A PTAS FOR THE MINIMIZATION OF POLYNOMIALS OF FIXED DEGREE OVER THE SIMPLEX

ETIENNE DE KLERK†, MONIQUE LAURENT‡, AND PABLO PARRILO

ABSTRACT. We consider the problem of computing the minimum value $p_{\min}$ taken by a polynomial $p(x)$ of degree d over the standard simplex Δ. This is an NP-hard problem already for degree $d = 2$. For any integer $k \geq 1$, by minimizing $p(x)$ over the set of rational points in Δ with denominator k, one obtains a hierarchy of upper bounds $p_{\Delta(k)}$ converging to $p_{\min}$ as $k \longrightarrow \infty$. These upper approximations are intimately linked to a hierarchy of lower bounds for $p_{\min}$ constructed via Pólya's theorem about representations of positive forms on the simplex. Revisiting the proof of Pólya's theorem allows us to give estimates on the quality of these upper and lower approximations for $p_{\min}$. Moreover, we show that the bounds $p_{\Delta(k)}$ yield a polynomial time approximation scheme for the minimization of polynomials of fixed degree d on the simplex, extending an earlier result of Bomze and De Klerk for degree $d = 2$.

March 30, 2004

1. Introduction

1.1. **A grid approximation.** We consider the problem of minimizing a polynomial $p(x)$ of degree d on the standard simplex

$$\Delta := \{x \in \mathbb{R}^n_+ \mid \sum_{i=1}^n x_i = 1\};$$

that is, the problem of computing

$$p_{\min} := \min_{x \in \Delta} \; p(x). \tag{1.1}$$

One may assume w.l.o.g. that $p(x)$ is a homogeneous polynomial (form). Indeed, as observed in [5], if $p(x) = \sum_{\ell=0}^d p_\ell(x)$, where $p_\ell(x)$ is homogeneous of degree ℓ, then minimizing $p(x)$ over Δ is equivalent to minimizing the degree d form $p'(x) := \sum_{\ell=0}^d p_\ell(x)(\sum_{i=1}^n x_i)^{d-\ell}$. Problem (1.1) is an NP-hard problem, already for forms of degree $d = 2$, as it contains the maximum stable set problem. Indeed, for a graph G with adjacency matrix A, the maximum size $\alpha(G)$ of a stable set in G can be expressed as

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta} x^T(I + A)x$$

1991 *Mathematics Subject Classification.* 90C22, 90C26, 49M20.

Key words and phrases. Semidefinite programming, global optimization, positive polynomial, sum of squares of polynomials, approximation algorithm.

† Supported by the Netherlands Organisation for Scientific Research grant NWO 613.000.214.

‡ Supported by the Netherlands Organisation for Scientific Research grant NWO 639.032.203.

by the theorem of Motzkin and Straus [7]. Given an integer $k \geq 1$, let

$$\Delta(k) := \{x \in \Delta \mid kx \in \mathbb{Z}^n\} \tag{1.2}$$

denote the set of rational points in Δ with denominator k and define

$$p_{\Delta(k)} := \min p(x) \text{ s.t. } x \in \Delta(k). \tag{1.3}$$

Thus, $p_{\min} \leq p_{\Delta(k)}$ for any $k \geq 1$. As $|\Delta(k)| = \binom{n+k-1}{k}$, one can compute the bound $p_{\Delta(k)}$ in polynomial time for any *fixed* k. Set

$$p_{\max} := \max_{x \in \Delta} p(x). \tag{1.4}$$

When $p(x)$ is a form of degree $d = 2$, Bomze and De Klerk [3] show that the following inequality holds:

$$p_{\Delta(k)} - p_{\min} \leq \frac{1}{k}(p_{\max} - p_{\min}) \tag{1.5}$$

for any $k \geq 1$. Nesterov [9] subsequently showed:

$$p_{\Delta(k)} - p_{\min} \leq \frac{1}{k}(\max_{i=1,\ldots,n} p_{2e_i} - p_{\min}), \tag{1.6}$$

where $e_1, \ldots, e_n$ denote the standard unit vectors. This inequality is stronger since, by evaluating $p(x)$ at e_i, one finds that $p(e_i) = p_{2e_i} \leq p_{\max}$. Bomze and De Klerk's result is based on an intimate link existing between the upper approximations $p_{\Delta(k)}$ and a hierarchy of lower bounds for $p_{\min}$ constructed using a result of Pólya about the representation of positive forms on the simplex. (This approach also permits, in fact, to establish the stronger inequality (1.6).) Nesterov's approach is different and uses a probabilistic argument for justifying the estimation (1.6). We describe both approaches in Section 2.1 below.

Using his probabilistic approach, Nesterov [9] proves the following result for polynomials of higher degree. Assume that $p(x)$ is a form of degree d which is a sum of square-free monomials; that is, a monomial x^α appears with a nonzero coefficient in $p(x)$ only if $\alpha_i \leq 1$ for all $i = 1, \ldots, n$. Then, for $k \geq d$,

$$p_{\Delta(k)} - p_{\min} \leq \left(1 - \frac{k!}{(k-d)!k^d}\right)(-p_{\min}) \leq \frac{d(d-1)}{2k}(-p_{\min}). \tag{1.7}$$

In this paper, we prove some estimates for the approximation $p_{\Delta(k)}$ for general degree d forms. For a polynomial $p(x) = \sum_\alpha p_\alpha x^\alpha$, define as in [13] the parameter:

$$L_p := \max_\alpha |p_\alpha| \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!} \tag{1.8}$$

and the following parameter, introduced in the next subsection:

$$p_{\max}^{(0)} := \max_\alpha p_\alpha \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!}. \tag{1.9}$$

Then, $p_{\max} \leq p_{\max}^{(0)} \leq L_p$. We show the inequalities:

$$p_{\Delta(k)} - p_{\min} \leq \left(1 - \frac{k!}{(k-d)!k^d}\right)(p_{\max}^{(0)} - p_{\min}) \leq \frac{d(d-1)}{2k}(p_{\max}^{(0)} - p_{\min}), \tag{1.10}$$

$$p_{\max}^{(0)} - p_{\min} \leq \binom{2d-1}{d} d^d (p_{\max} - p_{\min}), \tag{1.11}$$

which imply:

$$p_{\Delta(k)} - p_{\min} \le \frac{d^d}{k}\binom{d}{2}\binom{2d-1}{d}(p_{\max} - p_{\min}). \tag{1.12}$$

Our argument for (1.10) follows closely the proof given by Powers and Reznick [13] for Pólya's theorem, and our proof for (1.11) uses some tools of Reznick [15] about powers of linear forms. As an application of (1.12), the approximations $p_{\Delta(k)}$ yield a polynomial time approximation scheme (PTAS) for the problem of minimizing a form of degree d on the simplex, for any fixed d.

A more detailed analysis permits to show sharper estimates in some cases. For instance, when $d = 2$, relations (1.5) and (1.6) hold and, in the case $d = 3$, we can show that

$$p_{\Delta(k)} - p_{\min} \le \frac{4}{k}(p_{\max} - p_{\min}). \tag{1.13}$$

1.2. **Pólya's representation theorem for positive forms on the simplex.** We recall Pólya's result about positive forms on the simplex.

Theorem 1.1. *Let p be a form of degree d which is positive on the simplex Δ, i.e., $p_{\min} > 0$. Then, the polynomial $\left(\sum_{i=1}^n x_i\right)^r p(x)$ has nonnegative coefficients for all r satisfying*

$$r \ge \binom{d}{2}\frac{p_{\max}^{(0)}}{p_{\min}} - d. \tag{1.14}$$

Pólya [12] proved that $\left(\sum_{i=1}^n x_i\right)^r p(x)$ has nonnegative coefficients for r large enough; Powers and Reznick [13] proved that this holds for any $r \ge \binom{d}{2}\frac{L_p}{p_{\min}} - d$; we observe here that this holds for any r satisfying the weaker condition (1.14) (with L_p replaced by $p_{\max}^{(0)}$). We review the proof of Theorem 1.1 in Section 2.1, as it permits, moreover, to estimate the quality of the bound $p_{\Delta(k)}$ for $p_{\min}$.

