1\}^n} C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) = \prod_{i=1}^n (1 + \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, \mathbf{z}) \Big|_{\mathbf{y}=\mathbf{z}=0}.$$

To see this identity, note that any monomial of  $p$  which is not of the form  $\mathbf{y}^{\boldsymbol{\kappa}} \mathbf{z}^{\boldsymbol{\kappa}}$  for some  $\boldsymbol{\kappa}$  is mapped to zero in the RHS. Furthermore, any monomial of the form  $C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa}) \mathbf{y}^{\boldsymbol{\kappa}} \mathbf{z}^{\boldsymbol{\kappa}}$  is mapped to  $C_p(\boldsymbol{\kappa}, \boldsymbol{\kappa})$ . So, throughout the proof we lower bound  $\prod_{i=1}^n (1 + \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, \mathbf{z}) \Big|_{\mathbf{y}=\mathbf{z}=0}$ . The main reason that we use this reformulation is that it allows us to use induction. In the following lemma we show that for any bistable polynomial  $p$ ,  $(1 + \partial_{y_i} \partial_{z_i}) p$  is also bistable. So, we can apply the operator  $\prod_{i=1}^n (1 + \partial_{y_i} \partial_{z_i})$  inductively.

**Lemma 4.4.** *For any multilinear bistable polynomial  $p(y_1, \dots, y_n, z_1, \dots, z_n)$  and any  $1 \leq i \leq n$ ,  $(1 + \partial_{y_i} \partial_{z_i}) p|_{y_i=z_i=0}$  is also bistable.*

*Proof.* Let  $q(\mathbf{y}_{-i}, \mathbf{z}_{-i}) = (1 + \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, \mathbf{z})|_{y_i=z_i=0}$ . We need to show that  $q(\mathbf{y}_{-i}, -\mathbf{z}_{-i})$  is stable. We will prove that

$$q(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) = (1 - \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, -\mathbf{z})|_{y_i=z_i=0}. \quad (18)$$

Then, the claim follows from the fact the the polynomial on the RHS is stable. In particular, since  $p(\mathbf{y}, -\mathbf{z})$  is stable, by [Lemma 2.6](#),  $(1 - \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, -\mathbf{z})$  is stable. Furthermore, by [Fact 2.4](#) when we substitute  $y_i = z_i = 0$  in the latter we still have a stable polynomial.

It remains to show (18). We can write  $p$  as follows:

$$p(\mathbf{y}, \mathbf{z}) = y_i z_i p_1(\mathbf{y}_{-i}, \mathbf{z}_{-i}) + y_i p_2(\mathbf{y}_{-i}, \mathbf{z}_{-i}) + z_i p_3(\mathbf{y}_{-i}, \mathbf{z}_{-i}) + p_4(\mathbf{y}_{-i}, \mathbf{z}_{-i}).$$

So,

$$\begin{aligned} q(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) &= (1 + \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, -z_1, \dots, -z_{i-1}, z_i, -z_{i+1}, \dots, z_n)|_{y_i=z_i=0} \\ &= p_1(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) + p_4(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) \end{aligned}$$

On the other hand,

$$\begin{aligned} (1 - \partial_{y_i} \partial_{z_i}) p(\mathbf{y}, -\mathbf{z})|_{y_i=z_i=0} &= (1 - \partial_{y_i} \partial_{z_i}) (-y_i z_i p(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) + y_i p_2(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) \\ &\quad - z_i p_3(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) + p_4(\mathbf{y}_{-i}, -\mathbf{z}_{-i}))|_{y_i=z_i=0} \\ &= p_1(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) + p_4(\mathbf{y}_{-i}, -\mathbf{z}_{-i}) \end{aligned}$$

This proves (18). □

We prove [Theorem 4.2](#) by induction. See [Lemma 4.5](#) for the base case of the induction. Suppose the statement of the theorem holds for any bistable multilinear polynomial with nonnegative coefficients and at most  $2n - 2$  variables. Given  $p(y_1, \dots, y_n, z_1, \dots, z_n)$  and some  $\epsilon > 0$  we show that there exists  $\mathbf{y}, \mathbf{z} > 0$  such that

$$\prod_{i=1}^n (1 + \partial_{y_i} \partial_{z_i}) p \Big|_{\mathbf{y}=\mathbf{z}=0} + \epsilon \geq (1 - \alpha)^{1-\alpha} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}. \quad (19)$$

Let

$$q(y_1, \dots, y_{n-1}, z_1, \dots, z_{n-1}) = (1 + \partial_{y_n} \partial_{z_n})p|_{y_n=z_n=0}.$$

Note that by definition,  $q$  has nonnegative coefficients and is multilinear. Also, by [Lemma 4.4](#),  $q$  is a bistable polynomial.

