

On the convergence rate of grid search for polynomial optimization over the simplex

Etienne de Klerk¹ · Monique Laurent² ·
Zhao Sun³ · Juan C. Vera¹

Received: 5 October 2015 / Accepted: 21 February 2016 / Published online: 7 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract We consider the approximate minimization of a given polynomial on the standard simplex, obtained by taking the minimum value over all rational grid points with given denominator $r \in \mathbb{N}$. It was shown in De Klerk et al. (SIAM J Optim 25(3):1498–1514, 2015) that the accuracy of this approximation depends on r as $O(1/r^2)$ if there exists a rational global minimizer. In this note we show that the rational minimizer condition is not necessary to obtain the $O(1/r^2)$ bound.

Keywords Polynomial optimization · Grid search · Convergence rate · Taylor’s theorem

✉ Zhao Sun
Zhao.Sun@polymtl.ca

Etienne de Klerk
E.deKlerk@uvt.nl

Monique Laurent
M.Laurent@cwil.nl

Juan C. Vera
J.C.Veralizcano@uvt.nl

- ¹ Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands
- ² Centrum Wiskunde and Informatica (CWI), Amsterdam and Tilburg University, CWI, Postbus 94079, 1090 GB Amsterdam, The Netherlands
- ³ Canada Excellence Research Chair in “Data Science for Real-time Decision making”, École Polytechnique de Montréal, C.P. 6079, Succ. Centre-ville, Montréal H3C 3A7, Canada

1 Introduction

We consider the problem of minimizing a polynomial f over the standard simplex

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

That is, the problem of finding

$$f_{\min, \Delta_n} = \min f(x) \text{ s.t. } x \in \Delta_n. \tag{1}$$

Analogously, we denote $f_{\max, \Delta_n} = \max_{x \in \Delta_n} f(x)$.

We consider the parameter $f_{\min, \Delta(n,r)}$ obtained by minimizing f over the regular grid $\Delta(n, r) = \{x \in \Delta_n : rx \in \mathbb{N}^n\}$, consisting of all rational points in Δ_n with denominator r . That is,

$$f_{\min, \Delta(n,r)} = \min f(x) \text{ s.t. } x \in \Delta(n, r).$$

Note that the calculation of $f_{\min, \Delta(n,r)}$ requires $|\Delta(n, r)| = \binom{n+r-1}{r}$ function evaluations. Thus it may be computed in polynomial time for fixed r .

Interestingly, the parameter $f_{\min, \Delta(n,r)}$ yields a polynomial-time approximation scheme (PTAS) for problem (1) for polynomials of fixed degree, in the sense of the following two theorems.

The first theorem deals with the quadratic function case, and is due to Bomze and De Klerk [3].

Theorem 1 ([3, Theorem 3.2]) *For any quadratic polynomial f and $r \geq 1$, one has*

$$f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{f_{\max, \Delta_n} - f_{\min, \Delta_n}}{r}.$$

One says that $f_{\min, \Delta(n,r)}$ approximates f_{\min, Δ_n} with relative accuracy $1/r$, where the relative accuracy is defined as the ratio $(f_{\min, \Delta(n,r)} - f_{\min, \Delta_n}) / (f_{\max, \Delta_n} - f_{\min, \Delta_n})$. Note that this definition of a PTAS is that one may approximate f_{\min, Δ_n} to within any fixed relative accuracy in polynomial time. (This definition was introduced in the late 1970s, see e.g. [1, 2] and the references therein.) In particular, for any fixed $\epsilon > 0$, one has relative accuracy at most ϵ for $r \geq 1/\epsilon$. (Recall that $f_{\min, \Delta(n,r)}$ may be computed in polynomial time for fixed r .)

The second theorem is an extension of the previous result to polynomial objectives of fixed degree, and is due to De Klerk, Laurent and Parrilo [7].

Theorem 2 ([7, Theorem 1.3]) *For any polynomial f of degree d and $r \geq 1$, one has*

$$\begin{aligned} f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} &\leq \left(1 - \frac{r^d}{r^d}\right) \binom{2d-1}{d} d^d (f_{\max, \Delta_n} - f_{\min, \Delta_n}) \\ &\leq \frac{C_d}{r} (f_{\max, \Delta_n} - f_{\min, \Delta_n}), \end{aligned}$$

where $r^d := r(r - 1) \cdots (r - d + 1)$ denotes the falling factorial and C_d is a constant depending only on d .

