Available online at www.sciencedirect.com
ScienceDirect
Theoretical Computer Science 361 (2006) 210–225
Theoretical Computer Science
www.elsevier.com/locate/tcs

A PTAS for the minimization of polynomials of fixed degree over the simplex

Etienne de Klerk[a,b,1], Monique Laurent[c,*,2], Pablo A. Parrilo[d]

[a] *Tilburg University, The Netherlands*
[b] *University of Waterloo, Canada*
[c] *CWI, Amsterdam, The Netherlands*
[d] *Massachusetts Institute of Technology, Cambridge, USA*

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 $k \geqslant 1$, 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 $d=2$.

MSC: 90C22; 90C26; 49M20

Keywords: Semidefinite programming; Global optimization; Positive polynomial; Sum of squares of polynomials; Approximation algorithm

1. Introduction

1.1. Problem definition and complexity

We consider the problem of minimizing a polynomial $p(x)$ of degree d on the standard simplex

$$\Delta := \left\{ x \in \mathbb{R}^n_+ \,\middle|\, \sum_{i=1}^{n} x_i = 1 \right\};$$

* Corresponding author.
E-mail addresses: E.deKlerk@UvT.nl (E. de Klerk), M.Laurent@cwi.nl (M. Laurent), parrilo@mit.edu (P.A. Parrilo).
[1] Supported by The Netherlands Organization for Scientific Research Grants NWO 613.000.214 and NWO 016.025.026, and the NSERC Grant 283331-04.
[2] Supported by The Netherlands Organization for Scientific Research Grant NWO 639.032.203.

0304-3975/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.tcs.2006.05.011

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 [7], 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 $p'(x) := \sum_{\ell=0}^{d} p_\ell(x)(\sum_{i=1}^{n} x_i)^{d-\ell}$. We will use the standard compact notation:

$$p(x) = \sum_{\alpha} p_\alpha x^\alpha,$$

where the summation is over $\alpha \in \mathbb{Z}_+^n$ with finitely many nonzero terms, and $x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$. Setting $|\alpha| := \sum_{i=1}^{n} \alpha_i$, then $|\alpha| = d$ for all nonzero terms when p is a degree d form.

Problem (1.1) is an NP-hard problem, already for forms of degree $d = 2$, as it contains the maximum stable set problem. Indeed, let G be a graph with adjacency matrix A and let I denote the identity matrix; then the maximum size $\alpha(G)$ of a stable set in G can be expressed as

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

by the theorem of Motzkin and Straus [9].

1.2. Upper bounds using a rational grid

Given an integer $k \geqslant 1$, 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

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

Thus, $p_{\min} \leqslant p_{\Delta(k)}$ for any $k \geqslant 1$. 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 [5] show that the following inequality holds:

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

for any $k \geqslant 1$.

Using a probabilistic approach, Nesterov [11] gave a different proof of (1.5), and proved the following result for the case when $p(x)$ 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 $\alpha_i \leqslant 1$ for all $i = 1, \dots, n$. Then, for $k \geqslant d$,

$$p_{\Delta(k)} - p_{\min} \leqslant \left(1 - \frac{k!}{(k-d)!k^d}\right)(-p_{\min}) \leqslant \frac{d(d-1)}{2k}(-p_{\min}). \tag{1.6}$$

1.3. Lower bounds using Pólya's representation theorem

A second—and closely related—way of obtaining approximations to $p_{\min}$, is via Pólya's representation theorem for positive forms on the simplex.

We need to introduce the following parameters for a polynomial $p(x) = \sum_\alpha p_\alpha x^\alpha$:

$$L_p := \max_\alpha |p_\alpha| \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!}, \tag{1.7}$$

$$p_{\max}^{(0)} := \max_\alpha p_\alpha \frac{\alpha_1! \cdots \alpha_n!}{|\alpha|!} \tag{1.8}$$

which obviously satisfy: $p_{\max}^{(0)} \leqslant L_p$.

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 $(\sum_{i=1}^n x_i)^r p(x)$ has nonnegative coefficients for all r satisfying*

$$r \geqslant \binom{d}{2} \frac{p_{\max}^{(0)}}{p_{\min}} - d. \tag{1.9}$$

Pólya [17] proved that $(\sum_{i=1}^n x_i)^r p(x)$ has nonnegative coefficients for r large enough. Powers and Reznick [18] proved that this holds for any $r \geqslant \binom{d}{2} L_p / p_{\min} - d$. We observe here that this holds for any r satisfying the weaker condition (1.9) (with L_p replaced by $p_{\max}^{(0)}$); see Section 2.1 for a proof.