Now, by the induction hypothesis, there exists  $\mathbf{y}^*, \mathbf{z}^* \in \mathbb{R}^{n-1}$  such that

$$\prod_{i=1}^{n-1} (1 + \partial_{y_i} \partial_{z_i})q|_{\mathbf{y}=\mathbf{z}=0} + \epsilon/2 \geq (1 - \alpha_{-n})^{1-\alpha_{-n}} \frac{q(\mathbf{y}^*, \mathbf{z}^*)}{(\mathbf{y}^* \mathbf{z}^* / \alpha_{-n})^{\alpha_{-n}}}.$$

Also, observe that by the definition of  $q$ , the LHS of the above is equal to the LHS of [\(19\)](#). So to prove [\(19\)](#), it is enough to show that there exist  $y_n, z_n$  such that

$$(1 - \alpha_{-n})^{1-\alpha_{-n}} \frac{q(\mathbf{y}^*, \mathbf{z}^*)}{(\mathbf{y}^* \mathbf{z}^* / \alpha_{-n})^{\alpha_{-n}}} + \epsilon/2 \geq (1 - \alpha)^{1-\alpha} \frac{p(\mathbf{y}^*, y_n, \mathbf{z}^*, z_n)}{(\mathbf{y}^* \mathbf{z}^* / \alpha_{-n})^{\alpha_{-n}} (y_n z_n / \alpha_n)^{\alpha_n}}.$$

Let  $f(y_n, z_n) = p(\mathbf{y}^*, y_n, \mathbf{z}^*, z_n)$ . To prove the above it is enough to show that there exists  $y_n, z_n$  such that

$$(1 + \partial_{y_n} \partial_{z_n})f(y_n, z_n)|_{y_n=z_n=0} + \epsilon/2 \geq (1 - \alpha_n)^{1-\alpha_n} \frac{f(y_n, z_n)}{(y_n z_n / \alpha_n)^{\alpha_n}}. \quad (20)$$

We prove this using [Lemma 4.5](#). All we need to apply this lemma is that  $f(y_n, z_n)$  is a multilinear bistable polynomial with nonnegative coefficients. Since  $p$  is multilinear with nonnegative coefficients, and we substitute  $y_1, \dots, y_{n-1}, z_1, \dots, z_{n-1}$  with positive numbers, we get that  $f$  is multilinear with nonnegative coefficients. Furthermore, since  $p(y_1, \dots, y_n, -z_1, \dots, -z_n)$  is stable, by [Fact 2.4](#),

$$f(y_n, -z_n) = p(y_1^*, \dots, y_{n-1}^*, y_n, z_1^*, \dots, z_{n-1}^*, -z_n)$$

is stable. So,  $f$  is a bistable polynomial. Therefore, [\(20\)](#) simply follows by [Lemma 4.5](#). This completes the proof of [Theorem 4.2](#).

**Lemma 4.5.** *Let  $p(y, z)$  be a multilinear bistable bivariate polynomial with nonnegative coefficients. Then, for any  $\alpha \geq 0$ ,*

$$(1 + \partial_y \partial_z)p|_{y=z=0} \geq (1 - \alpha)^{1-\alpha} \inf_{y, z > 0} \frac{p(y, z)}{(yz/\alpha)^\alpha}$$

*Proof.* Since  $p$  is multilinear, if  $\alpha > 1$ , then the RHS of the above is zero and we are done. In particular, for a fixed  $z$ , as  $y \rightarrow \infty$ ,  $p(y, z)/y^\alpha \rightarrow 0$ . So, throughout the proof we assume that  $\alpha \leq 1$ .