Once again, one has that $f_{\min, \Delta(n,r)}$ approximates f_{\min, Δ_n} with relative accuracy $O(1/r)$, if d is fixed. (Here the constant in the big-O notation depends on d only, i.e., for fixed d it is an absolute constant not depending on the polynomial f .)

The authors of [8] show that there does not exist an $\epsilon > 0$ and a constant $C > 0$ such that, for any quadratic form f ,

$$f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{C}{r^{1+\epsilon}} (f_{\max, \Delta_n} - f_{\min, \Delta_n}) \quad \forall r \in \mathbb{N},$$

so in this sense the $1/r$ bound on the relative accuracy is tight in Theorem 1.

On the other hand if, as opposed to the PTAS property, one is only interested in the dependence of the accuracy $f_{\min, \Delta(n,r)} - f_{\min, \Delta_n}$ on r , then one may obtain $O(1/r^2)$ bounds, as shown in [9]. Here the constant in the big-O notation may depend on the polynomial f . For example, for a quadratic polynomial f , De Klerk et al. [9] show the following result.

Theorem 3 ([9, Theorem 2.2]) *Let f be a quadratic polynomial, and let x^* be a global minimizer of f over Δ_n , with denominator m , i.e. $mx^* \in \mathbb{N}^n$. For all integers $r \geq 1$, one has*

$$f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{m}{r^2} (f_{\max, \Delta_n} - f_{\min, \Delta_n}).$$

Note that this result does not give a PTAS, since the relative error is m/r^2 . (To get a given relative accuracy $\epsilon > 0$, one needs $r \geq \sqrt{m/\epsilon}$, so that r then depends on the problem size.)

The proof of [9, Theorem 2.2] relied on the fact that for quadratic objective functions the problem (1) has a rational global minimizer. For higher degree objective functions, the authors of [9] could only prove the $O(1/r^2)$ bound under the (restrictive) assumption of the existence of a rational minimizer.

Theorem 4 ([9, Corollary 4.5 and Lemma 4.6]) *Let f be a polynomial of degree d and assume that f has a rational global minimizer over Δ_n (say, in $\Delta(n, m)$). Then, one has*

$$f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{mc_d}{r^2} (f_{\max, \Delta_n} - f_{\min, \Delta_n}),$$

for some constant c_d depending only on d , namely¹ $c_d = (d - 1)(d! - 1)d^{2d-1} \binom{2d-1}{d}$.

In this note we prove that the accuracy $f_{\min, \Delta(n,r)} - f_{\min, \Delta_n}$ is $O(1/r^2)$ without the rational minimizer assumption. More precisely we show that for any polynomial f there exists a constant C_f (depending on f) such that $f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{C_f}{r^2}$ for all $r \in \mathbb{N}$. We will give several bounds, involving different constants C_f . For the first

¹ This value of c_d can be easily derived from results in [9] (specifically from Theorem 4.1, Lemma 4.6 and its proof).

bound in Theorem 5, the constant C_f depends on the support of a global minimizer of f and the coefficients of f while, for the second bound in Theorem 7, it depends on the smallest positive component of the minimizer, the range of values $f_{\max, \Delta_n} - f_{\min, \Delta_n}$ and the degree of f .

The results in this note complement a growing literature on the complexity of polynomial optimization and interpolation on a simplex; see [3–5, 7–11] and the references therein.

Notation

For an integer $n \geq 1$, we let $[n] = \{1, 2, \dots, n\}$. We denote $\mathbb{N}_d^n = \{\alpha \in \mathbb{N}^n : \sum_{i=1}^n \alpha_i \leq d\}$, with \mathbb{N} the set of nonnegative integers. For $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{N}^n$, we set $x^\alpha = \prod_{i=1}^n x_i^{\alpha_i}$. Moreover, given a subset $I \subseteq [n]$, x_I denotes the vector in $\mathbb{R}^{|I|}$ that contains the components x_i with $i \in I$. Finally, the support of $x \in \mathbb{R}^n$ is the set $\text{Supp}(x) = \{i \in [n] : x_i \neq 0\}$.

2 Preliminary results

First we will show some auxiliary results about approximations by grid points.