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 $p(x) - \lambda \geqslant 0$ for all $x \in \Delta$. Equivalently,

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

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

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

For the problem of maximizing $p(x)$ over Δ, one can analogously define the parameter $p_{\max}^{(r)}$ as the minimum scalar λ for which the polynomial $-p^{(r)}(x)$ has nonnegative coefficients. One can verify (see Section 2.1) that

$$p_{\min}^{(0)} = \min_\alpha p_\alpha \frac{\alpha_1! \cdots \alpha_n!}{d!}, \quad p_{\max}^{(0)} = \max_\alpha p_\alpha \frac{\alpha_1! \cdots \alpha_n!}{d!}; \tag{1.10}$$

that is, we find again the value from (1.8) for $p_{\max}^{(0)}$. Obviously,

$$p_{\min}^{(0)} \leqslant p_{\min}^{(r)} \leqslant p_{\min}^{(r+1)} \leqslant p_{\min}.$$

Moreover, $p_{\min}^{(r)} \geqslant 0$ if and only if the polynomial $(\sum_{i=1}^n x_i)^r p(x)$ has nonnegative coefficients. Hence, Theorem 1.1 asserts that $p_{\min}^{(r)} \geqslant 0$ for any r satisfying (1.9). 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 relation (2.3). As a consequence of Pólya's theorem, the bounds $p_{\min}^{(r)}$ converge asymptotically to $p_{\min}$ as $r \to \infty$.

This idea of using Pólya's result for constructing converging approximations goes back to the work of Parrilo [15,16], 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 [7], and Zuluaga et al. [24]. Faybusovich [7] proved:

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

$$p_{\min} - p_{\min}^{(r)} \leqslant (L_p - p_{\min}) \left(\frac{1}{w_r(d)} - 1 \right), \tag{1.11}$$

where

$$w_r(d) := \frac{(r+d)!}{r!(r+d)^d} = \prod_{i=1}^{d-1}\left(1 - \frac{i}{r+d}\right). \tag{1.12}$$

One can verify that

$$1 - \binom{d}{2}\frac{1}{r+d} \leqslant w_r(d) \leqslant 1, \tag{1.13}$$

which implies that $\lim_{r\to\infty} w_r(d) = 1$. Thus one finds again that the sequence $p_{\min}^{(r)}$ converges to $p_{\min}$ as $r \to \infty$.

1.4. Main results of this paper

In this paper we prove new bounds on the quality of the approximations $p_{\Delta(k)}$ and $p_{\min}^{(r)}$. In particular, we show the following.

Theorem 1.3. *Let $p(x)$ be a form of degree d and $r \geqslant 0$ an integer. Then*,

$$p_{\min} - p_{\min}^{(r)} \leqslant \left(\frac{1}{w_r(d)} - 1\right)\binom{2d-1}{d} d^d (p_{\max} - p_{\min}), \tag{1.14}$$

$$p_{\Delta(r+d)} - p_{\min} \leqslant (1 - w_r(d))\binom{2d-1}{d} d^d (p_{\max} - p_{\min}). \tag{1.15}$$

The motivation for proving bounds of this type comes from approximation theory. In order to explain this we recall the definition of an ε-approximation in nonlinear programming. The next definition has been used by several authors, in particular, by Ausiello, d'Atri and Protasi [2], Bellare and Rogaway [3], Nemhauser et al. [10], Nesterov et al. [12], Vavasis [23].

Definition 1.4. Consider the optimization problems:

$$\phi_{\max} := \max\{\phi(x) : x \in S\}, \quad \phi_{\min} := \min\{\phi(x) : x \in S\},$$

where $\phi : \mathbb{R}^n \to \mathbb{R}$ is continuous and S is a compact convex set. Given $\varepsilon > 0$, a value ψ_ε is said to *approximate* $\phi_{\min}$ *with relative accuracy* ε if

$$|\psi_\varepsilon - \phi_{\min}| \leqslant \varepsilon(\phi_{\max} - \phi_{\min}).$$

Then one also says that ψ_ε is a *ε-approximation* of $\phi_{\min}$. The approximation is called *implementable* if $\psi_\varepsilon = \phi(x)$ for some $x \in S$.

As is customary, we speak of a polynomial time approximation scheme (PTAS) if an implementable ε-approximation can be computed in polynomial time for every fixed $\varepsilon > 0$. Formally, we have the following definition.

