A refined error analysis for fixed-degree polynomial optimization over the simplex

Zhao Sun

Tilburg University
PO Box 90153, 5000 LE Tilburg, The Netherlands

Abstract

We consider the problem of minimizing a fixed-degree polynomial over the standard simplex. This problem is well known to be NP-hard, since it contains the maximum stable set problem in combinatorial optimization as a special case. In this paper, we revisit a known upper bound obtained by taking the minimum value on a regular grid, and a known lower bound based on Pólya's representation theorem. More precisely, we consider the difference between these two bounds and we provide upper bounds for this difference in terms of the range of function values. Our results refine the known upper bounds in the quadratic and cubic cases, and they asymptotically refine the known upper bound in the general case.

Keywords: Polynomial optimization over the simplex, Global optimization, Nonlinear optimization

2010 MSC: 90C26, 90C30

1. Introduction

Consider the problem of minimizing a homogeneous polynomial $f \in \mathbb{R}[x]$ of degree d on the (standard) simplex

$$\Delta_n := \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\}.$$

That is, the global optimization problem:

$$\underline{f} := \min_{x \in \Delta_n} f(x), \quad \text{or} \quad \overline{f} := \max_{x \in \Delta_n} f(x). \tag{1}$$

Here we focus on the problem of computing the minimum $\underline{f}$ of f over Δ_n. This problem is well known to be NP-hard, as it contains the maximum stable set problem as a special case (when f is quadratic). Indeed, given a graph $G = (V, E)$ with adjacency matrix A, Motzkin and Straus [8] show that the maximum stability number $\alpha(G)$ can be obtained by

$$\frac{1}{\alpha(G)} = \min_{x \in \Delta_{|V|}} x^T (I + A)x,$$

where I denotes the identity matrix. For problem (1), many approximation algorithms have been studied in the literature. In fact, when f has fixed degree d, there is a polynomial time approximation scheme (PTAS) for this problem, see [1] for the case $d = 2$ and [5, 7] for $d \geq 2$. For more results on its computational complexity, we refer to [3, 4].

We consider the following two bounds for $\underline{f}$: an upper bound $f_{\Delta(n,r)}$ obtained by taking the minimum value on a regular grid and a lower bound $f_{\min}^{(r-d)}$ based on Pólya's representation theorem. They both have been studied in the literature, see e.g. [1, 5, 7] for $f_{\Delta(n,r)}$ and [5, 14, 15] for $f_{\min}^{(r-d)}$. The two ranges $f_{\Delta(n,r)} - \underline{f}$ and $\underline{f} - f_{\min}^{(r-d)}$ have been studied separately and upper bounds for each of them have been shown in the above mentioned works.

In this paper, we study these two ranges at the same time. More precisely, we analyze the larger range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ and provide upper bounds for it in terms of the range of function values $\overline{f} - \underline{f}$. Of course, upper bounds for the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ can be obtained by combining the known upper bpunds for each of the two ranges $f_{\Delta(n,r)} - \underline{f}$ and $\underline{f} - f_{\min}^{(r-d)}$. Our new upper bounds for $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ refine these known bounds in the quadratic and cubic cases and provide an asymptotic refinement for general degree d.

1.1. Preliminaries

Throughout $\mathcal{H}_{n,d}$ denotes the set of all homogeneous polynomials in n variables with degree d.

Upper bounds using regular grids

One can construct an upper bound for $\underline{f}$ by taking the minimum of f on the regular grid

$$\Delta(n,r) := \{x \in \Delta_n : rx \in \mathbb{N}^n\},$$

for an integer $r \geq 0$. We define

$$f_{\Delta(n,r)} := \min_{x \in \Delta(n,r)} f(x).$$

Obviously, $\underline{f} \leq f_{\Delta(n,r)} \leq \overline{f}$, and $f_{\Delta(n,r)}$ can be computed by $|\Delta(n,r)| = \binom{n+r-1}{r}$ evaluations of f. In fact, when considering polynomials f of fixed degree d, the parameters $f_{\Delta(n,r)}$ (with increasing values of r) provide a PTAS for (1), as was proved by Bomze and de Klerk [1] (for $d = 2$), and by de Klerk et al. [5] (for $d \geq 2$). Recently, de Klerk et al. [7] provide an alternative proof for this PTAS and refine the error bound for $f_{\Delta(n,r)} - \underline{f}$ from [5] for cubic f.

In addition, some researchers study the properties of the regular grid $\Delta(n,r)$. For instance, given a point $x \in \Delta_n$, Bomze et al. [2] show a scheme to find the closest point to x on $\Delta(n,r)$ with respect to some class of norms including ℓ_p-norms for $p \geq 1$.

