SIAM J. OPTIM.
Vol. 25, No. 3, pp. 1498–1514

AN ERROR ANALYSIS FOR POLYNOMIAL OPTIMIZATION OVER THE SIMPLEX BASED ON THE MULTIVARIATE HYPERGEOMETRIC DISTRIBUTION*

ETIENNE DE KLERK†, MONIQUE LAURENT‡, AND ZHAO SUN†

Abstract. We study the minimization of fixed-degree polynomials over the simplex. This problem is well-known to be NP-hard, as it contains the maximum stable set problem in graph theory as a special case. In this paper, we consider a rational approximation by taking the minimum over the regular grid, which consists of rational points with denominator r (for given r). We show that the associated convergence rate is $O(1/r^2)$ for quadratic polynomials. For general polynomials, if there exists a rational global minimizer over the simplex, we show that the convergence rate is also of the order $O(1/r^2)$. Our results answer a question posed by De Klerk, Laurent, and Sun [*Math. Program.*, 151 (2015), pp. 433–457]. and improves on previously known $O(1/r)$ bounds in the quadratic case.

Key words. polynomial optimization over the simplex, global optimization, nonlinear optimization

AMS subject classifications. 90C26, 90C30

DOI. 10.1137/140976650

1. Introduction and preliminaries.

We consider optimization of polynomials over the standard simplex:

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

More precisely, given a polynomial $f \in \mathcal{H}_{n,d}$, where $\mathcal{H}_{n,d}$ denotes the set of n-variate homogeneous real polynomials of degree d, we define

$$\underline{f} := \min_{x \in \Delta_n} f(x) \tag{1.1}$$

and $\overline{f} := \max_{x \in \Delta_n} f(x)$. For computational complexity reasons, we assume throughout that the polynomial f has integer coefficients.

For quadratic $f \in \mathcal{H}_{n,2}$, Vavasis [18] shows that problem (1.1) admits a rational global minimizer x^*, whose bit-size is polynomial in the bit-size of the input data. On the other hand, when the degree of f is larger than 2, there exist polynomials f for which problem (1.1) does not have any rational global minimizer. This is the case, for instance, for the polynomial $f(x) = 2{x_1}^3 - x_1 \left(\sum_{i=1}^n x_i\right)^2$, whose global minimizer always has the irrational component $x_1 = 1/\sqrt{6}$.

Complexity and approximation results. The global optimization problem (1.1) is known to be NP-hard and contains the maximum stable set problem in graphs

---

*Received by the editors July 8, 2014; accepted for publication (in revised form) June 2, 2015; published electronically July 30, 2015.
http://www.siam.org/journals/siopt/25-3/97665.html

†Tilburg University, 5000 LE Tilburg, Netherlands (E.deKlerk@uvt.nl, Z.Sun@uvt.nl).
‡Centrum Wiskunde and Informatica (CWI), Amsterdam, and Tilburg University, CWI, 1090 GB Amsterdam, Netherlands (M.Laurent@cwi.nl).

as a special case. Indeed, for a graph $G = (V, E)$, Motzkin and Straus [12] show that its stability number $\alpha(G)$ can be calculated via

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

where I denotes the identity matrix and A_G denotes the adjacency matrix of graph G.

On the other hand, there exists a polynomial time approximation scheme (PTAS) for problem (1.1) over the class of polynomials $f \in \mathcal{H}_{n,d}$ with fixed degree d, as was shown by Bomze and De Klerk [2] for degree $d = 2$ and by De Klerk, Laurent, and Parrilo [7] for degree $d \geq 3$. The PTAS is easily described: It takes the minimum of f over the regular grid

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

for increasing values of $r \in \mathbb{N}$. Note that

$$f_{\Delta(n,r)} := \min_{x \in \Delta(n,r)} f(x) \tag{1.2}$$

may be computed by performing $|\Delta(n, r)| = \binom{n+r-1}{r}$ evaluations of f. Thus, for fixed r, $f_{\Delta(n,r)}$ can be obtained in polynomial (in n) time. The following error estimates have been shown for the range $f_{\Delta(n,r)} - \underline{f}$ in terms of the range $\overline{f} - \underline{f}$ of function values.