For now, let us indicate how Pólya's result can be used for constructing an asymptotically converging hierarchy of lower bounds for $p_{\min}$. Observe first that $p_{\min}$ can alternatively be formulated as the maximum scalar λ for which $p(x) - \lambda \ge 0$ for all $x \in \Delta$. Equivalently,

$$p_{\min} = \max\ \lambda \text{ such that } p(x) - \lambda\left(\sum_{i=1}^n x_i\right)^d \ge 0\ \forall x \in \mathbb{R}^n_+. \tag{1.15}$$

For any integer $r \ge 0$, define the parameter:

$$p_{\min}^{(r)} := \max\ \lambda \text{ s.t. } \quad \begin{array}{c}\text{the polynomial } \left(\sum_{i=1}^n x_i\right)^r \left(p(x) - \lambda\left(\sum_{i=1}^n x_i\right)^d\right) \\ \text{has nonnegative coefficients.}\end{array} \tag{1.16}$$

Obviously,

$$p_{\min}^{(r)} \le p_{\min}^{(r+1)} \le p_{\min}.$$

Moreover, $p_{\min}^{(r)} \ge 0$ if and only if the polynomial $\left(\sum_{i=1}^n x_i\right)^r p(x)$ has nonnegative coefficients. Hence, Theorem 1.1 asserts that $p_{\min}^{(r)} \ge 0$ for any r satisfying (1.14). The bound $p_{\min}^{(r)}$ can be computed in polynomial time for any *fixed* r, as it can be expressed as the minimum over the grid $\Delta(r + d)$ of a perturbation of the polynomial $p(x)$; see (2.3). As a consequence of Pólya's theorem, the bounds $p_{\min}^{(r)}$ converge asymptotically to $p_{\min}$ as $r \longrightarrow \infty$. This idea of using Pólya's result for constructing converging approximations goes back to the work of Parrilo [10, 11], who used it for constructing hierarchies of conic relaxations for the cone of copositive matrices (corresponding to degree 2 positive semidefinite forms). The construction was extended to general positive semidefinite forms by Faybusovich [5], Zuluaga et al. [18].

For the problem (1.4) of maximizing $p(x)$ over Δ, one can analogously define the bounds:

$$p^{(r)}_{\max} := \min\ \lambda \text{ s.t. } \begin{array}{l}\text{the polynomial } \left(\sum_{i=1}^n x_i\right)^r \left(\lambda \left(\sum_{i=1}^n x_i\right)^d - p(x)\right) \\ \text{has nonnegative coefficients,}\end{array} \tag{1.17}$$

satisfying:

$$p_{\max} \le p^{(r+1)}_{\max} \le p^{(r)}_{\max} \quad \text{for } r \ge 0.$$

As we see in Section 2.1, the following holds:

$$p^{(0)}_{\min} = \min_{\alpha} p_\alpha \frac{\alpha_1! \cdots \alpha_n!}{d!},\ p^{(0)}_{\max} = \max_{\alpha} p_\alpha \frac{\alpha_1! \cdots \alpha_n!}{d!}, \tag{1.18}$$

which justifies the definition of $p^{(0)}_{\max}$ given earlier in (1.9).

1.3. **Quality of the upper and lower approximations.** Let p be a form of degree d and let $p_{\Delta(k)}$, $p_{\max}$, L_p, $p^{(r)}_{\min}$, and $p^{(r)}_{\max}$, be as defined in (1.3), (1.4), (1.8), (1.16), and (1.17). Therefore,

$$p^{(r)}_{\min} \le p_{\min} \le p_{\Delta(r+d)} \le p_{\max} \le p^{(r)}_{\max}.$$

Finally, define the parameter

$$w_r(d) := \frac{(r+d)!}{r!(r+d)^d} = \prod_{i=1}^{d-1} \left(1 - \frac{i}{r+d}\right). \tag{1.19}$$

One can verify that

$$1 - \binom{d}{2} \frac{1}{r+d} \le w_r(d) \le 1, \tag{1.20}$$

which implies that $\lim_{r\to\infty} w_r(d) = 1$. For this, note that $\prod_{i=1}^m (1-x_i) \ge 1 - \sum_{i=1}^m x_i$ if $0 \le x_i \le 1$ for $i = 1, \ldots, m$. (Use induction on $m \ge 1$.)

The following result estimates the quality of the approximations $p^{(r)}_{\min}$ and $p_{\Delta(r+d)}$ for $p_{\min}$. Faybusovich [5] proved the weaker version of the inequality (1.21) where the parameter $p^{(0)}_{\max}$ is replaced by L_p; (1.22) coincides with (1.10) with $k = r + d$.

Theorem 1.2. *Let p be a form of degree d and $r \ge 0$ an integer. Then,*

$$p_{\min} - p^{(r)}_{\min} \le (p^{(0)}_{\max} - p_{\min}) \left(\frac{1}{w_r(d)} - 1\right), \tag{1.21}$$

$$p_{\Delta(r+d)} - p_{\min} \le (p^{(0)}_{\max} - p_{\min})(1 - w_r(d)). \tag{1.22}$$

The next two results give analogous inequalities involving the parameter $p_{\max}$ instead of $p^{(0)}_{\max}$ in the case when $d = 2, 3$.

Theorem 1.3. *Let p be a form of degree $d = 2$ and $r \geq 0$ an integer. Then,*

$$p_{\min} - p_{\min}^{(r)} \leq \frac{1}{r+1} \left(\max_{i=1,\ldots,n} p_{2e_i} - p_{\min} \right) \leq \frac{1}{r+1} (p_{\max} - p_{\min}), \tag{1.23}$$

$$p_{\Delta(r+2)} - p_{\min} \leq \frac{1}{r+2} \left(\max_{i=1,\ldots,n} p_{2e_i} - p_{\min} \right) \leq \frac{1}{r+2} (p_{\max} - p_{\min}). \tag{1.24}$$

The upper bounds in (1.23), (1.24) involving $p_{\max}$ are proved by Bomze and De Klerk [3] and the upper bound in (1.24) involving $\max_i p_{2e_i}$ is proved by Nesterov [9]. We show in Section 2.1 the following extension for forms of degree 3.

Theorem 1.4. *Let p be a form of degree $d = 3$ and $r \geq 0$ an integer. Then,*

$$p_{\min} - p_{\min}^{(r)} \leq \frac{4(r+3)}{(r+1)(r+2)} (p_{\max} - p_{\min}), \tag{1.25}$$

$$p_{\Delta(r+3)} - p_{\min} \leq \frac{4}{r+3} (p_{\max} - p_{\min}). \tag{1.26}$$

The results from Theorems 1.2, 1.3 and 1.4 can be proved by a close inspection of the proof of Pólya's theorem; see Section 2.1. For forms of arbitrary degree d, we can show the following result, whose proof needs an additional argument and will be given in Section 2.2.

Theorem 1.5. *The following holds for a form $p(x)$ of degree d.*

$$p_{\max}^{(0)} - p_{\min}^{(0)} \leq \binom{2d-1}{d} d^d (p_{\max} - p_{\min}). \tag{1.27}$$

Combined with Theorem 1.2, this implies:

Theorem 1.6. *Let $p(x)$ be a form of degree d and $r \geq 0$ an integer. Then,*

$$p_{\min} - p_{\min}^{(r)} \leq \left(\frac{1}{w_r(d)} - 1 \right) \binom{2d-1}{d} d^d (p_{\max} - p_{\min}), \tag{1.28}$$

$$p_{\Delta(r+d)} - p_{\min} \leq (1 - w_r(d)) \binom{2d-1}{d} d^d (p_{\max} - p_{\min}). \tag{1.29}$$

As an application, the grid approximations $p_{\Delta(k)}$ ($k \geq d$) provide a polynomial time approximation scheme for the problem of minimizing a form of degree d on the simplex. See Section 3 for details.