Lower bounds based on Pólya's representation theorem

Given a polynomial $f \in \mathcal{H}_{n,d}$, Pólya [12] shows that if f is positive over the simplex Δ_n, then the polynomial $(\sum_{i=1}^{n} x_i)^r f$ has nonnegative coefficients for any r large enough (see [13] for an explicit bound for r). Based on this result of Pólya, an asymptotically converging hierarchy of lower bounds for $\underline{f}$ can be constructed as follows: for any integer $r \geq d$, we define the parameter $f_{\min}^{(r-d)}$ as

$$f_{\min}^{(r-d)} := \max \lambda \ \text{ s.t. } \ \left(\sum_{i=1}^{n} x_i\right)^{r-d}\left(f - \lambda\left(\sum_{i=1}^{n} x_i\right)^d\right) \ \text{ has nonnegative coefficients.} \quad (2)$$

Notice that $\underline{f}$ can be equivalently formulated as

$$\underline{f} = \max \ \lambda \ \text{ s.t. } \ f(x) - \lambda\left(\sum_{i=1}^{n} x_i\right)^d \geq 0 \ \ \forall x \in \mathbb{R}_+^n.$$

Then, one can easily check the following inequalities:

$$f_{\min}^{(0)} \leq f_{\min}^{(1)} \leq \cdots \leq \underline{f} \leq f_{\Delta(n,r)} \leq \overline{f}.$$

Parrilo [9, 10] first introduces the idea of applying Pólya's representation theorem to construct hierarchical approximations in copositive optimization. De Klerk et al. [5] consider $f_{\min}^{(r-d)}$ and show upper bounds for $\underline{f} - f_{\min}^{(r-d)}$ in terms of $\overline{f} - \underline{f}$. Furthermore, Yildirim [15] and Sagol and Yildirim [14] analyze error bounds for $f_{\min}^{(r-2)}$ for quadratic f.

Now we give an explicit formula for the parameter $f_{\min}^{(r-d)}$, which follows from [13, relation (3)]; note that the quadratic case of this formula has also been observed in [11, 14, 15].

3

Lemma 1. *For $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$, one has*

$$f_{\min}^{(r-d)} = \min_{\substack{\alpha \in I(n,r)}} \sum_{\beta \in I(n,d)} f_\beta \frac{\alpha^{\underline{\beta}}}{r^{\underline{d}}}, \tag{3}$$

where we set $I(n,d) = \{\alpha \in \mathbb{N}^n : |\alpha| = d\}$, $r^{\underline{d}} := r(r-1)(r-2)\cdots(r-d+1)$, and $\alpha^{\underline{\beta}} := \prod_{i=1}^n \alpha_i^{\underline{\beta_i}}$.

PROOF. By using the multinomial theorem $(\sum_{i=1}^n x_i)^d = \sum_{\alpha \in I(n,d)} \frac{d!}{\alpha!} x^\alpha$, we obtain

$$\left(\sum_{i=1}^n x_i\right)^{r-d} f - \lambda \left(\sum_{i=1}^n x_i\right)^r = \left(\sum_{\gamma \in I(n,r-d)} \frac{(r-d)!}{\gamma!} x^\gamma\right)\left(\sum_{\beta \in I(n,d)} f_\beta x^\beta\right) - \lambda \left(\sum_{\alpha \in I(n,r)} \frac{r!}{\alpha!} x^\alpha\right)$$

$$= \sum_{\alpha \in I(n,r)} \left(\sum_{\beta \in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}}\right) \frac{r!}{\alpha!} x^\alpha - \lambda \left(\sum_{\alpha \in I(n,r)} \frac{r!}{\alpha!} x^\alpha\right)$$

$$= \sum_{\alpha \in I(n,r)} \left(\sum_{\beta \in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}} - \lambda\right) \frac{r!}{\alpha!} x^\alpha.$$

Hence, by definition (2), we obtain

$$f_{\min}^{(r-d)} = \max \quad \lambda \quad \text{s.t} \quad \sum_{\beta \in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}} - \lambda \geq 0 \quad \forall \alpha \in I(n,r)$$

$$= \min \sum_{\beta \in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}} \quad \text{s.t} \quad \alpha \in I(n,r).$$

$\square$

Similarly as $f_{\Delta(n,r)}$, by (3), the computation of $f_{\min}^{(r-d)}$ requires $|I(n,r)| = \binom{n+r-1}{r}$ evaluations of the polynomial $\sum_{\beta \in I(n,d)} f_\beta \alpha^{\underline{\beta}} \frac{1}{r^{\underline{d}}}$.

Bernstein coefficients

For any polynomial $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$, we can write it as

$$f = \sum_{\beta \in I(n,d)} f_\beta x^\beta = \sum_{\beta \in I(n,d)} \left(f_\beta \frac{\beta!}{d!}\right) \frac{d!}{\beta!} x^\beta. \tag{4}$$

For any $\beta \in I(n,d)$, we call $f_\beta \frac{\beta!}{d!}$ the *Bernstein coefficients* of f (this terminology has also been used in [6, 7]), since they are the coefficients of the polynomial f when f is expressed in the Bernstein basis $\{\frac{d!}{\beta!} x^\beta : \beta \in I(n,d)\}$ of $\mathcal{H}_{n,d}$. Applying the multinomial theorem together with (4), one can

4

obtain that when evaluating f at a point $x \in \Delta_n$, $f(x)$ is a convex combination of the Bernstein coefficients $f_\beta \frac{\beta!}{d!}$. Therefore, we have

$$\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \leq \underline{f} \leq f_{\Delta(n,r)} \leq \overline{f} \leq \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!}. \tag{5}$$

For the analysis in Section 5, we need the following result of [5], which bounds the range of the Bernstein coefficients of f in terms of its range of values $\overline{f} - \underline{f}$.