THEOREM 1.1 (see [2, Theorem 3.2]). *For any polynomial $f \in \mathcal{H}_{n,2}$ and $r \geq 1$, one has*

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

THEOREM 1.2 (see [7, Theorem 1.3]). *For any polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq 1$, one has*

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

where $r^{\underline{d}} := r(r-1)\cdots(r-d+1)$ denotes the falling factorial.

For more results about the computational complexity of problem (1.1), see [4, 5]; for properties of the grid $\Delta(n, r)$, see [3], and for recent studies of the approximation $f_{\Delta(n,r)}$, see [1, 8, 15, 16, 17].

De Klerk, Laurent, and Sun [8] recently provided alternative proofs of the PTAS results in Theorems 1.1 and 1.2. The idea of these proofs is to define a suitable discrete probability distribution on $\Delta(n, r)$ (seen as a sample space) by using the multinomial distribution. (This idea is an extension of a probabilistic argument by Nesterov [13]; for the exact connection, see [8, section 6].)

Recall that the multinomial distribution may be explained by considering a box filled with balls of n different colors, and where the fraction of balls of color $i \in \{1, \ldots, n\}$ is denoted by x_i, say. If one draws r balls randomly *with replacement* and we let the random variable Y_i denote the number of times that a ball of color i was drawn, then

$$\mathbf{Pr}\,[Y_1 = \alpha_1, \ldots, Y_n = \alpha_n] = \frac{r!}{\alpha!} x^\alpha, \qquad \alpha \in r\Delta(n, r),$$

where $\alpha! := \prod_{i=1}^n \alpha_i!$ and $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$. Defining the normalized random variable $X = \frac{1}{r}Y \in \Delta(n,r)$, one has

$$\mathbb{E}[f(X)] = \sum_{\alpha \in r\Delta(n,r)} f\left(\frac{\alpha}{r}\right)\frac{r!}{\alpha!}x^\alpha =: B_r(f)(x), \tag{1.3}$$

where $B_r(f)(x)$ is called the Bernstein approximation of f of order r at x. Therefore, since $f_{\Delta(n,r)} \leq \mathbb{E}[f(X)]$, the new PTAS proof in [8] is essentially a consequence of the properties of Bernstein approximation on the standard simplex.

This approach can be put in the more general context of the framework introduced by Lasserre [10, 11] based on reformulating any polynomial optimization problem as an optimization problem over measures. When applied to our setting, this implies the following upper bound:

$$f_{\Delta(n,r)} \leq \mathbb{E}_\mu(f) = \int_{\Delta(n,r)} f(x)\mu(dx)$$

for any probability measure μ on $\Delta(n,r)$. So the work [8] is based on selecting the multinomial distribution with appropriate parameters as measure μ. In this paper we will select another measure, as explained below.

Contribution of this paper. In this paper, we give a partial answer to a question posed in [8], concerning the error bound in Theorems 1.1 and 1.2, that may be rewritten as

$$\rho_r(f) := \frac{f_{\Delta(n,r)} - \underline{f}}{\overline{f} - \underline{f}} = O\left(\frac{1}{r}\right). \tag{1.4}$$

In [8] several examples are given where this error is in fact of the order $O(1/r^2)$ and the question is posed whether this could be true in general.

Here, we give an affirmative answer for quadratic polynomials. More precisely, we show that $\rho_r(f) \leq m/r^2$ if f has a global minimizer with denominator m (see Theorem 2.2). In view of Vavasis's result [18] on the existence of rational minimizers for quadratic programming, this implies that $\rho_r(f) = O(1/r^2)$ for quadratic f. For polynomials f of degree $d \geq 3$, when f admits a rational global minimizer, we show that $\rho_r(f) = O(1/r^2)$ (see Corollaries 3.3 and 4.5).

The main idea of our proof is to replace the multinomial distribution above by the *hypergeometric distribution*, and we therefore review some necessary background on the hypergeometric distribution next.