2. Approximating Forms on the Simplex

2.1. **Estimating the upper and lower approximations $p_{\Delta(r+d)}$ and $p_{\min}^{(r)}$ for $p_{\min}$ via Pólya's theorem.** We will use the following notation. Given $\alpha \in \mathbb{N}^n$, set

$$\alpha! := \alpha_1! \cdots \alpha_n!$$

and, following [13], given scalars x, t and a nonnegative integer m, set

$$(x)_t^m := x(x-t)\cdots(x-(m-1)t) = \prod_{i=0}^{m-1} (x - it).$$

Then, $(1)^d_{\frac{1}{r+d}} = w_r(d)$, the parameter defined in (1.19). Define

$$I(n,m) := \{\alpha \in \mathbb{N}^n : |\alpha| = \sum_{i=1}^n \alpha_i = m\}.$$

We use the multinomial identity:

$$\left(\sum_{i=1}^n x_i\right)^m = \sum_{\alpha \in I(n,m)} \frac{m!}{\alpha!} x^\alpha \tag{2.1}$$

and its generalization, known as the Vandermonde-Chu identity:

$$\left(\sum_{i=1}^n x_i\right)^m_t = \sum_{\alpha \in I(n,m)} \frac{m!}{\alpha!} (x_1)^{\alpha_1}_t \cdots (x_n)^{\alpha_n}_t. \tag{2.2}$$

(See [13] for a proof. Alternatively, use induction on $m \geq 1$.)

In what follows, $p(x)$ is a form of degree d and $r \geq 0$ is an integer. By definition, $p^{(r)}_{\min}$ is the maximum scalar λ for which the polynomial $\left(\sum_i x_i\right)^r p(x) - \lambda \left(\sum_i x_i\right)^{r+d}$ has nonnegative coefficients. We begin with evaluating the coefficients of this polynomial. We have:

$$\left(\sum_{i=1}^n x_i\right)^{r+d} = \sum_{\beta \in I(n,r+d)} \frac{(r+d)!}{\beta!} x^\beta,$$

$$p(x)\left(\sum_{i=1}^n x_i\right)^r = \sum_{\beta \in I(n,r+d)} A_\beta x^\beta$$

where

$$A_\beta := \sum_{\alpha \in I(n,d),\ \alpha \leq \beta} \frac{r!}{(\beta-\alpha)!} p_\alpha = \frac{r!(r+d)^d}{\beta!} \sum_{\alpha \in I(n,d)} p_\alpha \prod_{i=1}^n \left(\frac{\beta_i}{r+d}\right)^{\alpha_i}_{\frac{1}{r+d}}.$$

Therefore, $p^{(r)}_{\min}$ is the maximum λ for which $A_\beta - \lambda \frac{(r+d)!}{\beta!} \geq 0$ for all $\beta \in I(n, r+d)$; that is,

$$\begin{aligned} p^{(r)}_{\min} &= \min_{\beta \in I(n,r+d)} \frac{\beta!}{(r+d)!} A_\beta \\ &= \min_{\beta \in I(n,r+d)} \frac{1}{w_r(d)} \sum_{\alpha \in I(n,d)} p_\alpha \prod_{i=1}^n \left(\frac{\beta_i}{r+d}\right)^{\alpha_i}_{\frac{1}{r+d}}. \end{aligned} \tag{2.3}$$

As the point $x := \frac{\beta}{r+d}$ belongs to $\Delta(r+d)$, it follows that

$$p^{(r)}_{\min} = \min_{x \in \Delta(r+d)} \frac{1}{w_r(d)} \sum_{\alpha \in I(n,d)} p_\alpha (x_1)^{\alpha_1}_{\frac{1}{r+d}} \cdots (x_n)^{\alpha_n}_{\frac{1}{r+d}}. \tag{2.4}$$

As in [13], define the polynomial

$$\begin{aligned} \phi(x) &:= p(x) - \sum_{\alpha \in I(n,d)} p_\alpha (x_1)^{\alpha_1}_{\frac{1}{r+d}} \cdots (x_n)^{\alpha_n}_{\frac{1}{r+d}} \\ &= \sum_{\alpha \in I(n,d)} p_\alpha \left(x^\alpha - (x_1)^{\alpha_1}_{\frac{1}{r+d}} \cdots (x_n)^{\alpha_n}_{\frac{1}{r+d}}\right) \end{aligned} \tag{2.5}$$

and set

$$\phi_{\max} := \max_{x\in\Delta(r+d)} \phi(x).$$

Then,

$$p_{\min}^{(r)} = \frac{1}{w_r(d)} \min_{x\in\Delta(r+d)} (p(x) - \phi(x)). \tag{2.6}$$

This implies:

$$p_{\min}^{(r)} \geq \frac{1}{w_r(d)} (p_{\Delta(r+d)} - \phi_{\max})$$

and thus, as $p_{\Delta(r+d)} \geq p_{\min} \geq p_{\min}^{(r)}$,

$$p_{\min}^{(r)} \geq \frac{1}{w_r(d)} (p_{\min} - \phi_{\max}), \tag{2.7}$$

$$p_{\Delta(r+d)} \leq w_r(d) p_{\min} + \phi_{\max}. \tag{2.8}$$

Therefore,

$$p_{\min} - p_{\min}^{(r)} \leq \left(1 - \frac{1}{w_r(d)}\right) p_{\min} + \frac{1}{w_r(d)} \phi_{\max}, \tag{2.9}$$

$$p_{\Delta(r+d)} - p_{\min} \leq w_r(d) \left(\left(1 - \frac{1}{w_r(d)}\right) p_{\min} + \frac{1}{w_r(d)} \phi_{\max} \right). \tag{2.10}$$

We can now prove the results from Theorems 1.1, 1.2, 1.3, 1.4, and relation (1.7).

2.1.1. *Proving Theorem 1.2.* As $x^\alpha \geq \prod_{i=1}^n (x_i)_{\frac{1}{r+d}}^{\alpha_i}$ and $p_\alpha \leq p_{\max}^{(0)} \frac{d!}{\alpha!}$ (by the definition (1.9) of $p_{\max}^{(0)}$), we find that

$$\phi(x) \leq p_{\max}^{(0)} \sum_{\alpha\in I(n,d)} \frac{d!}{\alpha!} \left(x^\alpha - \prod_{i=1}^n (x_i)_{\frac{1}{r+d}}^{\alpha_i} \right).$$

In view of relations (2.1) and (2.2), the right hand side is equal to

$$p_{\max}^{(0)} \left(\left(\sum_i x_i \right)^d - \left(\sum_i x_i \right)_{\frac{1}{r+d}}^d \right) = p_{\max}^{(0)} (1 - w_r(d)).$$

Therefore,

$$\phi_{\max} \leq p_{\max}^{(0)} (1 - w_r(d)). \tag{2.11}$$

This inequality, combined with (2.9) and (2.10), gives the inequalities (1.21) and (1.22) from Theorem 1.2.

2.1.2. *Proving Theorem 1.1.* Assume that $p_{\min} > 0$ and $r \geq \binom{d}{2} \frac{p_{\max}^{(0)}}{p_{\min}} - d$. Then, relation (2.11) combined with the bound on $w_r(d)$ from (1.20) implies that $\phi_{\max} \leq p_{\max}^{(0)} \binom{d}{2} \frac{1}{r+d}$. Now, (2.7) implies that $p_{\min}^{(r)} \geq \frac{1}{w_r(d)} \left(p_{\min} - p_{\max}^{(0)} \binom{d}{2} \frac{1}{r+d} \right)$, which is nonnegative by our assumption on r. This shows that $p_{\min}^{(r)} \geq 0$ for such r; that is, Theorem 1.1 holds.

2.1.3. *Proving the inequality (1.7) for polynomials that are sums of square-free monomials.* Assume that $p_\alpha = 0$ whenever $\alpha_i \geq 2$ for some $i = 1, \ldots, n$. Then the polynomial $\phi(x)$ from (2.5) is identically zero. Thus $\phi_{\max} = 0$ and the estimate (1.7) follows directly from (2.10) (with $k = r + d$ and using (1.20)).