Lemma 1 *Let $x^* \in \Delta_n$ with support $I = \{i \in [n] : x_i^* > 0\}$. Then, for each integer $r \geq 1$, there exists a point $\tilde{x} \in \Delta(n, r)$ such that*

$$\|x^* - \tilde{x}\|_\infty \leq \frac{1}{r} \left(1 - \frac{1}{|I|}\right) \text{ and } \tilde{x}_i = 0 \ \forall i \in [n] \setminus I. \tag{2}$$

Proof First, we set $\tilde{x}_i = 0$ at the positions $i \in [n] \setminus I$. Then, we define the values of \tilde{x}_i with $i \in I$. By [4, Theorem 7], there exists a grid point $x' \in \Delta(|I|, r)$ such that $\|x_I^* - x'\|_\infty \leq \frac{1}{r} \left(1 - \frac{1}{|I|}\right)$. Set $\tilde{x}_I = x'$ and we get a point $\tilde{x} \in \Delta(n, r)$ satisfying (2). □

Lemma 2 *Let x^* be a global minimizer of the polynomial f in Δ_n and let \tilde{x} be a point in $\Delta(n, r)$ satisfying (2). Then, one has*

$$\nabla f(x^*)^T (\tilde{x} - x^*) = 0.$$

Proof By assumption, x^* is an optimal solution of the optimization problem $\min\{f(x) : x \geq 0, e^T x = 1\}$. From the KKT (necessary) conditions (see, e.g., [6, Chapter 5.5.3]), we have that there exist $\mu \in \mathbb{R}$ and $\lambda \in \mathbb{R}_+^n$ such that $\nabla f(x^*) = -\mu e + \lambda$, and $\lambda_i x_i^* = 0$ for all $i \in [n]$. Then we have $\nabla f(x^*)^T (\tilde{x} - x^*) = -\mu e^T (\tilde{x} - x^*) + \lambda^T (\tilde{x} - x^*)$. Moreover, $e^T (\tilde{x} - x^*) = e^T \tilde{x} - e^T x^* = 1 - 1 = 0$ and $\lambda_i > 0$ implies $x_i^* = 0$ and thus $\tilde{x}_i = 0$, so that $\lambda^T \tilde{x} = 0 = \lambda^T x^*$. This shows $\nabla f(x^*)^T (\tilde{x} - x^*) = 0$. □

Lemma 3 Consider a polynomial $f = \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha x^\alpha$ of degree d . Then, for any point $x \in [0, 1]^n$, one has

$$\sum_{i,j \in [n]} \left| \nabla^2 f(x)_{i,j} \right| \leq d(d-1) \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha|.$$

Proof As $f(x) = \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha x^\alpha$, one has

$$\nabla^2 f(x)_{i,j} = \begin{cases} \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha \alpha_i \alpha_j x^{\alpha - e_i - e_j} & \text{for } i \neq j, \\ \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha \alpha_i (\alpha_i - 1) x^{\alpha - 2e_i} & \text{for } i = j + . \end{cases}$$

Thus, we have

$$\begin{aligned} & \sum_{i,j \in [n]} |\nabla^2 f(x)_{i,j}| \\ & \leq \sum_{i,j \in [n]: i \neq j} \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \alpha_i \alpha_j x^{\alpha - e_i - e_j} + \sum_{i=1}^n \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \alpha_i (\alpha_i - 1) x^{\alpha - 2e_i} \\ & \leq \sum_{i,j \in [n]: i \neq j} \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \alpha_i \alpha_j + \sum_{i=1}^n \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \alpha_i (\alpha_i - 1) \\ & = \sum_{i,j \in [n]} \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \alpha_i \alpha_j - \sum_{i=1}^n \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \alpha_i \\ & = \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha| \left(\left(\sum_{i=1}^n \alpha_i \right)^2 - \left(\sum_{i=1}^n \alpha_i \right) \right) \\ & \leq (d^2 - d) \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha|, \end{aligned}$$

where for the second inequality we use $x_i \in [0, 1]$ for any $i \in [n]$. □

3 Bounds in terms of the support of a global minimizer

In this section we prove the following result, which shows the $O(1/r^2)$ convergence for the upper bounds $f_{\min, \Delta(n,r)}$ without the restrictive assumption of a rational minimizer.