Multivariate hypergeometric distribution. Consider a box containing m balls, of which m_i are of color i for $i = 1, \dots, n$. Thus $\sum_{i=1}^n m_i = m$. We draw r balls randomly from the box without replacement. This defines the random variable Y_i as the number of balls of color i in a random sample of r balls. Then, $Y = (Y_1, \dots, Y_n)$ has the *multivariate hypergeometric distribution*, with parameters m, r, and n. Given $\alpha \in \mathbb{N}^n$ with $\sum_{i=1}^n \alpha_i = r$, the probability of obtaining the outcome α, with α_i balls of color i, is equal to

$$\mathbf{Pr}\left[Y_1 = \alpha_1, \dots, Y_n = \alpha_n\right] = \frac{\prod_{i=1}^n \binom{m_i}{\alpha_i}}{\binom{m}{r}}. \tag{1.5}$$

Note that if $r = m$, then there is only one possible outcome, since all the balls are drawn from the box. For $\beta \in \mathbb{N}^n$, the βth moment of the multivariate hypergeometric distribution Y is defined as

$$m^{\beta}_{(n,r)}(Y) := \mathbb{E}\left(\prod_{i=1}^{n} Y_i^{\beta_i}\right) = \sum_{\alpha \in I(n,r)} \alpha^{\beta} \frac{\prod_{i=1}^{n} \binom{m_i}{\alpha_i}}{\binom{m}{r}},$$

where $I(n,r) := \{\alpha \in \mathbb{N}^n : |\alpha| := \sum_{i=1}^{n} \alpha_i = r\}$. Combining [9, relation (34.18)] and [9, relation (39.6)], we can obtain the explicit formula for $m^{\beta}_{(n,r)}(Y)$ in terms of the Stirling numbers of the second kind. For integers $a, b \in \mathbb{N}$, the *Stirling number of the second kind* $S(a,b)$ counts the number of ways of partitioning a set of a objects into b nonempty subsets. Note that $S(a,b) = 0$ if $b > a$, and define the base cases $S(a,0) = 0$ if $a > 0$, and $S(0,0) = 1$.

Moreover, we will denote $r^{\underline{d}} := r(r-1)\cdots(r-d+1)$, with the conventions that $r^{\underline{d}} = 0$ if $r < d$ and $r^{\underline{0}} = 1$.

THEOREM 1.3. *For $\beta \in \mathbb{N}^n$, one has*

$$m^{\beta}_{(n,r)}(Y) = \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i).$$

Define the random variables

$$X = (X_1, \ldots, X_n), \quad \text{where } X_i := Y_i/r \quad (i = 1, \ldots, n). \tag{1.6}$$

Thus X takes its values in $\Delta(n,r)$. Theorem 1.3 gives the explicit formula for the moments of X.

COROLLARY 1.4. *For $\beta \in \mathbb{N}^n$, one has*

$$m^{\beta}_{(n,r)}(X) := \mathbb{E}\left(\prod_{i=1}^{n} X_i^{\beta_i}\right) = \frac{1}{r^{|\beta|}} \sum_{\alpha \in \mathbb{N}^n : \alpha \le \beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^{n} m_i^{\underline{\alpha_i}} S(\beta_i, \alpha_i).$$

The multivariate hypergeometric distribution can be used for upper bounding the minimum of f over $\Delta(n,r)$.

LEMMA 1.5. *Let $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$ and let $X := (X_1, X_2, \ldots, X_n)$ be as in* (1.5) *and* (1.6). *Then, one has*

$$f_{\Delta(n,r)} \le \mathbb{E}\left(f(X)\right),$$

and the above inequality can be strict.

Proof. By definition (1.6), the random variable X takes its values in $\Delta(n,r)$, which implies directly that the expected value of $f(X)$ is at least the minimum of f over $\Delta(n,r)$. In order to show the inequality can be strict, we consider the following example: $f = 2x_1^2 + x_2^2 - 5x_1x_2$. One has $\underline{f} = -\frac{17}{32}$ attained at the unique minimizer $(\frac{7}{16}, \frac{9}{16})$. Then we let $m = 16$, $m_1 = 7$ and $m_2 = 9$. When $r = 2$, one can easily check that $f_{\Delta(2,2)} = -\frac{1}{2}$ (attained at the unique minimizer $(\frac{1}{2}, \frac{1}{2})$). On the other hand, when $r = 2$, $\mathbb{E}\left(f(X)\right) = \frac{31}{80}$, and thus $\mathbb{E}\left(f(X)\right) > f_{\Delta(2,2)}$. ∎

To motivate the choice of the multivariate hypergeometric distribution over the multinomial distribution, consider the case where f has a rational minimizer $x^* \in \Delta(n,m)$, i.e., each component of x^* has denominator m.