Definition 1.5 (*PTAS*). If a problem allows an implementable approximation $\psi_\varepsilon = \phi(x_\varepsilon)$ for each $\varepsilon > 0$, such that $x_\varepsilon \in S$ can be computed in time polynomial in n and the bit size required to represent ϕ, 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 1.6. 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 $\varepsilon > 0$, there is an algorithm that returns a solution $x \in \Delta$ satisfying

$$p(x) - p_{\min} \leqslant (p_{\max} - p_{\min})\varepsilon \tag{1.16}$$

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

In view of these definitions, our results in Theorem 1.3 imply the following complexity result.

Theorem 1.7. *There exists a PTAS for the problem class of minimizing a form of fixed degree $d \geqslant 2$ over the simplex Δ.*

In contrast, Bellare and Rogaway [3] proved that if $P \neq NP$ and $\varepsilon \in (0, \frac{1}{3})$, there is no polynomial time ε-approximation algorithm for the problem of minimizing a polynomial of total degree $d \geqslant 2$ over a feasible region $S = \{x \in [0,1]^n \mid Ax \leqslant b\}$. What Theorem 1.7 shows is that there is a PTAS in the special case when S is the standard simplex.

Note that the approximation result from Theorem 1.2 does not constitute a PTAS since it is not clear how to bound $L_p - p_{\min}$ in terms of $p_{\max} - p_{\min}$. The reason is that the quantity L_p is independent of $p_{\max}$ in general as Example 1.8 below illustrates. As argued by Vavasis [23] (see also [2]), the definition of an ε-approximation adopted in Definition 1.4 has some useful invariance properties. For instance, it is invariant under dilation of the objective function, as well as under the addition of a function $g(x)$ constant on the feasible region (e.g., $g(x) = t(\sum_{i=1}^n x_i)^r$ in our case of optimization over the simplex Δ); that is, if the objective function ϕ is replaced by $a\phi + g$ ($a > 0$) then an ε-approximation for ϕ remains an ε-approximation for the new objective function. This invariance property would be lost if one would replace in (1.16) the parameter $p_{\max}$ by the parameter L_p (as in (1.11)). The following example illustrates this too.

Example 1.8. Consider the problem of minimizing a quadratic form $p(x) := x^{\mathrm{T}} Q x$ on the standard simplex Δ. Thus, $L_p = \max_{1 \leqslant i \leqslant j \leqslant n} |Q_{ij}|$. Now let us transform the problem data by replacing Q by $Q - tJ$ for some $t > 0$, where J denotes the all-ones matrix. That is, define

$$p_t(x) := x^{\mathrm{T}}(Q - tJ)x = p(x) - t\left(\sum_{i=1}^n x_i\right)^2.$$

Thus, $p_t(x) = p(x) - t$ if $x \in \Delta$. This transformation does not change the global minimizer, nor does it change the range $(p_t)_{\max} - (p_t)_{\min}$ of the modified objective function. In other words, if $x \in \Delta$ defines an ε-approximation for the original problem, it also defines an ε-approximation for the modified problem.

However, $L_{p_t} - (p_t)_{\min}$ is clearly an increasing function of t for sufficiently large t. Therefore, one cannot bound $L_{p_t} - (p_t)_{\min}$ in terms of $(p_t)_{\max} - (p_t)_{\min} = p_{\max} - p_{\min}$.

Let us now argue that, if we replace in (1.16) the parameter $p_{\max}$ by L_p (as in (1.11)), then we obtain an alternative notion of ε-approximation that is not invariant under addition of a function constant on Δ. Indeed, given $x \in \Delta$ and $t \geqslant 0$, define ε_t via

$$p_t(x) - (p_t)_{\min} = \varepsilon_t(L_{p_t} - (p_t)_{\min}).$$

That is, x is an ε_t-approximation for $(p_t)_{\min}$ using the "new definition". If the invariance property would hold for the new definition, then ε_t would be independent of t. This is not the case, since ε_t goes to 0 as t goes to ∞.

2. Approximating forms on the simplex: Proofs

2.1. Estimating the upper and lower approximations $p_{\Delta(r+d)}$ and $p_{\min}^{(r)}$ for $p_{\min}$ via Pólya's theorem

We use the following notation. Given $\alpha \in \mathbb{N}^n$, set

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

and, following [18], 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, $(1)^d_{1/(r+d)} = w_r(d)$, the parameter defined in (1.12). 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 [18] for a proof. Alternatively, use induction on $m \geqslant 1$.)

In what follows, $p(x)$ is a form of degree d and $r \geqslant 0$ is an integer. By definition, $p^{(r)}_{\min}$ is the maximum scalar λ for which the polynomial $(\sum_i x_i)^r p(x) - \lambda(\sum_i x_i)^{r+d}$ 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

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

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

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

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

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

As in [18], define the polynomial

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

and set

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

Then,

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

This implies

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

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

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

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

Therefore,

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

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

We are now in a position to prove the following result which implies the bound (1.11) by Faybusovich [7], since $p_{\max}^{(0)} \leqslant L_p$.