Theorem 1. *[5, Theorem 2.2] For any polynomial $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$, one has*

$$\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \leq \binom{2d-1}{d} d^d (\overline{f} - \underline{f}).$$

1.2. Contribution of the paper

In this paper, we consider upper bounds for $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$ in terms of $\overline{f} - \underline{f}$. More precisely, we provide tighter upper bounds in the quadratic, cubic, and square-free (aka multilinear) cases and, in the general case $d \geq 2$, our upper bounds are asymptotically tighter when r is large enough. We will apply the formula (3) directly for the quadratic, cubic and square-free cases, while for the general case we will use Theorem 1.

There are some relevant results in the literature. De Klerk et al. [5] give upper bounds for $f_{\Delta(n,r)} - \underline{f}$ (the upper bound for cubic f has been refined by de Klerk et al. [7]) and for $\underline{f} - f_{\min}^{(r-d)}$ in terms of $\overline{f} - \underline{f}$, and by adding them up one can easily derive upper bounds for $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$. Furthermore, for quadratic polynomial f, Yildirim [15] considers the upper bound $\min_{k \leq r} f_{\Delta(n,k)}$ for $\underline{f}$ (for $r \geq 2$) and upper bounds the range $\min_{k \leq r} f_{\Delta(n,k)} - f_{\min}^{(r-d)}$ in terms of $\overline{f} - \underline{f}$. Our results in this paper refine the results in [5, 7, 15] for the quadratic and cubic cases (see Sections 2 and 3 respectively), while for the general case our result refines the result of [5] when r is sufficiently large (see Section 5).

Structure

The paper is organized as follows. In Sections 2 and 3, we consider the quadratic and cubic cases respectively, and refine the relevant results obtained from [5, 7, 15]. Then, we look at the square-free (aka multilinear) case in Section 4. Moreover, in Section 5, we consider general (fixed-degree) polynomials and compare our new result with the one of [5].

Notation

Throughout, we let $[n] := \{1, 2, \ldots, n\}$. We denote $\mathbb{R}^n_+$ as the set of all nonnegative real vectors, and $\mathbb{N}^n$ as the set of all nonnegative integral vectors. For $\alpha \in \mathbb{N}^n$, we define $|\alpha| := \sum_{i=1}^n \alpha_i$ and $\alpha! := \alpha_1! \alpha_2! \cdots \alpha_n!$. As before, denote $I(n, d) := \{\alpha \in \mathbb{N}^n : |\alpha| = d\} = d\Delta(n, d)$. We let e denote the all-ones vector and e_i denote the i-th standard unit vector. We denote $\mathbb{R}[x]$ as the set of all multivariate polynomials in n variables (i.e. $x_1, x_2 \ldots, x_n$) and denote $\mathcal{H}_{n,d}$ as the set of all multivariate homogeneous polynomials in n variables with degree d. For $\alpha \in \mathbb{N}^n$, we denote $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$, while for $I \subseteq [n]$, we let $x^I := \prod_{i \in I} x_i$. Moreover, we denote $x^{\underline{d}} := x(x-1)(x-2)\cdots(x-d+1)$ for integer $d \geq 0$ and $x^{\underline{\alpha}} := \prod_{i=1}^n x_i^{\underline{\alpha_i}}$ for $\alpha \in \mathbb{N}^n$. Thus, $x^{\underline{d}} = 0$ if x is an integer with $0 \leq x \leq d-1$.

2. The quadratic case

For any quadratic polynomial f, we consider the range $f_{\Delta(n,r)} - f_{\min}^{(r-2)}$ and derive the following upper bound in terms of $\overline{f} - \underline{f}$.

Theorem 2. *For any quadratic $f = x^T Q x$ and $r \geq 2$, one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-2)} \leq \frac{1}{r-1}(Q_{\max} - f_{\Delta(n,r)}) \leq \frac{1}{r-1}(\overline{f} - \underline{f}), \tag{6}$$

where $Q_{\max} := \max_{i \in [n]} Q_{ii}$.