If we now define the random variable X as in (1.5) and (1.6) with $m_i = mx_i^*$ ($i \in [n]$), and $r \leq m$, then

$$\mathbb{E}\left(f(X)\right) = \sum_{\alpha \in r\Delta(n,r)} \prod_{i=1}^{n} \frac{1}{\binom{m}{r}} \binom{mx_i^*}{\alpha_i} f\left(\frac{\alpha}{r}\right) =: H_r(f)(x^*).$$

Note that $H_r(f)(x^*)$ is the analogue of the Bernstein approximation $B_r(f)(x^*)$ in (1.3).

If $r = m$, then the only possible value that X can take is x^*. In other words, $H_m(f)(x^*) = f(x^*) = \underline{f}$, which means *finite convergence* of $H_r(f)(x^*)$ ($r = 1, 2, \ldots$) to $\underline{f}$, whereas the convergence $\lim_{r\to\infty} B_r(f)(x^*) = \underline{f}$ is not finite in general.

Bernstein coefficients. Any polynomial $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$ can be written as

$$f = \sum_{\beta \in I(n,d)} f_\beta x^\beta = \sum_{\beta \in I(n,d)} \left(f_\beta \frac{\beta!}{d!}\right) \frac{d!}{\beta!} x^\beta. \tag{1.7}$$

Then, the scalars $f_\beta \frac{\beta!}{d!}$ (for $\beta \in I(n,d)$) are called the *Bernstein coefficients* of f since they are the coefficients of f when expressing f in the Bernstein basis $\{\frac{d!}{\beta!}x^\beta : \beta \in I(n,d)\}$ of $\mathcal{H}_{n,d}$ (see, e.g., [6, 8, 16]). Combining (1.7) with the multinomial theorem

$$\left(\sum_{i=1}^{n} x_i\right)^d = \sum_{\alpha \in I(n,d)} \frac{d!}{\alpha!} x^\alpha, \tag{1.8}$$

it follows that, when $x \in \Delta_n$, $f(x)$ is a convex combination of its Bernstein coefficients $f_\beta \frac{\beta!}{d!}$. Hence, for any $x \in \Delta_n$, we have

$$\min_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!} \leq f(x) \leq \max_{\beta \in I(n,d)} f_\beta \frac{\beta!}{d!}. \tag{1.9}$$

In section 4, we will make use of the following theorem by De Klerk, Laurent, and Parrilo [7], which bounds the range of the Bernstein coefficients in terms of the range of function values $\overline{f} - \underline{f}$.

THEOREM 1.6 (see [7, Theorem 2.2]). *For any polynomial $f = \sum_{\beta \in I(n,d)} f_\beta x^\beta \in \mathcal{H}_{n,d}$, one has*

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

Notation. We denote $[n] := \{1, 2, \ldots, n\}$ and let $\mathbb{N}^n$ be the set of all n-dimensional nonnegative integral vectors. For $\alpha \in \mathbb{N}^n$, we define $|\alpha| := \sum_{i=1}^n \alpha_i$ and $\alpha! := \alpha_1!\alpha_2! \cdots \alpha_n!$. For vectors $\alpha, \beta \in \mathbb{N}^n$, the inequality $\alpha \leq \beta$ means $\alpha_i \leq \beta_i$ for any $i \in [n]$. As before, set $I(n,d) := \{\alpha \in \mathbb{N}^n : |\alpha| = d\}$ and let $\mathcal{H}_{n,d}$ be the set of all multivariate real homogeneous polynomials in n variables with degree d. Then, for $\alpha \in \mathbb{N}^n$, we denote $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$. Similarly, for $I \subseteq [n]$, we let $x^I := \prod_{i \in I} x_i$. A monomial x^α is called *square-free* (aka *multilinear*) if $\alpha_i \in \{0, 1\}$ ($i \in [n]$), and a polynomial f is called *square-free* if all its monomials are square-free. Moreover, 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 $\alpha \in \mathbb{N}^n$, with the conventions that $x^{\underline{d}} = 0$ if $x < d$ and $x^{\underline{0}} = 1$. We let e denote the all-ones vector and e_i denote the ith standard unit vector. Furthermore, for a random variable W, $\mathbb{E}(W)$ is its expectation.