Theorem 2.1. *Let p be a form of degree d and $r \geqslant 0$ an integer. Then,*

$$p_{\min} - p_{\min}^{(r)} \leqslant (p_{\max}^{(0)} - p_{\min}) \left(\frac{1}{w_r(d)} - 1\right), \tag{2.11}$$

$$p_{\Delta(r+d)} - p_{\min} \leqslant (p_{\max}^{(0)} - p_{\min})(1 - w_r(d)). \tag{2.12}$$

Proof. As $x^\alpha \geqslant \prod_{i=1}^n (x_i)_{1/(r+d)}^{\alpha_i}$ and $p_\alpha \leqslant p_{\max}^{(0)} d!/\alpha!$ (by the definition of $p_{\max}^{(0)}$), we find that

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

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

$$p_{\max}^{(0)} \left(\left(\sum_{i=1}^n x_i\right)^d - \left(\sum_{i=1}^n x_i\right)_{1/(r+d)}^d\right) = p_{\max}^{(0)}(1 - w_r(d)).$$

Therefore,

$$\phi_{\max} \leqslant p_{\max}^{(0)}(1 - w_r(d)). \tag{2.13}$$

This inequality, combined with (2.9) and (2.10), gives the inequalities (2.11) and (2.12) from Theorem 2.1. □

Proof of Theorem 1.1. Assume that $p_{\min} > 0$ and $r \geqslant \binom{d}{2} p_{\max}^{(0)}/p_{\min} - d$. Then, (2.13) combined with the bound on $w_r(d)$ from (1.13) implies that $\phi_{\max} \leqslant p_{\max}^{(0)} \binom{d}{2} 1/(r+d)$. Now, (2.7) implies that $p_{\min}^{(r)} \geqslant (1/w_r(d))(p_{\min} - p_{\max}^{(0)} \binom{d}{2} 1/(r+d))$, which is nonnegative by our assumption on r. This shows that $p_{\min}^{(r)} \geqslant 0$ for such r; that is, Theorem 1.1 holds. □

Proving the inequality (1.6) for polynomials that are sums of square-free monomials: Assume that $p_\alpha = 0$ whenever $\alpha_i \geqslant 2$ for some $i = 1, \ldots, n$. Then the polynomial $\phi(x)$ from (2.5) is identically zero. Thus $\phi_{\max} = 0$ and the estimate (1.6) follows directly from (2.10) (with $k = r + d$ and using (1.13)).

2.2. Estimating the maximum coefficient range of a polynomial

In this section we prove our main result, Theorem 1.3. As $p^{(0)}_{\max} - p_{\min} \leqslant p^{(0)}_{\max} - p^{(0)}_{\min}$, Theorem 1.3 will follow directly from the next result, which estimates $p^{(0)}_{\max} - p^{(0)}_{\min}$ in terms of $p_{\max} - p_{\min}$, combined with Theorem 2.1.

Theorem 2.2. *The following holds for a form $p(x)$ of degree d*:

$$p^{(0)}_{\max} - p^{(0)}_{\min} \leqslant \binom{2d-1}{d} d^d (p_{\max} - p_{\min}). \tag{2.14}$$

We now prove Theorem 2.2. Following Reznick [20], let us introduce the following definitions.

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

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

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

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

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

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

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

$$\langle p, P_\alpha \rangle = p(\alpha) \quad \text{for } \alpha \in \mathbb{R}^n. \tag{2.15}$$

Moreover,

$$\langle p, x^\alpha \rangle = a(p, \alpha) \quad \text{for } \alpha \in I(n, d). \tag{2.16}$$

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

$$h_\alpha(x) := \prod_{j=1}^{n} \prod_{\ell_j=0}^{\alpha_j - 1} (dx_j - \ell_j(x_1 + \cdots + x_n)), \quad h^*_\alpha(x) := \frac{1}{\alpha! d^d} h_\alpha(x). \tag{2.17}$$

Lemma 2.3 (*Reznick [20]*). *For $\alpha, \alpha' \in I(n, d)$, $\langle h^*_\alpha, P_{\alpha'} \rangle = 1$ if $\alpha = \alpha'$ and 0 otherwise.*

Proof. Direct verification. □

Corollary 2.4 (*Reznick [20]*). *The set $\{P_\alpha \mid \alpha \in I(n, d)\}$ is a basis of the vector space $F_{n,d}$.*

The above results can be found in this form in Reznick [20, Section 2], who extended an old result of Biermann in 1903 for the ternary case. They claim in fact the existence and uniqueness of an interpolation homogeneous polynomial of degree d taking prescribed values at the points of $I(n, d)$ (equivalently, at the points of the rational grid $\Delta(d)$) and as such can also be found in Nicolaides [13].