Theorem 5 Consider a polynomial $f = \sum_{\alpha \in \mathbb{N}_d^n} f_\alpha x^\alpha$ of degree d . Let x^* be a global minimizer of f in Δ_n with support $I = \{i \in [n] : x_i^* > 0\}$. Then, for all integers $r \geq 1$, one has

$$f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{d(d-1)}{2r^2} \left(1 - \frac{1}{|I|} \right)^2 \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha|.$$

Proof Let $x^* \in \Delta_n$ be a global minimizer of f in Δ_n , let $\tilde{x} \in \Delta(n, r)$ satisfying the condition (2) from Lemma 1, and set $h = \tilde{x} - x^*$. Using Taylor’s theorem, we can write:

$$f(\tilde{x}) - f(x^*) = f(x^* + h) - f(x^*) = \nabla f(x^*)^T h + \frac{1}{2} h^T \nabla^2 f(\zeta) h, \tag{3}$$

for some point ζ lying in the segment $[x^*, x^* + h] = [x^*, \tilde{x}] \subseteq \Delta_n$. By Lemma 2, we know that $\nabla f(x^*)^T h = 0$. Using (2) and Lemma 3, we can upper bound the second term as follows:

$$\frac{1}{2} h^T \nabla^2 f(\zeta) h \leq \frac{1}{2} \|h\|_\infty^2 \sum_{i,j=1}^n |\nabla^2 f(\zeta)_{i,j}| \leq \frac{d(d-1)}{2r^2} \left(1 - \frac{1}{|I|}\right)^2 \sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha|.$$

Combining with $f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq f(\tilde{x}) - f(x^*)$, this concludes the proof. \square

Note that when the support I of the global minimizer x^* is a singleton (i.e., x^* is a standard unit vector), $f_{\min, \Delta_n} = f_{\min, \Delta(n,r)}$ for any $r \geq 1$, which is consistent with the inequality in Theorem 5 (whose right hand side is equal to zero).

Note also that one can tighten the result of Theorem 5 by replacing the sum $\sum_{\alpha \in \mathbb{N}_d^n} |f_\alpha|$ by $\sum_{\alpha \in \mathbb{N}_d^n: \text{Supp}(\alpha) \subseteq I} |f_\alpha|$. For this, it suffices to apply Theorem 5 to the polynomial $g(x_1, \dots, x_{|I|}) = f(x_1, \dots, x_{|I|}, 0, \dots, 0) = \sum_{\alpha \in \mathbb{N}_d^n: \text{Supp}(\alpha) \subseteq I} f_\alpha x^\alpha$, after observing that x_I^* is a global minimizer of g over the simplex $\Delta_{|I|}$ and that $f_{\min, \Delta(n,r)} \leq g_{\min, \Delta(|I|,r)}$ and $f_{\min, \Delta_n} = g_{\min, \Delta_{|I|}}$.

We mention another variation of the bound in Theorem 5, where the quantity $\sum_\alpha |f_\alpha|$ is now replaced by $\sum_\alpha |g_\alpha|$ for an appropriate polynomial g (depending on the support of a global minimizer of f).

Corollary 1 *Consider a polynomial f of degree d . Let x^* be a global minimizer of f in Δ_n with support $I = \{i \in [n] : x_i^* > 0\}$, assumed to be equal to $\{1, \dots, |I|\}$. Define the $(|I| - 1)$ -variate polynomial $g(x_1, \dots, x_{|I|-1}) = f(x_1, \dots, x_{|I|-1}, 1 - \sum_{i=1}^{|I|-1} x_i, 0, \dots, 0)$ (with 0 at the positions $i \notin I$). For all integers $r \geq 1$, one has*

$$f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} \leq \frac{d(d-1)}{2r^2} \left(1 - \frac{1}{|I|}\right)^2 \sum_{\alpha \in \mathbb{N}_d^{|I|-1}} |g_\alpha|.$$

4 Bounds in terms of the smallest positive component of a global minimizer

We now give a different approach for the convergence rate of the bounds $f_{\min, \Delta(n,r)}$. We will use the following well-known Euler’s identity for homogeneous polynomials.

Theorem 6 (Euler’s Identity) *Let f be an n -variate homogeneous polynomial of degree d . Then, for all $k \leq d$,*

$$\sum_{i_1, \dots, i_k \in [n]} \frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} x_{i_1} \dots x_{i_k} = \frac{d!}{(d - k)!} f(x).$$

We start with several preliminary results that we will need for our main result in Theorem 7 below.