2.1.4. *Proving Theorem 1.3 for degree 2 forms.* By looking more closely at the form of the function $\phi(x)$, one can give an upper bound for $\phi_{\max}$ depending on $p_{\max}$ and $p_{\min}$ only. Indeed, when $d = 2$, one can verify that

$$\phi(x) = \frac{1}{r+2} \sum_i p_{2e_i} x_i.$$

Therefore,

$$\phi_{\max} = \frac{1}{r+2} \max_{i=1,\ldots,n} p_{2e_i} \leq \frac{1}{r+2} p_{\max}. \tag{2.12}$$

As $w_r(2) = \frac{r+1}{r+2}$, together with (2.9) and (2.10), this implies that the inequalities (1.23) and (1.24) from Theorem 1.3 hold.

Moreover, $p_{\min}^{(r)} \geq 0$ for $r \geq \frac{\max_i p_{2e_i}}{p_{\min}} - 2$. That is, for degree 2 forms, Theorem 1.1 remains valid for such r (instead of (1.14)).

2.1.5. *Proving Theorem 1.4 for degree 3 forms.* When $d = 3$, one can verify that

$$\phi(x) = \sum_{i=1}^n p_{3e_i}(3tx_i^2 - 2t^2 x_i) + \sum_{1 \leq i < j \leq n} (p_{2e_i+e_j} + p_{e_i+2e_j}) t x_i x_j, \tag{2.13}$$

after setting $t := \frac{1}{r+3}$. Evaluating p at the simplex points e_i and $\frac{1}{2}(e_i + e_j)$ yields, respectively, the relations:

$$p_{\min} \leq p(e_i) = p_{3e_i} \leq p_{\max}, \tag{2.14}$$

$$p_{3e_i} + p_{3e_j} + p_{2e_i+e_j} + p_{e_i+2e_j} \leq 8p_{\max}. \tag{2.15}$$

Using (2.15), we can bound the second sum in (2.13):

$$\begin{aligned} \textstyle\sum_{i<j} (p_{2e_i+e_j} + p_{e_i+2e_j}) x_i x_j &\leq \textstyle\sum_{i<j} (8p_{\max} - p_{3e_i} - p_{3e_j}) x_i x_j \\ &= 8p_{\max} \textstyle\sum_{i<j} x_i x_j - \sum_i p_{3e_i} x_i (1 - x_i). \end{aligned}$$

Therefore,

$$\phi(x) \leq 8tp_{\max} \sum_{i<j} x_i x_j + 4t \sum_i p_{3e_i} x_i^2 - t(1+2t) \sum_i p_{3e_i} x_i.$$

Using the fact that $p_{3e_i} \leq p_{\max}$ and $\sum_i x_i = 1$, the sum of the first two terms can be bounded by $4tp_{\max}$. Using the fact that $-p_{3e_i} \leq -p_{\min}$, the third term can be bounded by $-t(1+2t)p_{\min} = -\frac{r+5}{(r+3)^2} p_{\min}$. This shows:

$$\phi_{\max} \leq \frac{4}{r+3} p_{\max} - \frac{r+5}{(r+3)^2} p_{\min}. \tag{2.16}$$

Together with (2.9), (2.10), and the fact that $w_r(3) = \frac{(r+1)(r+2)}{(r+3)^2}$, this implies that the relations (1.25) and (1.26) from Theorem 1.4 hold.

To conclude, let us mention that it is not clear whether this type of argument for bounding $\phi_{\max}$ in terms of $p_{\max}$ and $p_{\min}$ extends for forms of degree 4. We use in the next subsection a different argument for dealing with the general case $d \geq 4$.

2.2. **Estimating the maximum coefficient range of a polynomial - Proof of Theorem 1.5.** In this section we prove the estimate of $p^{(0)}_{\max} - p^{(0)}_{\min}$ in terms of $p_{\max} - p_{\min}$ given in Theorem 1.5. Following Reznick [15], let us introduce the following definitions that will be useful for us.

Recall that $I(n,d) = \{\alpha \in \mathbb{Z}^n_+ : |\alpha| = d\}$ and let $F_{n,d}$ denote the set of forms of degree d in n variables. For $p \in F_{n,d}$, write

$$p(x) = \sum_{\alpha \in I(n,d)} p_\alpha x^\alpha = \sum_{\alpha \in I(n,d)} a(p,\alpha)\frac{d!}{\alpha!}x^\alpha,$$

after setting

$$a(p,\alpha) := p_\alpha \frac{\alpha!}{d!} \ \text{ for } \alpha \in I(n,d).$$

For $\alpha \in \mathbb{R}^n$, define the degree d form:

$$P_\alpha(x) := (\alpha^T x)^d \ \text{ for } x \in \mathbb{R}^n.$$

Define the inner product on $F_{n,d}$:

$$\langle p, q\rangle := \sum_{\alpha \in I(n,d)} a(p,\alpha)a(q,\alpha)\frac{d!}{\alpha!} \ \text{ for } p,q \in F_{n,d}.$$

As $P_\alpha(x) = \sum_{\beta \in I(n,d)} \frac{d!}{\beta!}\alpha^\beta x^\beta$, it follows that, for any $p \in F_{n,d}$,

$$\langle p, P_\alpha\rangle = p(\alpha) \ \text{ for } \alpha \in \mathbb{R}^n. \tag{2.17}$$

Moreover,

$$\langle p, x^\alpha\rangle = a(p,\alpha) \ \text{ for } \alpha \in I(n,d). \tag{2.18}$$

Finally, given $\alpha \in I(n,d)$, define the polynomials:

$$h_\alpha(x) := \prod_{j=1}^{n}\prod_{\ell_j=0}^{\alpha_j-1}(dx_j - \ell_j(x_1 + \ldots + x_n)), \quad h^*_\alpha(x) := \frac{1}{\alpha! d^d}h_\alpha(x). \tag{2.19}$$

Lemma 2.1. [15] *For $\alpha, \alpha' \in I(n,d)$, $\langle h^*_\alpha, P_{\alpha'}\rangle = 1$ if $\alpha = \alpha'$ and 0 otherwise.*

Proof. Direct verification. □

Corollary 2.2. *(Biermann's theorem, see* [15, §2]*)*
The set $\{P_\alpha \mid \alpha \in I(n,d)\}$ is a basis of the vector space $F_{n,d}$.

Let A be the $|I(n,d)| \times |I(n,d)|$ matrix permitting to express the monomial basis $\{x^\alpha \mid \alpha \in I(n,d)\}$ in terms of the basis $\{P_\beta \mid \beta \in I(n,d)\}$. That is,

$$x^\alpha = \sum_{\beta \in I(n,d)} A(\alpha,\beta)P_\beta(x). \tag{2.20}$$

For $p \in F_{n,d}$, by taking the inner product in (2.20) with p and using (2.17) and (2.18), we find:

$$a(p,\alpha) = \sum_{\beta \in I(n,d)} A(\alpha,\beta)p(\beta) \ \text{ for } \alpha \in I(n,d). \tag{2.21}$$

Taking the inner product in (2.20) with h^*_β, we find:

$$a(h^*_\beta, \alpha) = \langle h^*_\beta, x^\alpha\rangle = A(\alpha,\beta) \ \text{ for } \alpha,\beta \in I(n,d). \tag{2.22}$$

In view of (1.18), our parameters $p^{(0)}_{\max}$ and $p^{(0)}_{\min}$ are given by:

$$p^{(0)}_{\max} = \max_{\alpha \in I(n,d)} a(p,\alpha),\ p^{(0)}_{\min} = \min_{\alpha \in I(n,d)} a(p,\alpha).$$

Define the vectors $x := (p(\alpha))_{\alpha \in I(n,d)}$ and $y := (a(p,\alpha))_{\alpha \in I(n,d)}$. In view of (2.9), they are related by the relation:

$$y = Ax. \tag{2.23}$$

Denote by $x_{\max}$ (resp., $x_{\min}$) the largest entry of x; similarly for y. Thus,

$$y_{\max} - y_{\min} = p^{(0)}_{\max} - p^{(0)}_{\min}.$$

For $\alpha \in I(n,d)$, $p(\alpha) = p(\alpha/d)d^d$. Thus,

$$x_{\max} - x_{\min} \le d^d(p_{\max} - p_{\min}). \tag{2.24}$$

Our strategy for proving Theorem 1.5 consists of showing the following two results:

Proposition 2.3. *Let A be an $N \times N$ matrix satisfying $Ae = \mu e$ for some scalar μ, where e denotes the all-ones vector, and set*

$$\|A\|_\infty := \max_{i=1,\dots,N} \sum_{k=1}^{N} |A(i,k)|.$$

If $y = Ax$, then $y_{\max} - y_{\min} \le \|A\|_\infty(x_{\max} - x_{\min})$.