PROOF. By (3), we have

$$f_{\min}^{(r-2)} = \min_{\alpha \in I(n,r)} \frac{1}{r(r-1)} \left[f(\alpha) - \sum_{i=1}^n Q_{ii} \alpha_i \right].$$

Hence, $\frac{r-1}{r} f_{\min}^{(r-2)} = \min_{\alpha \in I(n,r)} \left[f(\frac{\alpha}{r}) - \sum_{i=1}^n Q_{ii} \frac{\alpha_i}{r} \frac{1}{r} \right]$. We obtain

$$\frac{r-1}{r} f_{\min}^{(r-2)} \geq \min_{\alpha \in I(n,r)} f(\frac{\alpha}{r}) - \max_{\alpha \in I(n,r)} \frac{1}{r} \sum_{i=1}^n Q_{ii} \frac{\alpha_i}{r} = f_{\Delta(n,r)} - \frac{1}{r} Q_{\max}. \tag{7}$$

One can easily obtain the first inequality in (6) by (7). For the second inequality in (6), we use the fact that $Q_{\max} \leq \overline{f}$ (since $Q_{ii} = f(e_i) \leq \overline{f}$ for $i \in [n]$), as well as the fact that $f_{\Delta(n,r)} \geq \underline{f}$. □

Now we point out that our result (6) refines the relevant result of [5]. De Klerk et al. [5] show the following theorem.

6

Theorem 3. *[5, Theorem 3.2] Suppose $f \in \mathcal{H}_{n,2}$ and $r \geq 2$. Then*

$$\underline{f} - f^{(r-2)}_{\min} \leq \frac{1}{r-1}(\overline{f} - \underline{f}), \tag{8}$$

$$f_{\Delta(n,r)} - \underline{f} \leq \frac{1}{r}(\overline{f} - \underline{f}). \tag{9}$$

By adding up (8) and (9), one gets

$$f_{\Delta(n,r)} - f^{(r-2)}_{\min} \leq \left(\frac{1}{r-1} + \frac{1}{r}\right)(\overline{f} - \underline{f}),$$

which is implied by our result (6).

Moreover, in [15], Yildirim considers one hierarchical upper bound of $\underline{f}$ (when f is quadratic), which is defined by $\min_{k \leq r} f_{\Delta(n,k)}$. One can easily verify that

$$f^{(r-2)}_{\min} \leq \underline{f} \leq \min_{k \leq r} f_{\Delta(n,k)} \leq f_{\Delta(n,r)}.$$

In [15, Theorem 4.1], Yildirim shows $\min_{k \leq r} f_{\Delta(n,k)} - f^{(r-2)}_{\min} \leq \frac{1}{r-1}(Q_{\max} - \underline{f})$, which can also be easily implied by our result (6).

The following example shows that the upper bound (6) can be tight.

Example 1. *[7, Example 2]* Consider the quadratic polynomial $f = \sum_{i=1}^{n} x_i^2$. As f is convex, one can check that $\underline{f} = \frac{1}{n}$ (attained at $x = \frac{1}{n}e$) and $\overline{f} = 1$ (attained at any standard unit vector). To compute $f_{\Delta(n,r)}$, we write r as $r = kn + s$, where $k \geq 0$ and $0 \leq s < n$. Then one can check that

$$f_{\Delta(n,r)} = \frac{1}{n} + \frac{1}{r^2}\frac{s(n-s)}{n}.$$

By (3), we have

$$f_{\Delta(n,r)} - f^{(r-2)}_{\min} = \frac{1}{r-1}(\overline{f} - \underline{f}) - \frac{1}{r^2(r-1)}\frac{s(n-s)}{n}.$$

Hence, for this example, the upper bound (6) is tight when $s = 0$.

3. The cubic case

For any cubic polynomial f, we consider the difference $f_{\Delta(n,r)} - f^{(r-3)}_{\min}$ and show the following result.

Theorem 4. *For any cubic polynomial f and $r \geq 3$, one has*

$$f_{\Delta(n,r)} - f^{(r-3)}_{\min} \leq \frac{4r}{(r-1)(r-2)}(\overline{f} - \underline{f}). \tag{10}$$

7

PROOF. We can write any cubic polynomial f as

$$f = \sum_{i=1}^{n} f_i x_i^3 + \sum_{i<j}(f_{ij}x_i x_j^2 + g_{ij}x_i^2 x_j) + \sum_{i<j<k} f_{ijk}x_i x_j x_k.$$

Then, by (3) one can check that

$$\frac{(r-1)(r-2)}{r^2} f_{\min}^{(r-3)}$$

$$= \min_{\alpha \in I(n,r)} \left\{ f(\frac{\alpha}{r}) - \frac{1}{r^3}\left(3\sum_{i=1}^{n} f_i \alpha_i^2 - 2\sum_{i=1}^{n} f_i \alpha_i + \sum_{i<j}(f_{ij}+g_{ij})\alpha_i \alpha_j\right)\right\}$$

$$\geq f_{\Delta(n,r)} - \frac{1}{r}\max_{\alpha \in I(n,r)}\left\{3\sum_{i=1}^{n} f_i\left(\frac{\alpha_i}{r}\right)^2 + \sum_{i<j}(f_{ij}+g_{ij})\left(\frac{\alpha_i}{r}\right)\left(\frac{\alpha_j}{r}\right)\right\} + \frac{1}{r^2}\min_{\alpha \in I(n,r)} 2\sum_{i=1}^{n} f_i \frac{\alpha_i}{r}$$

$$\geq f_{\Delta(n,r)} - \frac{1}{r}\max_{x \in \Delta_n}\left\{3\sum_{i=1}^{n} f_i x_i^2 + \sum_{i<j}(f_{ij}+g_{ij})x_i x_j\right\} + \frac{1}{r^2}\min_{x \in \Delta_n} 2\sum_{i=1}^{n} f_i x_i. \tag{11}$$