Since  $p$  is multilinear we can write it as

$$p(y, z) = ayz + by + cz + d,$$

for  $a, b, c, d \geq 0$ . To prove the lemma's conclusion, it is enough to show

$$a + d \geq (1 - \alpha)^{1-\alpha} \inf_{y, z > 0} \frac{ayz + by + cz + d}{(yz/\alpha)^\alpha}. \quad (21)$$

When  $\alpha = 1$ , the RHS of [\(21\)](#) is at most  $a$  since the ratio converges to  $a$  as  $y, z \rightarrow \infty$ . When  $\alpha = 0$ , the RHS of [\(21\)](#) is at most  $d$  since the ratio converges to  $d$  as  $y, z \rightarrow 0$ . In both cases the RHS is at most  $a + d$  and we are done. So for the rest of the proof assume that  $0 < \alpha < 1$ .

Since  $p(x, -y)$  is stable, by [Theorem 2.7](#) we have

$$-b \cdot c = \partial_y p(y, -z)|_{y=z=0} \cdot \partial_z p(y, -z)|_{y=z=0} \geq p(0, 0) \cdot \partial_y \partial_z p(y, -z)|_{y=z=0} = -a \cdot d,$$

or, in other words  $b \cdot c \leq a \cdot d$ . If this inequality is not tight, we can increase  $b, c$  for it to become tight. In other words, let  $b' \geq b$  and  $c' \geq c$  be such that  $b'c' = ad$ . It is easy to see that such  $b', c'$  can always be found. Then, to prove [\(21\)](#), it is enough to show that

$$a + d \geq (1 - \alpha)^{1-\alpha} \inf_{y, z > 0} \frac{ayz + b'y + c'z + d}{(yz/\alpha)^\alpha},$$

Consider the following matrix:

$$M := \begin{bmatrix} a & b' \\ c' & d \end{bmatrix}.$$

Since  $\det(M) = ad - b'c' = 0$ , we have  $\text{rank}(M) \leq 1$ , and in particular we can write  $M$  as  $uv^\top$  for  $u, v \in \mathbb{R}^2$ . Since the entries of  $M$  are nonnegative we can further assume that  $u, v \in \mathbb{R}_+^2$ . Therefore we have

$$M = \begin{bmatrix} a & b' \\ c' & d \end{bmatrix} = \begin{bmatrix} e \\ f \end{bmatrix} \cdot [g \quad h],$$

where  $e, f, g, h \geq 0$ . This implies the factorization  $ayz + b'y + c'z + d = (ey + f)(gz + h)$ . So to prove the conclusion of the lemma it is enough to show the following inequality:

$$a + d = eg + fh \geq (1 - \alpha)^{1-\alpha} \inf_{y, z > 0} \frac{(ey + f)(gz + h)}{(yz/\alpha)^\alpha} = \alpha^\alpha (1 - \alpha)^{1-\alpha} \inf_{y > 0} \frac{ey + f}{y^\alpha} \cdot \inf_{z > 0} \frac{gz + h}{z^\alpha}. \quad (22)$$

To prove the above inequality we will use the following claim.

**Claim 4.6.** For any  $0 < \alpha < 1$  and  $\beta, \gamma \geq 0$ ,

$$\inf_{t > 0} \frac{\beta t + \gamma}{t^\alpha} = \frac{\beta^\alpha \gamma^{1-\alpha}}{\alpha^\alpha (1 - \alpha)^{1-\alpha}}.$$

Before proving the claim, let us use it to prove [\(22\)](#). Because of [Claim 4.6](#), we can write the RHS of [\(22\)](#) as

$$\alpha^\alpha (1 - \alpha)^{1-\alpha} \cdot \frac{e^\alpha f^{1-\alpha}}{\alpha^\alpha (1 - \alpha)^{1-\alpha}} \cdot \frac{g^\alpha h^{1-\alpha}}{\alpha^\alpha (1 - \alpha)^{1-\alpha}} = \frac{(eg)^\alpha (fh)^{1-\alpha}}{\alpha^\alpha (1 - \alpha)^{1-\alpha}}.$$

Applying [Claim 4.6](#) once more, with  $\beta = eg, \gamma = fh$ , we get

$$\frac{(eg)^\alpha (fh)^{1-\alpha}}{\alpha^\alpha (1 - \alpha)^{1-\alpha}} = \inf_{t > 0} \frac{egt + fh}{t^\alpha} \leq \left( \frac{egt + fh}{t^\alpha} \right) \Big|_{t=1} = eg + fh,$$

which is the LHS of [\(22\)](#).