Proposition 2.4. *Let A be the matrix defined in (2.22); that is, $A = (A(\alpha,\beta) := a(h^*_\beta,\alpha))_{\alpha,\beta \in I(n,d)}$. Then, $Ae = \frac{1}{d^d}e$ and $\|A\|_\infty \le \binom{2d-1}{d}$.*

2.2.1. *Proof of Proposition 2.3.* Assume $y = Ax$ where $A = (a_{ik})_{i,k=1}^N$. At row i, $y_i = \sum_{k=1}^N a_{ik}x_k$. Thus,

$$y_i \le \left(\sum_{k|a_{ik}\ge 0} a_{ik}\right) x_{\max} - \left(\sum_{k|a_{ik}\le 0} |a_{ik}|\right) x_{\min} = r_i^+ x_{\max} - r_i^- x_{\min},$$

after setting

$$r_i^+ := \sum_{k|a_{ik}\ge 0} a_{ik},\ r_i^- := \sum_{k|a_{ik}\le 0} |a_{ik}|.$$

Similarly,

$$y_j \ge r_j^+ x_{\min} - r_j^- x_{\max}.$$

Therefore, for any $i,j = 1,\dots,N$,

$$y_i - y_j \le (r_i^+ + r_j^-)x_{\max} - (r_i^- + r_j^+)x_{\min}.$$

Note that $ax_{\max} - bx_{\min} \le \frac{a+b}{2}(x_{\max} - x_{\min})$ if and only if $(b-a)(x_{\max} + x_{\min}) \ge 0$. Here, $a = r_i^+ + r_j^-$, $b = r_i^- + r_j^+$, $b - a = \sum_k a_{jk} - \sum_k a_{ik} = 0$, and $b + a = \sum_k |a_{ik}| + |a_{jk}|$. Therefore,

$$y_i - y_j \le \frac{1}{2}\left(\sum_{k=1}^{N} |a_{ik}| + |a_{jk}|\right)(x_{\max} - x_{\min})$$

and

$$y_j - y_i \le \frac{1}{2}\left(\sum_{k=1}^{N} |a_{ik}| + |a_{jk}|\right)(x_{\max} - x_{\min}).$$

This implies that, for any i, j,

$$|y_i - y_j| \le \frac{1}{2}\left(\sum_{k=1}^{N} |a_{ik}| + |a_{jk}|\right)(x_{\max} - x_{\min}) \le \left(\max_i \sum_k |a_{ik}|\right)(x_{\max} - x_{\min});$$

that is,

$$y_{\max} - y_{\min} \le \|A\|_\infty (x_{\max} - x_{\min}).$$

2.2.2. *Proof of Proposition 2.4.* Let us first prove that $Ae = d^{-d}e$. For this consider the polynomial $p(x) := (\sum_{i=1}^n x_i)^d$. Then, $a(p,\alpha) = 1$ and $p(\alpha) = d^d$ for all $\alpha \in I(n,d)$. Thus, $x := (p(\alpha))_{\alpha \in I(n,d)} = d^d e$ and $y := (a(p,\alpha))_{\alpha \in I(n,d)} = e$. As $y = Ax$ by (2.23), it follows that $Ae = d^{-d}e$.

Recall that $A(\alpha,\beta) = a(h^*_\beta, \alpha) = a(h_\beta,\alpha)\frac{1}{\beta! d^d}$. Thus, $A(\alpha,\beta)$ is equal to the coefficient of x^α in $h_\beta(x)$ scaled by the factor $\frac{\alpha!}{d!}\frac{1}{\beta! d^d}$. We proceed in two steps for proving that

$$\|A\|_\infty = \max_{\alpha \in I(n,d)} \sum_{\beta \in I(n,d)} |A(\alpha,\beta)| \le \binom{2d-1}{d}.$$

(1) First, we show that each entry of A is bounded in absolute value by 1.
(2) Second, we show that there are at most $\binom{2d-1}{d}$ nonzero entries in any row of A.
Those two facts imply obviously the desired result.

Step (1). Bounding the entries of A. By definition, $h_\beta(x)$ is defined as the product of d linear forms: $f_1(x) = \sum_{i=1}^n f_{1i}x_i, \ldots, f_d(x) = \sum_{i=1}^n f_{di}x_i$. Thus,

$$h_\beta(x) = \sum_{i_1=1,\ldots,n} \cdots \sum_{i_d=1,\ldots,n} f_{1i_1}\cdots f_{di_d} x_{i_1}\cdots x_{i_d} =: \sum_{\alpha \in I(n,d)} s_\alpha x^\alpha,$$

where $s_\alpha = \sum f_{1i_1}\cdots f_{di_d}$ and the summation is over all d-tuples $(i_1,\ldots,i_d) \in \{1,\ldots,n\}^d$ containing 1 exactly α_1 times, 2 exactly α_2 times, …, n exactly α_n times.

First of all, each product $f_{1i_1}\cdots f_{di_d}$ is bounded in absolute value by d^d. Indeed, the linear forms $f_j(x)$ are of the form: $(d-\ell)x_1 - \ell x_2 \ldots - \ell x_n$; thus their coefficients belong to $\{-d, -d+1, \ldots, 0, 1, \ldots, d\}$.

Let us now count the number of terms in the summation defining s_α. It is equal to $\binom{d}{\alpha_1}\cdot\binom{d-\alpha_1}{\alpha_2}\cdots\binom{d-\alpha_1-\ldots-\alpha_{n-1}}{\alpha_n}$, which is equal to $\frac{d!}{\alpha!}$.

Summarizing, we find that $|s_\alpha| \le d^d\frac{d!}{\alpha!}$. Hence, $|A(\alpha,\beta)| = |s_\alpha|\frac{\alpha!}{d!}\frac{1}{\beta! d^d} \le \frac{1}{\beta!} \le 1$.

Step (2). Bounding the number of nonzero entries in a row of A. Write $h_\beta(x) = \prod_{j=1}^n P_j(x)$, where

$$P_j(x) = \prod_{\ell_j=0}^{\beta_j-1}\left((d-\ell_j)x_j - \sum_{i=1,\ldots,n,\ i\ne j} \ell_j x_i\right).$$

Fix $\alpha \in I(n,d)$ and consider the α-th row of A. We want to bound the number of β's for which $A(\alpha,\beta) \ne 0$; that is, the number of β's for which x^α occurs with a nonzero coefficient in $h_\beta(x)$.

Consider some β for which $A(\alpha,\beta) \neq 0$. Say, $supp(\beta) = \{1,\ldots,t\}$; that is, $\beta_1 \geq 1,\ldots,\beta_t \geq 1, \beta_{t+1} = \ldots = \beta_n = 0$. Then, $P_j(x) = 1$ for $j = t+1,\ldots,n$ and

$$P_j(x) = dx_j \prod_{\ell_j=1}^{\beta_j-1} \left((d-\ell_j)x_j - \sum_{i=1,\ldots,n,\ i\neq j} \ell_j x_i \right)$$

for $j = 1,\ldots,t$. Therefore,

$$h_\beta(x) = d^t \prod_{j=1}^{t} x_j \prod_{j=1}^{t} \prod_{\ell_j=1}^{\beta_j-1} \left((d-\ell_j)x_j - \sum_{i=1,\ldots,n,\ i\neq j} \ell_j x_i \right).$$

Hence, if x^α has a nonzero coefficient in $h_\beta(x)$, then necessarily $\alpha_1 \geq 1,\ldots,\alpha_t \geq 1$; that is, the support of β is contained in the support of α.