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,

$$x^\alpha = \sum_{\beta \in I(n,d)} A(\alpha,\beta) P_\beta(x). \tag{2.18}$$

For $p \in F_{n,d}$, by taking the inner product in (2.18) with p and using (2.15) and (2.16), we find

$$a(p,\alpha) = \sum_{\beta \in I(n,d)} A(\alpha,\beta) p(\beta) \quad \text{for } \alpha \in I(n,d). \tag{2.19}$$

Taking the inner product in (2.18) with h^*_β, we find

$$a(h^*_\beta, \alpha) = \langle h^*_\beta, x^\alpha \rangle = A(\alpha,\beta) \quad \text{for } \alpha, \beta \in I(n,d). \tag{2.20}$$

In view of (1.10), our parameters $p^{(0)}_{\max}$ and $p^{(0)}_{\min}$ can be expressed as

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

Define the vectors $x := (p(\alpha))_{\alpha \in I(n,d)}$ and $y := (a(p,\alpha))_{\alpha \in I(n,d)}$. In view of (2.19), they are related by the relation

$$y = Ax. \tag{2.21}$$

Denote by $x_{\max}$ (resp., $x_{\min}$) the largest entry (resp., smallest 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,

$$x_{\max} - x_{\min} \leqslant d^d (p_{\max} - p_{\min}). \tag{2.22}$$

Our strategy for proving Theorem 2.2 consists of showing the following two results.

Proposition 2.5. *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 $y_{\max} - y_{\min} \leqslant \|A\|_\infty (x_{\max} - x_{\min})$.

Proposition 2.6. *Let A be the matrix defined in* (2.20); *that is, $A = (A(\alpha,\beta) := a(h^*_\beta, \alpha))_{\alpha,\beta \in I(n,d)}$. Then, $Ae = (1/d^d)e$ and $\|A\|_\infty \leqslant \binom{2d-1}{d}$.*

2.2.1. Proof of Proposition 2.5

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,

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

after setting

$$r_i^+ := \sum_{k \mid a_{ik} \geqslant 0} a_{ik}, \quad r_i^- := \sum_{k \mid a_{ik} \leqslant 0} |a_{ik}|.$$

Similarly,

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

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

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

Note that $ax_{\max} - bx_{\min} \leqslant (a+b)/2(x_{\max} - x_{\min})$ if and only if $(b-a)(x_{\max} + x_{\min}) \geqslant 0$. 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,

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

and

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

This implies that, for any i, j,

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

that is,

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

2.2.2. Proof of Proposition 2.6

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 (2.21), it follows that $Ae = d^{-d}e$.

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

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

(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: $(d - \ell)x_1 - \ell x_2 \cdots - \ell x_n$; 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 $\binom{d}{\alpha_1} \cdot \binom{d-\alpha_1}{\alpha_2} \cdots \binom{d-\alpha_1-\cdots-\alpha_{n-1}}{\alpha_n}$, which is equal to $d!/\alpha!$.

Summarizing, we find that $|s_\alpha| \leqslant d^d(d!/\alpha!)$. Hence, $|A(\alpha, \beta)| = |s_\alpha|(\alpha!/d!)(1/\beta! d^d) \leqslant 1/\beta! \leqslant 1$.

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

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

Fix $\alpha \in I(n, d)$ and consider the αth row of A. We want to bound the number of β's for which $A(\alpha, \beta) \neq 0$; 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, $\beta_1 \geqslant 1, \ldots, \beta_t \geqslant 1, \beta_{t+1} = \cdots = \beta_n = 0$. Then, $P_j(x) = 1$ for $j = t+1, \ldots, n$ and

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

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 $\alpha_1 \geqslant 1, \ldots, \alpha_t \geqslant 1$; 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 $|\alpha| = d$, $s \leqslant d$ and thus $\binom{s+d-1}{d} \leqslant \binom{2d-1}{d}$.

3. Concluding remarks

3.1. On the definition of a PTAS

Definitions 1.4 (ε-approximation) and 1.5 (PTAS) are crucial for our complexity result in Theorem 1.7 to hold. Indeed, consider defining an ε-approximation via an $x_\varepsilon \in \Delta$ satisfying

$$p(x_\varepsilon) - p_{\min} \leqslant \varepsilon |p_{\min}| \tag{3.1}$$

for a given $\varepsilon > 0$. Such definition mimics the definition of an ε-approximation that is classically used for combinatorial optimization problems (see, e.g., [14, Chapter 17]).