It only remains to prove [Claim 4.6](#).

*Proof.* By the weighted AM-GM inequality we have

$$\beta t + \gamma = \alpha \cdot \frac{\beta t}{\alpha} + (1 - \alpha) \cdot \frac{\gamma}{1 - \alpha} \geq \left( \frac{\beta t}{\alpha} \right)^\alpha \cdot \left( \frac{\gamma}{1 - \alpha} \right)^{1-\alpha} = t^\alpha \frac{\beta^\alpha \gamma^{1-\alpha}}{\alpha^\alpha (1 - \alpha)^{1-\alpha}}.$$



Therefore

$$\inf_{t>0} \frac{\beta t + \gamma}{t^\alpha} \geq \frac{\beta^\alpha \gamma^{1-\alpha}}{\alpha^\alpha (1-\alpha)^{1-\alpha}}.$$

Note that we have equality in AM-GM when  $\beta t/\alpha = \gamma/(1-\alpha)$ . So it is enough to set  $t = \frac{\alpha\gamma}{(1-\alpha)\beta}$  to get equality, which is possible if  $\beta, \gamma > 0$ .

It remains to prove the claim when either  $\beta = 0$  or  $\gamma = 0$ . In both cases the RHS is 0 so we just need to show that the LHS is 0 as well. When  $\beta = 0$ , it is enough to let  $t \rightarrow \infty$ , and when  $\gamma = 0$ , it is enough to let  $t \rightarrow 0$  to get

$$\inf_{t>0} \frac{\beta t + \gamma}{t^\alpha} = 0.$$

□

This completes the proof of [Lemma 4.5](#). □

## 5 The Upper Bound

In this section we prove the RHS of (3). This will complete the proof of [Theorem 1.1](#).

**Theorem 5.1.** *For any two real stable polynomials  $p(y_1, \dots, y_n)$  and  $q(z_1, \dots, z_n)$  with nonnegative coefficients,*

$$\sum_{\kappa \in \mathbb{Z}_+^n} \kappa^\kappa \cdot C_p(\kappa) C_q(\kappa) \leq \sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}. \quad (23)$$

Note that  $\kappa^\kappa \geq \kappa!$  for any  $\kappa$ , so the above theorem immediately implies the RHS of (3). Note that unlike [Theorem 4.1](#) here we only prove our theorem for product of stable polynomials (as opposed to bistable polynomials). This is because the above statement fails for (multilinear) bistable polynomials. For example, the polynomial  $p(y, z) = 1 + yz$  is bistable; however by [Claim 4.6](#), for any fixed  $0 < \alpha < 1$  we have

$$\inf_{y, z > 0} \frac{p(y, z)}{(yz/\alpha)^\alpha} = \inf_{y, z > 0} \frac{1 + yz}{(yz/\alpha)^\alpha} = (1 - \alpha)^{-(1-\alpha)}.$$

It is an easy exercise to see that  $(1 - \alpha)^{-(1-\alpha)} \leq e^{1/e}$ . But the quantity  $\sum_{\kappa \in \mathbb{Z}_+^n} \kappa^\kappa \cdot C_p(\kappa, \kappa)$  is 2 which is strictly more than  $e^{1/e}$ .

In order to prove [Theorem 5.1](#) we will crucially use the following minimax result about the RHS of (23).

**Lemma 5.2.** *For any bistable polynomial  $p(\mathbf{y}, \mathbf{z})$  with nonnegative coefficients,*

$$\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} = \inf_{\mathbf{y} > 0} \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}.$$

Before proving [Lemma 5.2](#), let us show how it can be used to prove [Theorem 5.1](#).