Evaluating f at e_i and $(e_i + e_j)/2$ yields, respectively, the relations:

$$\underline{f} \leq f_i \leq \overline{f}, \tag{12}$$

$$f_i + f_j + f_{ij} + g_{ij} \leq 8\overline{f}. \tag{13}$$

Using (13) and the fact that $\sum_{i=1}^{n} x_i = 1$, one can obtain

$$\sum_{i<j}(f_{ij}+g_{ij})x_i x_j \leq \sum_{i<j}(8\overline{f} - f_i - f_j)x_i x_j = 8\overline{f}\sum_{i<j} x_i x_j - \sum_{i=1}^{n} f_i x_i(1 - x_i). \tag{14}$$

By (11), (12), (14) and the fact that $\sum_{i=1}^{n} x_i = 1$, one can get

$$(r-1)(r-2)f_{\min}^{(r-3)} \geq r^2 f_{\Delta(n,r)} - 4r\overline{f} + (r+2)\min_{x \in \Delta_n}\sum_{i=1}^{n} f_i x_i \geq r^2 f_{\Delta(n,r)} - 4r\overline{f} + (r+2)\underline{f}.$$

Hence, one has

$$(r-1)(r-2)\left(f_{\Delta(n,r)} - f_{\min}^{(r-3)}\right) \leq 4r\overline{f} - (3r-2)f_{\Delta(n,r)} - (r+2)\underline{f} \leq 4r(\overline{f} - \underline{f}).$$

$\square$

Now we observe that our result (10) refines the relevant upper bound obtained from [5, 7]. De Klerk et al. [5] show the following result.

Theorem 5. *[5, Theorem 3.3] Suppose $f \in \mathcal{H}_{n,3}$ and $r \geq 3$. Then*

$$\underline{f} - f_{\min}^{(r-3)} \leq \frac{4r}{(r-1)(r-2)}(\overline{f} - \underline{f}), \tag{15}$$

$$f_{\Delta(n,r)} - \underline{f} \leq \frac{4}{r}(\overline{f} - \underline{f}). \tag{16}$$

Recently, de Klerk et al. [7, Corollary 2] refine (16) to

$$f_{\Delta(n,r)} - \underline{f} \leq \left(\frac{4}{r} - \frac{4}{r^2}\right)(\overline{f} - \underline{f}). \tag{17}$$

As in the quadratic case (in Section 2), our new upper bound (10) implies the upper bound obtained by adding up (15) and (17).

4. The square-free case

Consider the square-free (aka multilinear) polynomial $f = \sum_{I:I\subseteq[n],|I|=d} f_I x^I \in \mathcal{H}_{n,d}$. We have the following result for the difference $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$.

Theorem 6. *For any square-free polynomial $f = \sum_{I:I\subseteq[n],|I|=d} f_I x^I$ and $r \geq d$, one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right)(\overline{f} - \underline{f}). \tag{18}$$

PROOF. From (3), one can easily check that

$$f_{\min}^{(r-d)} = \min_{\alpha \in I(n,r)} \sum_{I:I\subseteq[n],|I|=d} f_I \frac{\alpha^I}{r^{\underline{d}}} = \frac{1}{r^{\underline{d}}} \min_{\alpha \in I(n,r)} f(\alpha).$$

As a result, one can obtain

$$\frac{f_{\min}^{(r-d)}}{f_{\Delta(n,r)}} = \frac{r^d}{r^{\underline{d}}}.$$

For $d = 1$, the result (18) is clear.

Now we assume $d \geq 2$. Considering $\overline{f} \geq 0$ (as $f(e_i) = 0$ for any $i \in [n]$), we obtain

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} = \left(1 - \frac{r^d}{r^{\underline{d}}}\right) f_{\Delta(n,r)} \leq \left(1 - \frac{r^d}{r^{\underline{d}}}\right)\underline{f} \leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right)(\overline{f} - \underline{f}).$$

$\square$

The following example shows that our upper bound (18) can be tight.

9

Example 2. [7, Example 4] Consider the square-free polynomial $f = -x_1 x_2$. One can check $\overline{f} = 0$, $\underline{f} = -\frac{1}{4}$, and

$$f_{\Delta(2,r)} = \begin{cases} -\frac{1}{4} & \text{if } r \text{ is even,} \\ -\frac{1}{4} + \frac{1}{4r^2} & \text{if } r \text{ is odd.} \end{cases}$$

By (3), we have

$$f_{\Delta(2,r)} - f_{\min}^{(r-2)} = \begin{cases} \frac{1}{r-1} \left(\overline{f} - \underline{f} \right) & \text{if } r \text{ is even,} \\ \left(\frac{1}{r} + \frac{1}{r^2} \right) \left(\overline{f} - \underline{f} \right) & \text{if } r \text{ is odd.} \end{cases}$$

For this example, the upper bound (18) is tight when r is even.