Therefore, the number of β's for which $A(\alpha,\beta) \neq 0$ is bounded by the number of sequences $\beta \in I(n,d)$ with $supp(\beta) \subseteq supp(\alpha)$, which is equal to $\binom{s+d-1}{d}$, setting $s := |supp(\alpha)|$. As $|\alpha| = d$, $s \leq d$ and thus $\binom{s+d-1}{d} \leq \binom{2d-1}{d}$.

3. A PTAS for the minimization of forms on the simplex

Consider the generic optimization problem:

$$\phi_{\max} := \max\left\{ f(x) \ : \ x \in S \right\},$$

for some continuous $f : \mathbb{R}^n \mapsto \mathbb{R}$ and compact convex set S, and let

$$\phi_{\min} := \min\left\{ f(x) \ : \ x \in S \right\}.$$

Definition 3.1 (Nesterov *et al.* [8]). A value ψ_μ approximates $\phi_{\min}$ with relative accuracy $\mu \in (0,1]$ if

$$|\psi_\mu - \phi_{\min}| \leq \mu(\phi_{\max} - \phi_{\min}).$$

Then one also says that ψ_μ is a μ-approximation of $p_{\min}$. The approximation is called implementable if $\psi_\mu = f(x)$ for some $x \in S$.

Definition 3.2 (PTAS). If a problem allows an implementable approximation $\psi_\mu = f(x_\mu)$ for each $\mu \in (0,1]$, such that $x_\mu \in S$ can be computed in time polynomial in n and the bit size required to represent f, we say that the problem allows a polynomial time approximation scheme (PTAS).

For the problem of minimizing a form over the simplex, this definition may be summarized as follows.

Definition 3.3. Consider the problem (1.1) of minimizing a degree d form p on the standard simplex. A PTAS for this problem exists if, for every $\epsilon > 0$, there is an algorithm that returns a solution $x \in \Delta$ satisfying

$$p(x) - p_{\min} \leq (p_{\max} - p_{\min})\epsilon \tag{3.1}$$

in time polynomial in n and the bit size of the coefficients of p.

Note that the approximation results in Theorem 1.2 do not constitute a PTAS, since it is not clear how to bound $p_{\max}^{(0)} - p_{\min}$ in terms of $p_{\max} - p_{\min}$. One reason for adopting Definition 3.3 is that, in general, nothing is known about the signs of $p_{\min}$ and $p_{\max}$. Let us mention a variation of this definition that could be considered as well.

Definition 3.4. A PTAS (of type 2) for the problem (1.1) of minimizing a form $p(x)$ on the standard simplex exists if, for every $\epsilon > 0$, there is an algorithm that returns a solution $x \in \Delta$ satisfying

$$p(x) - p_{\min} \leq \epsilon |p_{\min}|, \quad \text{i.e., } p(x) \leq p_{\min}(1 + \epsilon \cdot sign(p_{\min})). \tag{3.2}$$

in time polynomial in n and the bit size of the coefficients of p.

This new definition is stronger, i.e., relation (3.2) implies relation (3.1), when $p_{\min} \leq 0 \leq p_{\max}$ and when $0 \leq 2p_{\min} \leq p_{\max}$. Note that Definition 3.4 corresponds to the classic definition used, e.g., for combinatorial optimization problems, where the signs of the minimum and maximum are often known in advance. Consider, for instance, the (unweighted) max-cut problem:

$$mc(G) := \max_{x \in \{\pm 1\}^n} p_0(x) := \frac{1}{2} \sum_{ij \in E} (1 - x_i x_j) \tag{3.3}$$

for a graph $G = (V, E)$ ($|V| = n$). Then, a PTAS for the max-cut problem should, for every $\epsilon > 0$, return a point $x \in \{\pm 1\}^n$ for which

$$p(x) \geq (1 - \epsilon) mc(G) \tag{3.4}$$

in time polynomial in n. Similarly, a PTAS for the maximum stable set problem should, for every $\epsilon > 0$, return a stable set S satisfying $|S| \geq (1 - \epsilon)\alpha(G)$.

As an application of Theorem 1.5 (or Theorems 1.3 and 1.4 in the case $d = 2, 3$), we find:

Theorem 3.5. *There is a PTAS (using Definition 3.3) for the problem of minimizing a form of degree $d \geq 2$ over the simplex Δ.* □

On the other hand, if we use Definition 3.4 instead of Definition 3.3, the above result does not remain valid.

Theorem 3.6. *There is no PTAS of type 2 (using Definition 3.4) for the problem of minimizing a form of degree $d \geq 2$ over the simplex Δ.*

Proof. It suffices to show the result for degree $d = 2$. The proof is based on a reduction from the maximum stable set problem. Given a graph $G = (V, E)$ with adjacency matrix A, consider the quadratic polynomial $p(x) := x^T(I + A)x$. By Motzkin-Straus theorem, the minimum of $p(x)$ over Δ is $\frac{1}{\alpha(G)}$, where $\alpha(G)$ is the maximum cardinality of a stable set in G. Thus, $p_{\min} = \frac{1}{\alpha(G)} > 0$ and $p_{\max} \geq 2p_{\min}$ if $\alpha(G) \geq 2$, since $p_{\max} \geq 1 = p(e_i)$.

Lemma 3.7. *Given $x^* \in \Delta$, one can construct a stable set S for which $\frac{1}{|S|} \leq p(x^*)$ in time polynomial in n.*

Proof. The proof is based on the same argument used for proving Motzkin-Straus theorem. Let T denote the support of x^*. First we construct another point $x \in \Delta$ whose support is stable and such that $p(x) \leq p(x^*)$. If T is stable, let $x := x^*$. Suppose that T contains two adjacent nodes, say, nodes 1 and 2. Consider the function $f(x_1, x_2) := p(x_1, x_2, x_3^*, \ldots, x_n^*)$ in two variables x_1, x_2. For any point $(x_1, x_2) \in \Delta_0 := \{(x_1, x_2) \mid x_1, x_2 \geq 0,\ x_1 + x_2 = 1 - \sum_{i \geq 3} x_i^*\}$, $f(x_1, x_2)$ has the form $ax_1 + bx_2 + c$ where a, b, c are constants depending on $x_3^*, \ldots, x_n^*$. As f is linear, it attains its minimum on the segment Δ_0 at one of its extremities, i.e., at (x_1, x_2) with $x_1 = 0$ or $x_2 = 0$. Thus one can construct a new point $x \in \Delta$ such that $p(x) \leq p(x^*)$, whose support is contained in T and does not contain $\{1, 2\}$. Iterating, we find a point $x \in \Delta$ whose support S is stable and such that $p(x) \leq p(x^*)$. Now, $p(x) = \sum_{i \in S} x_i^2$ which, using Cauchy-Schwartz inequality, implies that $p(x) \geq \frac{1}{|S|}$. □

Assume that there is a PTAS of type 2, using Definition 3.4, for the problem of minimizing a quadratic form on the simplex. Let ϵ be given such that $0 < \epsilon < 1$ and set $\epsilon' := \frac{\epsilon}{1-\epsilon}$. Then one can construct in polynomial time a point $x \in \Delta$ satisfying (3.2), i.e., $p(x) \leq \frac{1}{\alpha(G)}(1+\epsilon')$. By Lemma 3.7, one can construct in polynomial time a stable set S such that $\frac{1}{|S|} \leq p(x) \leq \frac{1}{\alpha(G)}(1+\epsilon')$, which implies $|S| \geq \alpha(G)(1-\epsilon)$. This shows therefore the existence of a PTAS for the maximum stable set problem, contradicting the inapproximability result of Arora et al. [1]. □

4. Concluding Remarks

4.1. **A probabilistic approach for estimating the grid bounds $p_{\Delta(k)}$.** Nesterov [9] proposes an alternative probabilistic argument for estimating the quality of the bounds $p_{\Delta(k)}$. He introduces a random walk on the simplex Δ, which generates a sequence of random points x_k ($k \geq 0$) in the simplex with the property that $x_k \in \Delta(k)$. Thus the expected value $E(p(x_k))$ of the evaluation of the polynomial $p(x)$ at x_k satisfies: $E(p(x_k)) \geq p_{\Delta(k)}$. Hence upper bounds for $p_{\Delta(k)}$ can be obtained by bounding $E(p(x_k))$.