Lemma 4 *Consider a homogeneous polynomial f of degree $d \geq 1$, assumed to have nonnegative coefficients. Let x^* be a global minimizer of f on Δ_n and let $\tilde{x} \in \Delta(n, r)$ satisfying (2). Consider a scalar $s > 0$ such that $|\tilde{x}_i - x_i^*| \leq s x_i^*$ for all $i \in [n]$. Then, for all integers $r \geq 1$,*

$$f_{\min, \Delta(n, r)} - f_{\min, \Delta_n} \leq ((1 + s)^d - (1 + ds)) f_{\min, \Delta_n}.$$

Proof First note that, as f has nonnegative coefficients then, for all $k \geq 1, i_1, \dots, i_k \in [n]$ and $x \in \Delta_n$, we have

$$\frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} \geq 0. \tag{4}$$

Set $h = \tilde{x} - x^*$. Then, we have:

$$\begin{aligned} & f_{\min, \Delta(n, r)} - f_{\min, \Delta_n} \\ & \leq f(\tilde{x}) - f(x^*) \\ & = \sum_{k=1}^d \frac{1}{k!} \sum_{i_1, \dots, i_k \in [n]} \frac{\partial^k f(x^*)}{\partial x_{i_1} \dots \partial x_{i_k}} h_{i_1} \dots h_{i_k} \quad \text{(From Taylor’s theorem)} \\ & = \sum_{k=2}^d \frac{1}{k!} \sum_{i_1, \dots, i_k \in [n]} \frac{\partial^k f(x^*)}{\partial x_{i_1} \dots \partial x_{i_k}} h_{i_1} \dots h_{i_k} \quad \text{(from Lemma 2)} \\ & \leq \sum_{k=2}^d \frac{s^k}{k!} \sum_{i_1, \dots, i_k \in [n]} \frac{\partial^k f(x^*)}{\partial x_{i_1} \dots \partial x_{i_k}} x_{i_1}^* \dots x_{i_k}^* \quad \text{(using 4)} \\ & = \sum_{k=2}^d s^k \binom{d}{k} f(x^*) \quad \text{(from Theorem 6)} \\ & = ((1 + s)^d - (1 + ds)) f(x^*). \end{aligned}$$

□

Lemma 5 *Let $0 < \epsilon \leq 2/3$. For any scalar $s \geq 0$ such that $ds \leq \epsilon$, we have*

$$(1 + s)^d - (1 + ds) \leq (1 + \epsilon) \binom{d}{2} s^2.$$

Proof From the binomial theorem, we have: $(1 + s)^d - \left(1 + ds + \binom{d}{2}s^2\right) = \sum_{k=3}^d \binom{d}{k}s^k$. Hence it suffices to show that $\sum_{k=3}^d \binom{d}{k}s^k \leq \epsilon \binom{d}{2}s^2$. One can verify that $\binom{d}{k} \leq \frac{d^{k-2}}{3} \binom{d}{2}$ for all $k \geq 3$. Using this and $0 \leq ds \leq \epsilon \leq 2/3$ one obtains:

$$\sum_{k=3}^d \binom{d}{k}s^k \leq \frac{s^2}{3} \binom{d}{2} \sum_{k=3}^d (ds)^{k-2} \leq \frac{s^2}{3} \binom{d}{2} \sum_{k=1}^{\infty} \epsilon^k = \frac{s^2}{3} \binom{d}{2} \frac{\epsilon}{1 - \epsilon} \leq \epsilon \binom{d}{2} s^2.$$

□

Lemma 6 *Let $x^* \in \Delta_n$ be given. Let $r \geq 1$ and let $\tilde{x} \in \Delta(n, r)$ satisfying relation (2). Let x_{\min}^* be the smallest positive component of x^* . Then $|\tilde{x}_i - x_i^*| \leq \frac{1}{rx_{\min}^*} x_i^*$ for all $i \in [n]$.*

Proof Fix $i \in [n]$. If $x_i^* = 0$, then $\tilde{x}_i = 0$ by (2) and thus the desired inequality holds. Otherwise, $x_{\min}^* \leq x_i^*$ and thus $|\tilde{x}_i - x_i^*| \leq \frac{1}{r} \leq \frac{1}{rx_{\min}^*} x_i^*$. □

We can now state our main result of this section, which shows again (but with a different constant) that the parameter $f_{\min, \Delta(n, r)}$ approximates f_{\min, Δ_n} with accuracy in $O(1/r^2)$.