*Proof of Theorem 5.1.* Because of Lemma 5.2, we have

$$\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} = \inf_{\mathbf{y} > 0} \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}.$$

So to prove (23), we just need to show that for any fixed  $\mathbf{y} > 0$ , the following inequality holds:

$$\sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \geq \sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} \boldsymbol{\kappa}^\boldsymbol{\kappa} \cdot C_p(\boldsymbol{\kappa}) \cdot C_q(\boldsymbol{\kappa}).$$

Note that  $p(\mathbf{y})$  is fixed on the LHS. So

$$\begin{aligned} \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} &= p(\mathbf{y}) \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \\ &= \sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} C_p(\boldsymbol{\kappa}) \mathbf{y}^\boldsymbol{\kappa} \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}. \end{aligned} \quad (24)$$

But for any fixed  $\boldsymbol{\kappa} \in \mathbb{Z}_+^n$ , we have

$$\begin{aligned} \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} &\geq \left( \inf_{\mathbf{z} > 0} \frac{q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \right) \Big|_{\alpha=\boldsymbol{\kappa}} = \inf_{\mathbf{z} > 0} \frac{q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\boldsymbol{\kappa})^\boldsymbol{\kappa}} \\ &\geq \inf_{\mathbf{z} > 0} \frac{C_q(\boldsymbol{\kappa}) \mathbf{z}^\boldsymbol{\kappa}}{(\mathbf{y}\mathbf{z}/\boldsymbol{\kappa})^\boldsymbol{\kappa}} = \boldsymbol{\kappa}^\boldsymbol{\kappa} \cdot \frac{C_q(\boldsymbol{\kappa})}{\mathbf{y}^\boldsymbol{\kappa}}. \end{aligned}$$

Plugging back into (24) gives us

$$\sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{p(\mathbf{y})q(\mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \geq \sum_{\boldsymbol{\kappa} \in \mathbb{Z}_+^n} C_p(\boldsymbol{\kappa}) \mathbf{y}^\boldsymbol{\kappa} \boldsymbol{\kappa}^\boldsymbol{\kappa} \frac{C_q(\boldsymbol{\kappa})}{\mathbf{y}^\boldsymbol{\kappa}} = \sum_{\boldsymbol{\kappa}} \boldsymbol{\kappa}^\boldsymbol{\kappa} \cdot C_p(\boldsymbol{\kappa}) \cdot C_q(\boldsymbol{\kappa}).$$

This concludes the proof of Theorem 5.1, assuming Lemma 5.2.  $\square$

## 5.1 Duality

In this section we prove Lemma 5.2. From basic properties of inf and sup, it is immediate that

$$\sup_{\alpha \geq 0} \inf_{\mathbf{y}, \mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha} \leq \inf_{\mathbf{y} > 0} \sup_{\alpha \geq 0} \inf_{\mathbf{z} > 0} \frac{p(\mathbf{y}, \mathbf{z})}{(\mathbf{y}\mathbf{z}/\alpha)^\alpha}.$$

So to prove equality, it is enough to prove that the LHS is greater than or equal to the RHS. We would ideally like to apply one of the classical minimax theorems, such as von Neumann's minimax theorem, which often require the function inside inf and sup to be convex-concave. However in its current form  $p(\mathbf{y}, \mathbf{z})/(\mathbf{y}\mathbf{z}/\alpha)^\alpha$  does not satisfy this property. Instead we transform both sides to a form suitable for minimax theorems.

Because of Fact 2.18, we can replace  $\sup_{\alpha \geq 0}$  by  $\inf_{\alpha \in [0, \deg p]^n}$ . This is crucial for applying Theorem 5.3, since the domain of  $\alpha$  has to be compact. Additionally, by Lemma 2.12 we can move