Fix a point $q \in \Delta$ (to be chosen later as a global minimizer of the polynomial $p(x)$ over Δ). Let ζ be a discrete random variable with values in $\{1, \ldots, n\}$ distributed as follows:

$$\text{Prob}(\zeta = i) = q_i \ (i = 1, \ldots, n). \tag{4.1}$$

Consider the random process:

$$y_0 = 0 \in \mathbb{R}^n, \ y_{k+1} = y_k + e_{\zeta_k} \ (k \geq 0)$$

where ζ_k are random independent variables distributed according to (4.1). In other words, y_{k+1} is $y_k + e_i$ with probability q_i. Finally, define

$$x_k = \frac{1}{k} y_k \ (k \geq 1).$$

Thus all x_k lie in the set $\Delta(k)$.

In order to evaluate $E(p(x_k))$, one needs to compute $E(x_k^\beta)$ for any monomial with $|\beta| = d$. One can verify that $E(x_k^\beta) = \frac{1}{k^d} E(y_k^\beta)$, with

$$E(y_k^\beta) = \sum_{\alpha \in \mathbb{N}^n, \ |\alpha| = k} \text{Prob}(y_k(1) = \alpha_1, \ldots, y_k(n) = \alpha_n) \alpha^\beta = \sum_{|\alpha| = k} \frac{k!}{\alpha!} q^\alpha \alpha^\beta.$$

(Here, $y_k(i)$ is the i-th coordinate of y_k and recall that $q^\alpha = q_1^{\alpha_1} \cdots q_n^{\alpha_n}$ and $\alpha^\beta = \alpha_1^{\beta_1} \cdots \alpha_n^{\beta_n}$.) The following computations are given in [9]:

$$E(x_k(i)) = q_i, \ E(x_k(i)^2) = \frac{1}{k} q_i + \left(1 - \frac{1}{k}\right) q_i^2,$$

$$E(x_k(i) x_k(j)) = \left(1 - \frac{1}{k}\right) q_i q_j \ (i \neq j),$$

$$E(x_k^\beta) = \frac{k!}{(k-d)!k^d} q^\beta \quad \text{if } \beta \in \{0,1\}^n \text{ with } |\beta| = d.$$

Assume now that q is a global minimizer of $p(x)$ over Δ. If $p(x)$ is a sum of square-free monomials of degree d, then

$$E(p(x_k)) = \sum_{|\beta|=d} p_\beta E(x_k^\beta) = p(q) \frac{k!}{(k-d)!k^d} = p_{\min} w_r(d),$$

with $k = r + d$. This gives the estimate (1.7) ([9], Lemma 3). If $p(x)$ is a form of degree 2, then

$$E(p(x_k)) = \frac{1}{k} \sum_i p_{2e_i} q_i + \left(1 - \frac{1}{k}\right) p(q) = w_r(2) p_{\min} + \frac{1}{r+2} \sum_{i=1}^n p_{2e_i} q_i,$$

which gives the estimate (1.6) ([9], Theorem 2).

Remark 4.1. In these two cases (sum of square-free monomials, or degree 2), it turns out that $E(p(x_k)) = w_r(d)p_{\min} + \phi(q)$. Hence, the upper bound $w_r(d)p_{\min} + \phi_{\max}$ for $p_{\Delta(r+d)}$ (recall (2.8)) remains an upper bound for $E(p(x_k))$. The identity $E(p(x_k)) = w_r(d)p_{\min} + \phi(q)$ is not true when $d = 3$. When $d = 3$, one can verify that $E(p(x_k)) = w_r(3)p_{\min} + \phi(q) + \phi'(q)$, where

$$\phi'(q) := \left(\frac{1}{r+3}\right)^2 \left(3 \sum_i p_{3e_i} q_i (1 - q_i) - \sum_{i<j} (p_{2e_i+e_j} + p_{e_i+2e_j}) q_i q_j\right).$$

One can verify that $\phi'(q) \leq \frac{4}{(r+3)^2}(p_{\max} - p_{\min})$. Combined with (2.16), this implies that

$$E(p(x_k)) \leq p_{\min} + \left(\frac{4}{r+3} + \frac{4}{(r+3)^2}\right)(p_{\max} - p_{\min}).$$

4.2. **Approximating polynomials over polytopes.** As observed by Nesterov [9], some results for the simplex can be extended to the problem of minimizing a degree d form $p(x)$ over a polytope

$$P := \operatorname{conv}(u_1, \ldots, u_N)$$

where $u_1, \ldots, u_N \in \mathbb{R}^n$. Indeed, if U denote the $n \times N$ matrix with columns $u_1, \ldots, u_N$, then minimizing the polynomial (in n variables) $p(x)$ over P is equivalent to minimizing the polynomial (in N variables) $\tilde{p}(x) := p(Ux)$ over the standard simplex Δ in $\mathbb{R}^N$. Thus,

$$p_{\min,P} := \min_{x \in P} p(x) = \min_{x \in \Delta} \tilde{p}(x)$$

and, for an integer $k \geq 1$, one can define the grid approximation:

$$p_{P(k)} := \tilde{p}_{\Delta(k)} = \min_{x \in \Delta(k)} p\left(\sum_{i=1}^N x_i u_i\right).$$

The bounds obtained earlier for $\tilde{p}_{\Delta(k)}$ translate into bounds for $p_{P(k)}$. For instance, when $p(x)$ has degree 2,

$$p_{P(k)} - p_{\min,P} \leq \frac{1}{k}\left(\max_{i=1,\ldots,N} p(u_i) - p_{\min,P}\right).$$

When $p(x)$ is a sum of square-free monomials,

$$p_{P(k)} - p_{\min,P} \le \frac{d(d-1)}{2k}(-p_{\min,P}).$$

Let us point out, however, that the complexity of computing the parameter $p_{P(k)}$ depends on the number N of vertices, which can be exponentially large in terms of the number n of variables.

Observe that the problem of maximizing a quadratic form over the cube $[-1,1]^n$ is NP-hard and no PTAS can exist, since it contains the max-cut problem. Indeed, given a graph $G=(V,E)$, define its Laplacian matrix L as the $V\times V$ matrix with entries $L_{ii} := -\deg(i)$ $(i \in V)$ and $L_{ij} := 1$ if $i \neq j$ are adjacent, $L_{ij} := 0$ otherwise. Then,

$$mc(G) = \max_{x\in\{\pm 1\}^n} \frac{1}{4} x^T L x = \max_{x\in[-1,1]^n} \frac{1}{4} x^T L x,$$

where the last equality follows from the fact that $L \succeq 0$.