Theorem 7 *Let f be a polynomial of degree $d \geq 1$. Let x^* be a global minimizer of f on Δ_n with smallest positive component x_{\min}^* . Then, for any $0 < \epsilon \leq 2/3$, and $r \geq \frac{d}{\epsilon x_{\min}^*}$, one has*

$$f_{\min, \Delta(n, r)} - f_{\min, \Delta_n} \leq \frac{1}{r^2} \frac{(1 + \epsilon)d^d \binom{d}{2}}{(x_{\min}^*)^2} \binom{2d - 1}{d} (f_{\max, \Delta_n} - f_{\min, \Delta_n}).$$

Proof First observe that it suffices to show the result for homogeneous polynomials. Indeed, if $f = \sum_{\alpha} f_{\alpha} x^{\alpha}$ is not homogeneous, then we may consider instead the homogeneous polynomial $F(x) = \sum_{\alpha} f_{\alpha} x^{\alpha} (\sum_{i=1}^n x_i)^{d - |\alpha|}$ and the result for F will imply the result for f . Hence we now assume that $f = \sum_{\alpha \in \mathbb{N}_{=d}^n} f_{\alpha} x^{\alpha}$ is homogeneous, where $\mathbb{N}_{=d}^n = \{\alpha \in \mathbb{N}^n : \sum_{i=1}^n \alpha_i = d\}$.

Set $s = \frac{1}{rx_{\min}^*}$, so that $ds \leq \epsilon$. Assume first that the polynomial f has nonnegative coefficients. Then, using Lemmas 4, 5 and 6, we can conclude that

$$\begin{aligned} f_{\min, \Delta(n, r)} - f_{\min, \Delta_n} &\leq ((1 + s)^d - (1 + ds)) f_{\min, \Delta_n} \\ &\leq (1 + \epsilon) \binom{d}{2} s^2 f_{\min, \Delta_n} = \frac{(1 + \epsilon) \binom{d}{2}}{r^2 (x_{\min}^*)^2} f_{\min, \Delta_n}. \end{aligned} \tag{5}$$

In the general case when no sign condition is assumed on the coefficients of f , we get back to the preceding case by doing a suitable ‘shift’ on f . For this, define the parameters

$$\hat{f}_{\min} := \min_{\alpha \in \mathbb{N}_{=d}^n} f_{\alpha} \frac{\alpha!}{d!}, \quad \hat{f}_{\max} = \max_{\alpha \in \mathbb{N}_{=d}^n} f_{\alpha} \frac{\alpha!}{d!}$$

known, respectively, as the minimum and maximum Bernstein coefficients of f . Observe that, for any $x \in \Delta_n$, $\sum_{\alpha \in \mathbb{N}_{=d}^n} \frac{d!}{\alpha!} x^\alpha = 1$, and thus $f(x) = \sum_{\alpha \in \mathbb{N}_{=d}^n} f_\alpha \frac{\alpha!}{d!} \left(\frac{d!}{\alpha!} x^\alpha\right)$ is a convex combination of the Bernstein coefficients $f_\alpha \alpha! / d!$, which implies

$$\hat{f}_{\min} \leq f_{\min, \Delta_n} \leq f_{\max, \Delta_n} \leq \hat{f}_{\max}. \tag{6}$$

We now define the polynomial

$$g(x) = f(x) - \hat{f}_{\min} \left(\sum_{i=1}^n x_i\right)^d = \sum_{\alpha \in \mathbb{N}_{=d}^n} \left(f_\alpha - \hat{f}_{\min} \frac{d!}{\alpha!}\right) x^\alpha,$$

which is homogeneous of degree d and with nonnegative coefficients. Hence we can apply the above relation (5) to g and, since g and f have the same global minimizers on Δ_n , we deduce that

$$\begin{aligned} f_{\min, \Delta(n,r)} - f_{\min, \Delta_n} &= g_{\min, \Delta(n,r)} - g_{\min, \Delta_n} \leq \frac{(1 + \epsilon) \binom{d}{2}}{r^2 (x_{\min}^*)^2} g_{\min, \Delta_n} \\ &= \frac{(1 + \epsilon) \binom{d}{2}}{r^2 (x_{\min}^*)^2} (f_{\min, \Delta_n} - \hat{f}_{\min}). \end{aligned} \tag{7}$$

In view of (6), we have: $f_{\min, \Delta_n} - \hat{f}_{\min} \leq \hat{f}_{\max} - \hat{f}_{\min}$. Finally, combining with the inequality: $\hat{f}_{\max} - \hat{f}_{\min} \leq \binom{2d-1}{d} d^d (f_{\max, \Delta_n} - f_{\min, \Delta_n})$ shown in [7, Theorem 2.2], we can conclude the proof. \square

Note that Theorem 7 does not imply Theorem 4. Indeed, if there is a rational global minimizer $x^* \in \Delta(n, m)$, then $x_{\min}^* \geq 1/m$ so that Theorem 7 gives a $O(m^2/r^2)$ bound in terms of m and r , as opposed to the $O(m/r^2)$ bound in Theorem 4.