$\alpha$  from the denominator to the numerator, and inside of  $p$ . Finally, we apply the change of variables  $\mathbf{y} \leftrightarrow \exp(\tilde{\mathbf{y}})$  and  $\mathbf{z} \leftrightarrow \exp(\tilde{\mathbf{z}})$ . Therefore, to prove the above inequality, it is enough to show that

$$\sup_{\alpha \in [0, \deg p]^n} \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} \in \mathbb{R}^n} \frac{p(\exp(\tilde{\mathbf{y}}), \alpha \exp(\tilde{\mathbf{z}}))}{\exp(\tilde{\mathbf{y}})^\alpha \exp(\tilde{\mathbf{z}})^\alpha} \geq \inf_{\tilde{\mathbf{y}} \in \mathbb{R}^n} \sup_{\alpha \in [0, \deg p]^n} \inf_{\tilde{\mathbf{z}} \in \mathbb{R}^n} \frac{p(\exp(\tilde{\mathbf{y}}), \alpha \exp(\tilde{\mathbf{z}}))}{\exp(\tilde{\mathbf{y}})^\alpha \exp(\tilde{\mathbf{z}})^\alpha}. \quad (25)$$

We use Sion's minimax theorem. We state a weaker form of this theorem below, which will be enough for our purposes; for the original stronger version, look at [Sio58].

**Theorem 5.3** (Immediate corollary of [Sio58]). *Let  $A$  be a compact convex subset of  $\mathbb{R}^n$  and  $B$  a (not necessarily compact) convex subset of  $\mathbb{R}^m$ . Suppose that  $f : A \times B \rightarrow \mathbb{R}$  has the following properties:*

- For each  $a \in A$ , the function  $f(a, \cdot) : B \rightarrow \mathbb{R}$  is continuous and quasi-convex.
- For each  $b \in B$ , the function  $f(\cdot, b) : A \rightarrow \mathbb{R}$  is continuous and quasi-concave.

Then we have

$$\sup_{a \in A} \inf_{b \in B} f(a, b) = \inf_{b \in B} \sup_{a \in A} f(a, b).$$

Recall that a real-valued function  $f$  is quasi-convex if  $f^{-1}(-\infty, t]$  is convex for any  $t \in \mathbb{R}$  and  $f$  is quasi-concave if  $-f$  is quasi-convex.

We will apply [Theorem 5.3](#) to the following function.

$$f(\alpha, (\tilde{\mathbf{y}}, \tilde{\mathbf{z}})) := \frac{p(\exp(\tilde{\mathbf{y}}), \alpha \exp(\tilde{\mathbf{z}}))}{\exp(\tilde{\mathbf{y}})^\alpha \exp(\tilde{\mathbf{z}})^\alpha}.$$

For a fixed  $\alpha \geq 0$ , the function  $f$  is log-convex, and therefore quasi-convex, in  $\tilde{\mathbf{y}}, \tilde{\mathbf{z}}$  because of [Fact 2.11](#). For a fixed  $\tilde{\mathbf{y}}, \tilde{\mathbf{z}}$ , the function  $f$  is log-concave, and therefore quasi-concave, in  $\alpha$  because of [Corollary 2.10](#). It is easy to see that  $f$  is continuous in each variable as well. Therefore, by [Theorem 5.3](#) we have

$$\begin{aligned} \sup_{\alpha \in [0, \deg p]^n} \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} \in \mathbb{R}^n} \frac{p(\exp(\tilde{\mathbf{y}}), \alpha \exp(\tilde{\mathbf{z}}))}{\exp(\tilde{\mathbf{y}})^\alpha \exp(\tilde{\mathbf{z}})^\alpha} &= \inf_{\tilde{\mathbf{y}}, \tilde{\mathbf{z}} \in \mathbb{R}^n} \sup_{\alpha \in [0, \deg p]^n} \frac{p(\exp(\tilde{\mathbf{y}}), \alpha \exp(\tilde{\mathbf{z}}))}{\exp(\tilde{\mathbf{y}})^\alpha \exp(\tilde{\mathbf{z}})^\alpha} \\ &\geq \inf_{\tilde{\mathbf{y}} \in \mathbb{R}^n} \sup_{\alpha \in [0, \deg p]^n} \inf_{\tilde{\mathbf{z}} \in \mathbb{R}^n} \frac{p(\exp(\tilde{\mathbf{y}}), \alpha \exp(\tilde{\mathbf{z}}))}{\exp(\tilde{\mathbf{y}})^\alpha \exp(\tilde{\mathbf{z}})^\alpha}. \end{aligned}$$

This proves (25) and finishes the proof of [Lemma 5.2](#).