4.3. **Semidefinite approximations.** Stronger semidefinite bounds can be defined for the minimum $p_{\min}$ of a degree d form $p(x)$ over the standard simplex Δ. For this, if $p(x) = \sum_\alpha p_\alpha x^\alpha$, consider the (even) polynomial

$$\tilde{p}(x) := \sum_\alpha p_\alpha x^{2\alpha}.$$

The problem of minimizing $p(x)$ over the simplex Δ is equivalent to the problem of minimizing $\tilde{p}(x)$ over the unit sphere $S := \{x \in \mathbb{R}^n \mid \sum_{i=1}^n x_i^2 = 1\}$; that is,

$$p_{\min} = \min_{x\in S} \tilde{p}(x).$$

Let Σ^2 denote the set of polynomials in $\mathbb{R}[x_1,\ldots,x_n]$ that can written as a sum of squares of polynomials. Given an integer $r \ge 0$, define the parameter:

$$p^{(r)}_{\min,sos} := \max\ \lambda \text{ for which } \left(\sum_{i=1}^n x_i^2\right)^r \left(\tilde{p}(x) - \lambda\left(\sum_{i=1}^n x_i^2\right)^d\right) \in \Sigma^2. \tag{4.2}$$

If the polynomial $(\sum_i x_i)^r(p(x) - \lambda\left(\sum_{i=1}^n x_i\right)^d)$ has nonnegative coefficients, then the polynomial $(\sum_i x_i^2)^r(\tilde{p}(x) - \lambda\left(\sum_{i=1}^n x_i^2\right)^d)$ is obviously a sum of squares. Therefore,

$$p^{(r)}_{\min} \le p^{(r)}_{\min,sos} \le p_{\min} \quad \text{for all } r \ge 0.$$

The bound $p^{(r)}_{\min,sos}$ can be computed in polynomial time with an arbitrary precision for any *fixed* r. This follows from the well known fact (see, e.g., [14]) that testing whether a polynomial can be written as a sum of squares of polynomials can be formulated as a semidefinite program. As a consequence of Pólya's theorem, the semidefinite bounds $p^{(r)}_{\min,sos}$ converge to $p_{\min}$ as $r \to \infty$.

Schmüdgen [17] proved (in a more general context) that every polynomial which is positive on the unit sphere S has a representation of the form $s_0(x) + (1 - \sum_i x_i^2)s_1(x)$, where $s_0(x) \in \Sigma^2$ and $s_1(x) \in \mathbb{R}[x_1,\ldots,x_n]$. This fact motivates the definition of the following alternative semidefinite lower bound for $p_{\min}$, for any integer $r \ge 0$:

$$\begin{array}{ll} \max\ \lambda & \text{such that } \tilde{p}(x) - \lambda = s_0(x) + (1 - \sum_{i=1}^n x_i^2)s_1(x) \\ & \text{where } s_0 \in \Sigma^2,\ s_1 \in \mathbb{R}[x_1,\ldots,x_n],\ \deg(s_0) \le 2(r+d) \end{array} \tag{4.3}$$

It follows from Schmüdgen's theorem that these bounds also converge to $p_{\min}$ as $r \to \infty$. In fact, De Klerk et al. [4] show that the bounds (4.3) coincide with the semidefinite bounds $p^{(r)}_{\min,sos}$. That is, both approaches based on Pólya's result and on Schmüdgen's result yield the same hierarchies of semidefinite bounds for the problem of minimizing a form on the simplex.

4.4. **Optimizing polynomials over the unit sphere.** We group here a few observations about the complexity of optimizing a form over the sphere.

As is well-known, minimizing a quadratic form over the unit sphere is an easy problem, as it amounts to computing the minimum eigenvalue of a matrix, a problem for which efficient algorithms exist.

As we saw in the previous subsection, the problem of minimizing an even form on the unit sphere can be reformulated as the problem of minimizing an associated form on the simplex. Hence upper and lower bounds are available as well as good estimates on their quality.

On the other hand, Nesterov [9] shows that maximizing a cubic form on the unit sphere is a NP-hard problem, using a reduction from the maximum stable set problem.

Let us finally mention a result of Faybusovich [5] about the quality of the semidefinite bounds for the optimization of forms on the unit sphere. Let $p(x)$ be a form of even degree $2d$, let S denote the unit sphere, and set $p_{\min,S} := \min_{x\in S} p(x)$, $p_{\max,S} := \max_{x\in S} p(x)$. For an integer $r \geq 0$, define the parameter

$$p^{(r)}_S := \max\ \lambda \text{ s.t. } \left(\sum_{i=1}^n x_i^2\right)^r \left(p(x) - \lambda\left(\sum_{i=1}^n x_i^2\right)^d\right) \in \Sigma^2.$$

Thus, $p^{(r)}_S \leq p_{\min,S}$ for all $r \geq 0$. Using a result of Reznick [16], Faybusovich [5] shows that, for $r \geq \frac{2nd(2d-1)}{4\ln 2} - \frac{n}{2} - d$,

$$p_{\min,S} - p^{(r)}_S \leq \frac{2nd(2d-1)}{2\ln 2(2r+n+2d) - 2nd(2d-1)}(p_{\max,S} - p_{\min,S}).$$

This does not yield a PTAS, since this estimate holds only for $r = O(n)$. It remains an open problem whether optimization of a fixed degree form over the unit sphere allows a PTAS.

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedi. Proof verification and intractability of approximations problems. In *Proceedings of the 33rd IEEE Symposium on Foundations of Computer Science*, IEEE Computer Science Press, Los Alamitos, CA, pp 14–23, 1992.

[2] P. Berman and M. Karpinski. On some tighter inapproximability results, further improvements. *Electronic Colloquium on Computational Complexity,* Report TR98-065, 1998. `http://www.eccc.uni-trier.de/eccc/`

[3] I. Bomze and E. De Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Global Optimization*, 24(2):163–185, 2002.

[4] E. De Klerk, M. Laurent, and P. Parrilo. On the equivalence of algebraic approaches to the minimization of forms on the simplex. In *Positive Polynomials in Control*, D. Henrion and A. Garulli, eds., LNCIS, Springer, forthcoming.

[5] L. Faybusovich. Global optimization of homogeneous polynomials on the simplex and on the sphere. In C. Floudas and P. Pardalos, editors, *Frontiers in Global Optimization*. Kluwer Academic Publishers, 2003.

[6] J. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J.Optim.*, 11(3):796–817, 2001.

[7] T.S. Motzkin and E.G. Straus. Maxima for graphs and a new proof of a theorem of Túran. *Canadian J. Math.*, 17:533-540, 1965.

[8] Yu. Nesterov, H. Wolkowicz, and Y. Ye. Semidefinite programming relaxations of nonconvex quadratic optimization. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *Handbook of semidefinite programming*, pages 361–419. Kluwer Academic Publishers, Norwell, MA, 2000.

[9] Yu. Nesterov. Random walk in a simplex and quadratic optimization over convex polytopes. CORE Discussion Paper 2003/71, CORE-UCL, Louvain-La-Neuve, September 2003.

[10] P.A. Parrilo. *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization.* PhD thesis, California Institute of Technology, May 2000.

[11] P.A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming, Ser. B*, 96:293–320, 2003.

[12] G. Pólya. *Collected Papers*, volume 2, pages 309–313. MIT Press, Cambridge, Mass., London, 1974.

[13] V. Powers and B. Reznick. A new bound for Pólya's theorem with applications to polynomials positive on polyhedra. *Journal of Pure and Applied Algebra*, 164:221–229, 2001.

[14] V. Powers and T. Wörmann. An algorithm for sums of squares of real polynomials. *Journal of Pure and Applied Algebra*, 127:99–104, 1998.

[15] B. Reznick. Sums of even powers of real linear forms. *Memoirs of the American Mathematical Society*, Number 463, 1992.

[16] B. Reznick. Uniform denominators in Hilbert's seventeenth problem. *Mathematische Zeitschrift*, 220:75–97, 1995.

[17] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289:203–206, 1991.

[18] L.F. Zuluaga, J.C. Vera, and J.F. Peña. LMI approximations for cones of positive semidefinite forms. Preprint, 2003.

E. De Klerk, University of Waterloo
E-mail address: `edeklerk@math.uwaterloo.ca`

M. Laurent, CWI, Amsterdam
E-mail address: `M.Laurent@cwi.nl`

P. Parrilo, Swiss Federal Institute of Technology, Zürich
E-mail address: `parrilo@control.ee.ethz.ch`