5. The general case

Now, we consider an arbitrary polynomial $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$. We need the following notation to formulate our result. Consider the univariate polynomial $t^d - t^{\underline{d}}$ (in the variable t), which can be written as

$$t^d - t^{\underline{d}} = \sum_{k=1}^{d-1} (-1)^{d-k-1} a_{d-k} t^k, \tag{19}$$

for some positive scalars $a_1, a_2, \ldots, a_{d-1}$. We can show the following error bound for the range $f_{\Delta(n,r)} - f_{\min}^{(r-d)}$.

Theorem 7. *For any polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq d$, one has*

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \frac{\sum_{k=1}^{d-1} a_{d-k} r^k}{r^{\underline{d}}} \binom{2d-1}{d} d^d (\overline{f} - \underline{f}). \tag{20}$$

In the proof we will need the following Vandermonde-Chu identity (see [13] for a proof, or alternatively use induction on $d \geq 1$):

$$(\sum_{i=1}^{n} x_i)^{\underline{d}} = \sum_{\alpha \in I(n,d)} \frac{d!}{\alpha!} x^{\underline{\alpha}} \quad \forall x \in \mathbb{R}^n, \tag{21}$$

which is an analogue of the multinomial theorem $(\sum_{i=1}^{n} x_i)^d = \sum_{\alpha \in I(n,d)} \frac{d!}{\alpha!} x^\alpha$.

Now we prove Theorem 7.

PROOF. *(of Theorem 7)* From (3), we have

$$\frac{r^{\underline{d}}}{r^d} f_{\min}^{(r-d)} = \min_{\alpha \in I(n,r)} \left\{ \sum_{\beta \in I(n,d)} f_\beta \frac{\alpha^{\underline{\beta}}}{r^d} - \sum_{\beta \in I(n,d)} f_\beta \frac{\alpha^\beta - \alpha^{\underline{\beta}}}{r^d} \right\}.$$

10

From this we obtain the inequality:

$$\frac{r^{\underline{d}}}{r^d} f^{(r-d)}_{\min} \geq f_{\Delta(n,r)} - \max_{\alpha \in I(n,r)} \sum_{\beta \in I(n,d)} f_\beta \frac{\alpha^\beta - \alpha^{\underline{\beta}}}{r^d}. \tag{22}$$

We now focus on the summation $\sum_{\beta \in I(n,d)} f_\beta (\alpha^\beta - \alpha^{\underline{\beta}})$.

For any $\beta \in I(n,d)$ and $x \in \mathbb{R}^n$, we can write the polynomial $x^\beta - x^{\underline{\beta}}$ as

$$x^\beta - x^{\underline{\beta}} = \sum_{\gamma:|\gamma| \leq d-1} (-1)^{d-|\gamma|-1} c^\beta_\gamma x^\gamma, \tag{23}$$

for some nonnegative scalars c^β_γ (which is an analogue of (19)). We now claim that, for any fixed $k \in [d-1]$, the following identity holds:

$$\sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c^\beta_\gamma x^\gamma = (-1)^{d-k-1} a_{d-k} \Big(\sum_{i=1}^n x_i\Big)^k. \tag{24}$$

For this, observe that the polynomials at both sides of (24) are homogeneous of degree k. Hence (24) will follow if we can show that the equality holds after summing each side over $k \in [d-1]$. In other words, it suffices to show the identity:

$$\sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c^\beta_\gamma x^\gamma = \sum_{k=1}^{d-1} (-1)^{d-k-1} a_{d-k} \Big(\sum_{i=1}^n x_i\Big)^k.$$

By the definition of a_{d-k} in (19), the right side of the above equation is equal to $(\sum_{i=1}^n x_i)^d - (\sum_{i=1}^n x_i)^{\underline{d}}$. Hence, we only need to show

$$\sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c^\beta_\gamma x^\gamma = \Big(\sum_{i=1}^n x_i\Big)^d - \Big(\sum_{i=1}^n x_i\Big)^{\underline{d}}. \tag{25}$$

Summing over (23), we obtain

$$\sum_{\beta \in I(n,d)} \frac{d!}{\beta!} \big(x^\beta - x^{\underline{\beta}}\big) = \sum_{\beta \in I(n,d)} \sum_{\gamma:|\gamma| \leq d-1} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c^\beta_\gamma x^\gamma = \sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} (-1)^{d-|\gamma|-1} c^\beta_\gamma x^\gamma.$$

We can now conclude the proof of (25) (and thus of (24)). Indeed, by using the multinomial theorem and the Vandermonde-Chu identity (21), we see that the left-most side in the above relation is equal to $(\sum_{i=1}^n x_i)^d - (\sum_{i=1}^n x_i)^{\underline{d}}$.