We show here that one cannot obtain an ε-approximation for problem (1.1) in the sense of (3.1) for each $\varepsilon > 0$ in polynomial time, unless P = NP, not even for $d = 2$. 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 $p(x) := x^{\mathrm{T}}(I + A)x$. By the Motzkin–Straus theorem, the minimum of $p(x)$ over Δ is $1/\alpha(G)$, where $\alpha(G)$ is the maximum cardinality of a stable set in G. Thus, $p_{\min} = 1/\alpha(G) > 0$ and $p_{\max} \geqslant 2p_{\min}$ if $\alpha(G) \geqslant 2$, since $p_{\max} \geqslant 1 = p(e_i)$.

Lemma 3.1. *Given $x^* \in \Delta$, one can construct a stable set S for which $1/|S| \leqslant p(x^*)$ 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 $p(x) \leqslant p(x^*)$. 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 $(x_1, x_2) \in \Delta_0 := \{(x_1, x_2) \mid x_1, x_2 \geqslant 0,\ x_1 + x_2 = 1 - \sum_{i \geqslant 3} x_i^*\}$, $f(x_1, x_2)$ 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 $p(x) \leqslant p(x^*)$, 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 $p(x) \leqslant p(x^*)$. Now, $p(x) = \sum_{i \in S} x_i^2$ which, using Cauchy–Schwartz inequality, implies that $p(x) \geqslant 1/|S|$. □

Now let ε be given such that $0 < \varepsilon < 1$ and set $\varepsilon' := \varepsilon/(1 - \varepsilon)$. Assume that one can construct in polynomial time a point $x \in \Delta$ satisfying (3.1), i.e., $p(x) \leqslant (1/\alpha(G))(1 + \varepsilon')$. By Lemma 3.1, one can construct in polynomial time a stable set S such that $1/|S| \leqslant p(x) \leqslant (1/\alpha(G))(1 + \varepsilon')$, which implies $|S| \geqslant \alpha(G)(1 - \varepsilon)$. This shows therefore the existence of a PTAS for the maximum stable set problem, contradicting an inapproximability result of Arora et al. [1].

3.2. Sharper bounds in the cases $d = 2, 3$

Theorem 2.2 involves the constant $\binom{2d-1}{d}d^d$ as factor of $p_{\max} - p_{\min}$ (and thus Theorem 1.3 as well). This is a large constant which could perhaps be improved via a tighter analysis, although this would have no impact on the claim of existence of a PTAS. As a matter of fact, in the case of degree $d = 2, 3$ forms, we can prove a better constant in Theorem 1.3 by looking more closely at the form of the function ϕ involved in the proof of Pólya's result. The result for $d = 2$ (Theorem 3.2) is known from Bomze and De Klerk [5] and Nesterov [11], but the result for $d = 3$ in Theorem 3.3 is new.

In what follows, $e_1, \ldots, e_n$ denote the standard unit vectors in $\mathbb{R}^n$. Thus, p_{2e_i} denotes the coefficient of the monomial x_i^2 in $p(x)$.

Theorem 3.2. *Let p be a form of degree $d = 2$ and $r \geqslant 0$ an integer. Then,*

$$p_{\min} - p_{\min}^{(r)} \leqslant \frac{1}{r+1}\left(\max_{i=1,\ldots,n} p_{2e_i} - p_{\min}\right) \leqslant \frac{1}{r+1}(p_{\max} - p_{\min}), \tag{3.2}$$

$$p_{\Delta(r+2)} - p_{\min} \leqslant \frac{1}{r+2}\left(\max_{i=1,\ldots,n} p_{2e_i} - p_{\min}\right) \leqslant \frac{1}{r+2}(p_{\max} - p_{\min}). \tag{3.3}$$

Proof. By looking more closely at the form of the function $\phi(x)$ in (2.5), 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,

$$\phi_{\max} = \frac{1}{r+2}\max_{i=1,\ldots,n} p_{2e_i} \leqslant \frac{1}{r+2} p_{\max}. \tag{3.4}$$

As $w_r(2) = (r+1)/(r+2)$, together with (2.9) and (2.10), this implies that the inequalities (3.2) and (3.3) from Theorem 3.2 hold.