## References

- [AO15a] Nima Anari and Shayan Oveis Gharan. Effective-Resistance-Reducing Flows, Spectrally Thin Trees, and Asymmetric TSP. In *FOCS*, pages 20–39, 2015. [2](#)
- [AO15b] Nima Anari and Shayan Oveis Gharan. The kadison-singer problem for strongly rayleigh measures and applications to asymmetric tsp. 2015. [2](#)
- [AOSS17] Nima Anari, Shayan Oveis Gharan, Amin Saberi, and Mohit Singh. Nash social welfare, matrix permanent, and stable polynomials. In *ITCS*, 2017. to appear. [11](#)

- [BB10] J. Borcea and P. Brädén. Multivariate Pólya-Schur classification problems in the Weyl algebra. *Proceedings of the London Mathematical Society*, 101(3):73–104, 2010. 5
- [BBL09] Julius Borcea, Petter Branden, and Thomas M. Liggett. Negative dependence and the geometry of polynomials. *Journal of American Mathematical Society*, 22:521–567, 2009. 4, 5, 6, 11
- [BC95] André Bouchet and William H. Cunningham. Delta-matroids, jump systems, and bisubmodular polyhedra. *SIAM J. Discrete Math.*, 8(1):17–32, 1995. 8
- [Brä07] Petter Brädén. Polynomials with the half-plane property and matroid theory. *Advances in Mathematics*, 216(1):302–320, 2007. 5, 7
- [BV06] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2006. 6
- [Csi14] Péter Csikvári. Lower matching conjecture, and a new proof of schrijver’s and gurvits’s theorems. arXiv preprint arXiv:1406.0766, 2014. 9
- [Gra02] JH Grace. The zeros of a polynomial. In *Proc. Cambridge Philos. Soc*, volume 11, pages 352–357, 1902. 5
- [GS14] Leonid Gurvits and Alex Samorodnitsky. Bounds on the permanent and some applications. In *FOCS*, pages 90–99. IEEE Computer Society, 2014. 9
- [Gül97] Osman Güler. Hyperbolic polynomials and interior point methods for convex programming. *MOR*, 22(2):350–77, 1997. 6
- [Gur06] Leonid Gurvits. Hyperbolic polynomials approach to van der waerden/schrijver-valiant like conjectures: Sharper bounds, simpler proofs and algorithmic applications. In *STOC*, STOC ’06, pages 417–426, 2006. 2
- [Gur11] L. Gurvits. Unleashing the power of schrijvers permanental inequality with the help of the bethe approximation. arXiv:1106.2844, 2011. 9
- [KT12] Alex Kulesza and Ben Taskar. Determinantal point processes for machine learning. *arXiv preprint arXiv:1207.6083*, 2012. 11
- [Lel17] Marc Lelarge. Counting matchings in irregular bipartite graphs. In *SODA*, 2017. to appear. 9
- [Lov97] László Lovász. The membership problem in jump systems. *J. Combin. Theory Ser. B*, 70(1):45–66, 1997. 7
- [MF08] S. Thomas McCormick and Satoru Fujishige. Strongly polynomial and fully combinatorial algorithms for bisubmodular function minimization. In *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 44–53. ACM, New York, 2008. 8, 9
- [MSS13] Adam Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing Families II: Mixed Characteristic Polynomials and the Kadison-Singer Problem. 2013. 2

- [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, 2016. [1](#), [3](#), [13](#)
- [Sch98] Alexander Schrijver. Counting 1-factors in regular bipartite graphs. *Journal of Combinatorial Theory, Series B*, 72(1):122–135, 1998. [9](#)
- [Sio58] Maurice Sion. On general minimax theorems. *Pacific J. Math.*, 8:171–176, 1958. [26](#)
- [SV16] Damian Straszak and Nisheeth K. Vishnoi. Real stable polynomials and matroids: Optimization and counting. 2016. [3](#)
- [Sze22] G Szegő. Bemerkungen zu einem satz von jh grace über die wurzeln algebraischer gleichungen. *Mathematische Zeitschrift*, 13(1):28–55, 1922. [5](#)
- [Wal22] JL Walsh. On the location of the roots of certain types of polynomials. *Transactions of the American Mathematical Society*, 24(3):163–180, 1922. [5](#)