11

We partition $[d-1]$ as $[d-1] = I_o \cup I_e$, where $I_o := \{k : k \in [d-1], d-k \text{ is odd}\}$ and $I_e := \{k : k \in [d-1], d-k \text{ is even}\}$. Then, from (23), the summation $\sum_{\beta \in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}})$ becomes

$$
\sum_{\beta \in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}}) = \sum_{\beta \in I(n,d)} f_\beta \sum_{\gamma:|\gamma|\leq d-1} (-1)^{d-|\gamma|-1} c_\gamma^\beta \alpha^\gamma
$$

$$
= \sum_{k=1}^{d-1} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} f_\beta (-1)^{d-|\gamma|-1} c_\gamma^\beta \alpha^\gamma
$$

$$
\leq \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} c_\gamma^\beta \alpha^\gamma - \left(\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} \sum_{\gamma \in I(n,k)} \sum_{\beta \in I(n,d)} \frac{d!}{\beta!} c_\gamma^\beta \alpha^\gamma.
$$

By (24) we obtain

$$
\sum_{\beta \in I(n,d)} f_\beta(\alpha^\beta - \alpha^{\underline{\beta}}) \leq \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} a_{d-k} (\sum_{i=1}^n \alpha_i)^k - \left(\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} a_{d-k} (\sum_{i=1}^n \alpha_i)^k.
$$

Combining with (22), we get

$$
r^{\underline{d}} f_{\min}^{(r-d)} \geq r^d f_{\Delta(n,r)} - \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} a_{d-k} r^k + \left(\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} a_{d-k} r^k.
$$

That is,

$$
r^{\underline{d}}(f_{\Delta(n,r)} - f_{\min}^{(r-d)}) \leq (r^{\underline{d}} - r^d) f_{\Delta(n,r)} + \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} a_{d-k} r^k - \left(\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} a_{d-k} r^k.
$$

Since $r^{\underline{d}} - r^d = \sum_{k=1}^{d-1} (-1)^{d-k} a_{d-k} r^k$, we obtain

$$
r^{\underline{d}}(f_{\Delta(n,r)} - f_{\min}^{(r-d)})
$$

$$
\leq \sum_{k=1}^{d-1} (-1)^{d-k} a_{d-k} r^k f_{\Delta(n,r)} + \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} a_{d-k} r^k - \left(\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} a_{d-k} r^k
$$

$$
= \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_o} a_{d-k} r^k + f_{\Delta(n,r)} \sum_{k \in I_e} a_{d-k} r^k - \left(\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k \in I_e} a_{d-k} r^k
$$

$$
- f_{\Delta(n,r)} \sum_{k \in I_o} a_{d-k} r^k.
$$

According to (5), one has $\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \leq f_{\Delta(n,r)} \leq \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!}$. Therefore, we have

$$
r^{\underline{d}}(f_{\Delta(n,r)} - f_{\min}^{(r-d)}) \leq \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right) \sum_{k=1}^{d-1} a_{d-k} r^k.
$$

That is,

$$
f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \frac{\sum_{k=1}^{d-1} a_{d-k} r^k}{r^{\underline{d}}} \left(\max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \right).
$$

Finally, together with Theorem 1, we can conclude the result of Theorem 7. $\square$

12

Now, we compare the following theorem by de Klerk et al. [5] with our new result (20).

Theorem 8. *[5, Theorem 1.3] Suppose $f \in \mathcal{H}_{n,d}$ and $r \geq d$. Then*

$$\underline{f} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - 1\right)\binom{2d-1}{d}d^d(\overline{f} - \underline{f}), \tag{26}$$

$$f_{\Delta(n,r)} - \underline{f} \leq \left(1 - \frac{r^{\underline{d}}}{r^d}\right)\binom{2d-1}{d}d^d(\overline{f} - \underline{f}). \tag{27}$$

By adding up (26) and (27), we obtain

$$f_{\Delta(n,r)} - f_{\min}^{(r-d)} \leq \left(\frac{r^d}{r^{\underline{d}}} - \frac{r^{\underline{d}}}{r^d}\right)\binom{2d-1}{d}d^d(\overline{f} - \underline{f}). \tag{28}$$

Lemma 2. *When r is large enough, the upper bound (20) refines the upper bound (28).*

PROOF. It suffices to show that $\frac{r^d}{r^{\underline{d}}} - \frac{r^{\underline{d}}}{r^d}$ is larger than $\frac{\sum_{k=1}^{d-1} a_{d-k}r^k}{r^{\underline{d}}}$ when r is sufficiently large. Since $\frac{r^d}{r^{\underline{d}}} - \frac{r^{\underline{d}}}{r^d} = \left(r^d - \frac{(r^{\underline{d}})^2}{r^d}\right)/r^{\underline{d}}$, we only need to compare $r^d - \frac{(r^{\underline{d}})^2}{r^d}$ and $\sum_{k=1}^{d-1} a_{d-k}r^k$. For the term $r^d - \frac{(r^{\underline{d}})^2}{r^d}$, one can check that the coefficient of r^d is 0 and the coefficient of r^{d-1} is $2a_1 > 0$. On the other hand, in the summation $\sum_{k=1}^{d-1} a_{d-k}r^k$, the coefficient of r^{d-1} is $a_1 > 0$. Therefore, when r is sufficiently large, $r^d - \frac{(r^{\underline{d}})^2}{r^d}$ is larger than $\sum_{k=1}^{d-1} a_{d-k}r^k$, by which we conclude the proof. $\square$

We illustrate the result in Lemma 2 in the case of quartic polynomials.