Moreover, $p_{\min}^{(r)} \geqslant 0$ for $r \geqslant \max_i p_{2e_i}/p_{\min} - 2$. That is, for degree 2 forms, Theorem 1.1 remains valid for such r (instead of (1.9)). □

Theorem 3.3. *Let p be a form of degree $d = 3$ and $r \geqslant 0$ an integer. Then,*

$$p_{\min} - p_{\min}^{(r)} \leqslant \frac{4(r+3)}{(r+1)(r+2)}(p_{\max} - p_{\min}), \tag{3.5}$$

$$p_{\Delta(r+3)} - p_{\min} \leqslant \frac{4}{r+3}(p_{\max} - p_{\min}). \tag{3.6}$$

Proof. When $d = 3$, one can verify that

$$\phi(x) = \sum_{i=1}^{n} p_{3e_i}(3tx_i^2 - 2t^2 x_i) + \sum_{1 \leqslant i < j \leqslant n} (p_{2e_i+e_j} + p_{e_i+2e_j}) t x_i x_j, \tag{3.7}$$

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

$$p_{\min} \leqslant p(e_i) = p_{3e_i} \leqslant p_{\max}, \tag{3.8}$$

$$p_{3e_i} + p_{3e_j} + p_{2e_i+e_j} + p_{e_i+2e_j} \leqslant 8 p_{\max}. \tag{3.9}$$

Using (3.9), we can bound the second sum in (3.7):

$$\sum_{i<j}(p_{2e_i+e_j}+p_{e_i+2e_j})x_ix_j \leqslant \sum_{i<j}(8p_{\max}-p_{3e_i}-p_{3e_j})x_ix_j = 8p_{\max}\sum_{i<j}x_ix_j-\sum_i p_{3e_i}x_i(1-x_i).$$

Therefore,

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

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

$$\phi_{\max}\leqslant \frac{4}{r+3}p_{\max}-\frac{r+5}{(r+3)^2}p_{\min}. \tag{3.10}$$

Together with (2.9), (2.10), and the fact that $w_r(3)=(r+1)(r+2)/(r+3)^2$, this implies that relations (3.5) and (3.6) from Theorem 3.3 hold. □

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.

3.3. A probabilistic approach for estimating the grid bounds $p_{\Delta(k)}$

Nesterov [11] 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 ($k\geqslant 0$) 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: $E(p(x_k))\geqslant p_{\Delta(k)}$. 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 $\{1,\dots,n\}$ distributed as follows:

$$\mathrm{Prob}(\zeta=i)=q_i \quad (i=1,\dots,n). \tag{3.11}$$

Consider the random process

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

where ζ_k are random independent variables distributed according to (3.11). In other words, y_{k+1} is y_k+e_i with probability q_i. Finally, define

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

Thus all x_k lie in the set $\Delta(k)$. The following computations are given in [11]:

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

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

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

Therefore, if q is a global minimizer of $p(x)$ over Δ and 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 $k = r + d$, which gives the estimate (1.6) [11, Lemma 3]. If $p(x)$ is a form of degree 2, then

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

with $k = r + 2$, which gives the estimate (3.3) [11, Theorem 2].

Remark 3.4. 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 $\phi'(q) \leqslant 4/(r+3)^2(p_{\max} - p_{\min})$. Combined with (3.10), this implies that

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

3.4. Approximating polynomials over polytopes

As observed by Nesterov [11], some results for the simplex can be extended to the problem of minimizing a degree d form $p(x)$ over a polytope

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

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 $k \geqslant 1$, 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,

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

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

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

However, 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,

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

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

Nevertheless, when the polytope P is given by its vertices and the number of vertices is polynomially bounded in terms of n, our results imply the existence of a PTAS for the minimization of a fixed degree form on P.

3.5. 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 $r \geqslant 0$, define the parameter

$$p_{\min,sos}^{(r)} := \max \lambda \quad \text{for which } \left(\sum_{i=1}^n x_i^2\right)^r \left(\tilde{p}(x) - \lambda \left(\sum_{i=1}^n x_i^2\right)^d\right) \in \Sigma^2. \tag{3.12}$$

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

$$p_{\min}^{(r)} \leqslant p_{\min,sos}^{(r)} \leqslant p_{\min} \quad \text{for all } r \geqslant 0.$$

The bound $p_{\min,sos}^{(r)}$ can be computed in polynomial time with an arbitrary precision for any *fixed* r. This follows from the well-known fact (see, e.g., [19]) 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 $p_{\min,sos}^{(r)}$ converge to $p_{\min}$ as $r \to \infty$.