5 Comparison of bounds

We now consider the following seven polynomials, for which we compare the upper bounds for $f_{\min, \Delta(n,r)} - f_{\min, \Delta_n}$ obtained in Theorems 3 or 4 (depending on the degree of f), Theorem 5, Corollary 1, and Theorem 7:

$$\begin{aligned} f_1 &= \sum_{i=1}^n \left(x_i - \frac{1}{n}\right)^2, & f_2 &= \sum_{i=1}^n x_i^2, \\ f_3 &= -\sum_{i=1}^n x_i^2, & f_4 &= \left(x_1 - \frac{1}{m}\right)^2 + \left(x_2 - \frac{m-1}{m}\right)^2, \\ f_5 &= \left(x_1 - \frac{m-1}{2m}\right)^2 + \left(x_2 - \frac{m+1}{2m}\right)^2, & f_6 &= \sum_{i=1}^n x_i^d, & f_7 &= -\prod_{i=1}^d x_i. \end{aligned}$$

Table 1 Comparison of upper bounds

Polynomial	Theorems 3 or 4	Theorem 5	Corollary 1	Theorem 7
f_1	$\frac{n-1}{r^2} = O\left(\frac{n}{r^2}\right)$	$\frac{(n-1)^2(n+1)^2}{n^3r^2} = O\left(\frac{n}{r^2}\right)$	$\frac{(n-1)^3(n+1)^2}{n^3r^2} = O\left(\frac{n^2}{r^2}\right)$	$\frac{12(1+\epsilon)(n^2-n)}{r^2} = O\left(\frac{n^2}{r^2}\right)$
f_2	$\frac{n-1}{r^2} = O\left(\frac{n}{r^2}\right)$	$\frac{(n-1)^2}{nr^2} = O\left(\frac{n}{r^2}\right)$	$\frac{(n-1)^2(n^2+n-1)}{nr^2} = O\left(\frac{n^2}{r^2}\right)$	$\frac{12(1+\epsilon)(n^2-n)}{r^2} = O\left(\frac{n^2}{r^2}\right)$
f_3	$\frac{n-1}{nr^2} = O\left(\frac{1}{r^2}\right)$	0	0	$\frac{12(1+\epsilon)(n-1)}{nr^2} = O\left(\frac{1}{r^2}\right)$
f_4	$\frac{2(m-1)^2}{mr^2} = O\left(\frac{m}{r^2}\right)$	$\frac{5m^2-2m+2}{4m^2r^2} = O\left(\frac{1}{r^2}\right)$	$\frac{(m+1)^2}{2m^2r^2} = O\left(\frac{1}{r^2}\right)$	$\frac{24(1+\epsilon)(m-1)^2}{r^2} = O\left(\frac{m^2}{r^2}\right)$
f_5	$\frac{(m+1)^2}{mr^2} = O\left(\frac{m}{r^2}\right)$	$\frac{9m^2+1}{8m^2r^2} = O\left(\frac{1}{r^2}\right)$	$\frac{9m^2-6m+1}{8m^2r^2} = O\left(\frac{1}{r^2}\right)$	$\frac{24(1+\epsilon)(m+1)^2}{(m-1)^2r^2} = O\left(\frac{1}{r^2}\right)$
f_6	$\frac{d^{2d-1}\left(1-\frac{1}{d}\right)(n^d-n)(2d-1)!}{nr^{d-1}r^2(d-2)!} = O\left(\frac{nd^{3d}}{r^2}\right)$	$\frac{d(d-1)n}{2r^2}\left(1-\frac{1}{n}\right)^2 = O\left(\frac{d^2n}{r^2}\right)$	$\frac{d^2-d)(n^d+(-1)^d(n-1))}{2r^2n^2/(n-1)^2} = O\left(\frac{d^2nd}{r^2}\right)$	$\frac{d^d(1+\epsilon)(n^d-1)}{nr^{d-3}r^2} = O\left(\frac{d^d}{r^2}\right)$ $= O\left(\frac{d^{2d+2}n^2}{r^2}\right)$
f_7	$\frac{d^d(d-1)(d-1)(2d-1)}{r^2} = O\left(\frac{d^{3d+1}}{r^2}\right)$	$\frac{d(d-1)}{2r^2}\left(1-\frac{1}{d}\right)^2 = O\left(\frac{d^2}{r^2}\right)$	$\frac{d^2(d-1)}{2r^2}\left(1-\frac{1}{d}\right)^2 = O\left(\frac{d^3}{r^2}\right)$	$\frac{(1+\epsilon)d^d}{r^2} = O\left(\frac{d^{d+4}}{r^2}\right)$ $= O\left(\frac{d^d}{r^2}\right)$