Example 3. Consider a polynomial $f \in \mathcal{H}_{n,4}$ written as

$$
\begin{aligned}
f &= \sum_{i=1}^n f_i x_i^4 + \sum_{i<j} f_{ij}x_i^3 x_j + g_{ij}x_i^2 x_j^2 + h_{ij}x_i x_j^3 + \sum_{i<j<k} f_{ijk}x_i^2 x_j x_k \\
&+ g_{ijk}x_i x_j^2 x_k + h_{ijk}x_i x_j x_k^2 + \sum_{i<j<k<l} f_{ijkl}x_i x_j x_k x_l.
\end{aligned}
$$

In this case, (20) reads

$$f_{\Delta(n,r)} - f_{\min}^{(r-4)} \leq \frac{6r^2 + 11r + 6}{(r-1)(r-2)(r-3)}\binom{7}{4}4^4(\overline{f} - \underline{f}), \tag{29}$$

while (28) reads

$$f_{\Delta(n,r)} - f_{\min}^{(r-4)} \leq \frac{12r^2 - 58r + 144 - \frac{193}{r} + \frac{132}{r^2} - \frac{36}{r^3}}{(r-1)(r-2)(r-3)}\binom{7}{4}4^4(\overline{f} - \underline{f}). \tag{30}$$

One can check that (29) refines (30) when $r \geq 10$.

13

Remark 1. We now consider the convergence rate of the sequence

$$\alpha_r := \frac{f_{\Delta(n,r)} - f_{\min}^{(r-d)}}{\overline{f} - \underline{f}} \quad r = 1, 2, \ldots$$

Suppose the degree of f is fixed. By (20), we have $\alpha_r = O(\frac{1}{r})$. As in Example 2, $\alpha_r = \Omega(\frac{1}{r})$ holds, we can conclude that the dependence of α_r on r in (20) is tight, in the sense that there does not exist any $\epsilon > 0$ such that $\alpha_r = O(\frac{1}{r^{1+\epsilon}})$.

In [7], de Klerk et al. consider the convergence rate of the sequence

$$\beta_r := \frac{f_{\Delta(n,r)} - \underline{f}}{\overline{f} - \underline{f}} \quad r = 1, 2, \ldots$$

They consider several examples, and all of them satisfy $\beta_r = O(\frac{1}{r^2})$. However, it is still an open question to determine the asymptotic convergence rate of β_r in general.

Acknowledgements

References

[1] Bomze, I.M., Klerk, E. de.: Solving standard quadratic optimization problems via semidefinite and copositive programming. *Journal of Global Optimization,* 24(2), 163-185 (2002)

[2] Bomze, I.M., Gollowitzer, S., Yildirim, E.A.: Rounding on the standard simplex: Regular grids for global optimization. *Journal of Global Optimization,* DOI: 10.1007/s10898-013-0126-2

[3] Klerk, E. de.: The complexity of optimizing over a simplex, hypercube or sphere: A short survey. *Central European Journal of Operations Research,* 16(2), 111-125 (2008)

[4] Klerk, E. de, Hertog, D. den, Elfadul, G.E.E.: On the complexity of optimization over the standard simplex. *European Journal of Operational Research,* 191, 773-785 (2008)

[5] Klerk, E. de, Laurent, M., Parrilo, P.: A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoretical Computer Science,* 361(2-3), 210-225 (2006)

[6] Klerk, E. de, Laurent, M.: Error bounds for some semidefinite programming approaches to polynomial optimization on the hyeprcube. *SIAM Journal on Optimization,* 20(6), 3104-3120, (2010)

[7] Klerk, E. de, Laurent, M., Sun, Z.: An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex. Preprint, arXiv:1311.0173 (2013)

[8] Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Túran, *Canadian Journal of Mathematics,* 17, 533-540 (1965)

[9] Parrilo, P.: Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization, Ph.D. Thesis, California Institute of Technology, May 2000.

[10] Parrilo, P.: Semidefinite programming relaxations for semialgebraic problems, *Mathematical Programming Series B,* 96, 293-320 (2003)

[11] Peña, J.C., Vera, J.C., Zuluaga, L.F.: Computing the stability number of a graph via linear and semidefinite programming. *SIAM Journal on Optimization,* 18(1), 87-105 (2007)

[12] Pólya, G.: Collected Papers, Vol.2, MIT Press, Cambridge, MA, London, 309-313 (1974)

[13] Powers, V., Reznick, B.: 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] Sagol, G., Yildirim, E.A.: Analysis of copositive optimization based bounds on standard quadratic optimization. Technical Report, Department of Industrial Engineering, Koc University, Sariyer, Istanbul, Turkey (2013)

[15] Yildirim, E.A.: On the accuracy of uniform polyhedral approximations of the copositive cone. *Optimization methods and software,* 27(1), 155-173 (2012)