Schmüdgen [22] proved (in a more general context) that every polynomial that 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 $r \geqslant 0$:

$$\begin{aligned} \max \lambda \quad & \text{such that } \tilde{p}(x) - \lambda = s_0(x) + \left(1 - \sum_{i=1}^n x_i^2\right) s_1(x) \\ & \text{where } s_0 \in \Sigma^2, \quad s_1 \in \mathbb{R}[x_1, \ldots, x_n], \quad \deg(s_0) \leqslant 2(r+d). \end{aligned} \tag{3.13}$$

It follows from Schmüdgen's theorem that these bounds also converge to $p_{\min}$ as $r \to \infty$. In fact, De Klerk et al. [6] show that the bounds (3.13) coincide with the semidefinite bounds $p_{\min,sos}^{(r)}$. 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.

3.6. 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 [11] 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 [7] 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 $r \geqslant 0$, define the parameter

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

Thus, $p_S^{(r)} \leqslant p_{\min,S}$ for all $r \geqslant 0$. Using a result of Reznick [21], Faybusovich [7] shows that, for $r \geqslant 2nd(2d-1)/4\ln 2 - n/2 - d$,

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

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

Acknowledgments

We thank two referees for their suggestions which helped improve the presentation of the paper. We also thank a referee for pointing out to us the references [2,10].

References

[1] S. Arora, C. Lund, R. Motwani, M. Sudan, 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, 1992, pp. 14–23.
[2] G. Ausiello, A. D'Atri, M. Protasi, Structure preserving reductions among convex optimization problems, J. Comput. System Sci. 21 (1980) 136–153.
[3] M. Bellare, P. Rogaway, The complexity of approximating a nonlinear program, Math. Programming 69 (1995) 429–441.
[5] I. Bomze, E. De Klerk, Solving standard quadratic optimization problems via linear, semidefinite and copositive programming, Global Optim. 24 (2) (2002) 163–185.
[6] E. De Klerk, M. Laurent, P.A. Parrilo, On the equivalence of algebraic approaches to the minimization of forms on the simplex, in: D. Henrion, A. Garulli (Eds.), Positive Polynomials in Control, Lecture Notes on Control and Information Sciences, Vol. 312, Springer, Berlin, 2005, pp. 121–133.
[7] L. Faybusovich, Global optimization of homogeneous polynomials on the simplex and on the sphere, in: C. Floudas, P. Pardalos (Eds.), Frontiers in Global Optimization, Kluwer Academic Publishers, Dordrecht, 2003.
[9] T.S. Motzkin, E.G. Straus, Maxima for graphs and a new proof of a theorem of Túran, Canad. J. Math. 17 (1965) 533–540.
[10] G.L. Nemhauser, L.A. Wolsey, M.L. Fisher, An analysis of approximations for maximizing submodular set functions, I, Math. Programming 14 (1978) 265–294.
[11] 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.
[12] Yu. Nesterov, H. Wolkowicz, Y. Ye, Semidefinite programming relaxations of nonconvex quadratic optimization, in: H. Wolkowicz, R. Saigal, L. Vandenberghe (Eds.), Handbook of Semidefinite Programming, Kluwer Academic Publishers, Norwell, MA, 2000, pp. 361–419.
[13] R.A. Nicolaides, On a class of finite elements generated by Lagrange interpolation, SIAM J. Numer. Anal. 9 (3) (1972) 435–445.
[14] C.H. Papadimitriou, K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Prentice-Hall, New Jersey, 1982.
[15] P.A. Parrilo, Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, Ph.D. Thesis, California Institute of Technology, May 2000.
[16] P.A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, Math. Programming Ser. B 96 (2003) 293–320.
[17] G. Pólya, Collected Papers, Vol. 2, MIT Press, Cambridge, MA, London, 1974 pp. 309–313.
[18] V. Powers, B. Reznick, A new bound for Pólya's theorem with applications to polynomials positive on polyhedra, J. Pure Appl. Algebra 164 (2001) 221–229.
[19] V. Powers, T. Wörmann, An algorithm for sums of squares of real polynomials, J. Pure Appl. Algebra 127 (1998) 99–104.
[20] B. Reznick, Sums of even powers of real linear forms, Mem. Amer. Math. Soc. 463 (1992).
[21] B. Reznick, Uniform denominators in Hilbert's seventeenth problem, Math. Z. 220 (1995) 75–97.
[22] K. Schmüdgen, The K-moment problem for compact semi-algebraic sets, Math. Ann. 289 (1991) 203–206.
[23] S. Vavasis, Approximation algorithms for concave quadratic programming, in: C.A. Floudas, P. Pardalos (Eds.), Recent Advances in Global Optimization, Princeton University Press, Princeton, NJ, 1992, pp. 3–18.
[24] L.F. Zuluaga, J.C. Vera, J.F. Peña, LMI approximations for cones of positive semidefinite forms, SIAM J. Optim. 16 (2006) 1076–1091.