Table 2 Possible relationships

	Theorems 3 or 4	Theorem 5	Corollary 1	Theorem 7
Theorems 3 or 4	–	$f_2^{(*)}, f_3, f_4, f_5, f_6, f_7$	f_3, f_4, f_5, f_7	f_5, f_7
Theorem 5	$f_1^{(*)}$	–	$f_4^{(*)}, f_5^{(*)}$	–
Corollary 1	f_1, f_2	f_1, f_2, f_6, f_7	–	–
Theorem 7	$f_1, f_2, f_3^{(*)}, f_4$	$f_1, f_2, f_3, f_4, f_5^{(*)}, f_6, f_7$	$f_1^{(*)}, f_2^{(*)}, f_3, f_4, f_5^{(*)}, f_7$	–

If entry ij in the table is f , it means that the bound indexed by column j is stronger than the bound indexed by row i for the function f

In the first three examples we restrict our attention to the cases when $n \geq 2$. In polynomial f_4 , we select $m \geq 2$, and in polynomial f_5 , we select m even, which implies $m - 1, m + 1$ and $2m$ are relatively prime. In polynomials f_6 and f_7 , we select $d \geq 3$. The results are shown in Table 1.

We now summarize the possible relationships between the various bounds in Table 2, which should be understood as follows. For instance, having the entry f_3 at the position (Theorems 3 or 4, Theorem 5) means that, for the polynomial f_3 , the bound of Theorem 5 is better than the bound of Theorems 3 or 4 and this is a strong dominance (since the improvement depends on the parameter n). When the improvement depends only on a constant we indicate this by marking the polynomial with an asterisk, as for instance for the entry $f_1^{(*)}$ at the position (Theorems 5, 3 or 4). In conclusion, we can see using the polynomials f_1, f_2, f_3, f_4, f_5 that there is no possible ordering of the bounds provided by Theorems 3 or 4, 5 and 7 and Corollary 1.

Acknowledgements We thank the associate editor and two anonymous referees for their comments which helped improve the presentation of the paper.

References

1. Ausiello, G., D’Atri, A., Protasi, M.: Structure preserving reductions among convex optimization problems. *J. Comput. Syst. Sci.* **21**(1), 136–153 (1980)
2. Bellare, M., Rogaway, P.: The complexity of approximating a nonlinear program. *Math. Program.* **69**(1), 429–441 (1995)
3. Bomze, I.M., De Klerk, E.: Solving standard quadratic optimization problems via semidefinite and copositive programming. *J. Global Optim.* **24**(2), 163–185 (2002)
4. Bomze, I.M., Gollowitzer, S., Yildirim, E.A.: Rounding on the standard simplex: Regular grids for global optimization. *J. Global Optim.* **59**(2–3), 243–258 (2014)
5. Bos, L.P.: Bounding the Lebesgue function for Lagrange interpolation in a simplex. *J. Approx. Theory* **38**, 43–59 (1983)
6. Boyd, S., Vandenberghe, L.: *Convex Optimization*. Cambridge University Press, Cambridge (2004)
7. De Klerk, E., Laurent, M., Parrilo, P.: A PTAS for the minimization of polynomials of fixed degree over the simplex. *Theoret. Comput. Sci.* **361**(2–3), 210–225 (2006)
8. De Klerk, E., Laurent, M., Sun, Z.: An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex. *Math. Program.* **151**(2), 433–457 (2015)
9. De Klerk, E., Laurent, M., Sun, Z.: An error analysis for polynomial optimization over the simplex based on the multivariate hypergeometric distribution. *SIAM J. Optim.* **25**(3), 1498–1514 (2015)

10. Faybusovich, L.: Global optimization of homogeneous polynomials on the simplex and on the sphere. In: Floudas, C., Pardalos, P. (eds.) *Frontiers in Global Optimization*, pp. 109–121. Kluwer Academic Publishers, Boston (2004)
11. Nesterov, Yu.: Random walk in a simplex and quadratic optimization over convex polytopes. CORE Discussion Paper 2003/71, CORE-UCL (2003)