Structure. The rest of the paper is organized as follows. In section 2, we consider the standard quadratic optimization problem, while in section 3 we treat the cubic and square-free (or *multilinear*) cases. In section 4, we focus on the general fixed-degree polynomial optimization over the simplex. Finally, we give all the proofs of results stated in section 3 in the appendix.

2. Standard quadratic optimization.

We consider the problem (1.1) where the polynomial f is assumed to be quadratic. The following result plays a key role for our refined error analysis in Theorem 2.2 below.

THEOREM 2.1. *Let $f = x^TQx \in \mathcal{H}_{n,2}$. For any integers r and $m \geq 2$ such that $1 \leq r \leq m$, one has*

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

Proof. Let $m \geq 2$ and let $x^* \in \Delta(n,m)$ be a minimizer of f over $\Delta(n,m)$, i.e., $f(x^*) = f_{\Delta(n,m)}$, and set $m_i = mx_i^*$ for $i \in [n]$. Consider the random variable $X = (X_1, \ldots, X_n)$ defined as in (1.5) and (1.6). By Corollary 1.4, one has

$$\begin{aligned}
\mathbb{E}[X_i^2] &= \left(\frac{m_i}{m}\right)^2 \left(1 - \frac{m-r}{r(m-1)} + \frac{m(m-r)}{rm_i(m-1)}\right) \quad (i \in [n]),\\
\mathbb{E}[X_iX_j] &= \frac{m_i}{m}\frac{m_j}{m}\left(1 - \frac{m-r}{r(m-1)}\right) \quad (i \neq j \in [n]).
\end{aligned}$$

Then, we have

$$\begin{aligned}
\mathbb{E}\left[f(X)\right] &= \sum_{i,j\in[n]:i\neq j} Q_{ij}\mathbb{E}[X_iX_j] + \sum_{i=1}^n Q_{ii}\mathbb{E}[X_i^2]\\
&= \sum_{i,j\in[n]:i\neq j} Q_{ij}\frac{m_i}{m}\frac{m_j}{m}\left(1 - \frac{m-r}{r(m-1)}\right)\\
&\quad + \sum_{i=1}^n Q_{ii}\left(\frac{m_i}{m}\right)^2\left(1 - \frac{m-r}{r(m-1)} + \frac{m(m-r)}{rm_i(m-1)}\right)\\
&= \sum_{i,j\in[n]} Q_{ij}x_i^*x_j^*\left(1 - \frac{m-r}{r(m-1)}\right) + \frac{m-r}{r(m-1)}\sum_{i=1}^n Q_{ii}x_i^*\\
&\leq f(x^*) - \frac{m-r}{r(m-1)}\underline{f} + \frac{m-r}{r(m-1)}\max_{i\in[n]} Q_{ii}\\
&\leq f(x^*) - \frac{m-r}{r(m-1)}\underline{f} + \frac{m-r}{r(m-1)}\overline{f}.
\end{aligned}$$

Hence, we obtain

$$\mathbb{E}\left[f(X)\right] - f_{\Delta(n,m)} = \mathbb{E}\left[f(X)\right] - f(x^*) \leq \frac{m-r}{r(m-1)}(\overline{f} - \underline{f}).$$

Using Lemma 1.5, we can conclude the proof. ∎

When f is quadratic, Vavasis [18] shows that there always exists a rational global minimizer x^* for problem (1.1). Say x^* has denominator m, i.e., $x^* \in \Delta(n,m)$. Our next result gives an upper bound for the error estimate $f_{\Delta(n,r)} - \underline{f}$ in terms of this denominator m.

THEOREM 2.2. *Let $f = x^T Q x \in \mathcal{H}_{n,2}$, and let x^* be a global minimizer of f over Δ_n, with denominator m. For any integer $r \geq 1$, one has*

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

Before proceeding with the proof, we note that one may give an upper bound on m in terms of Q, if $Q \in \mathbb{Z}^{n\times n}$. To this end, let $\bar{q} = \max_{ij}|Q_{ij}|$, and assume $x^* \in \Delta_n$ is a minimizer of f with the largest number (say, ℓ) of zero entries of all minimizers. Then one may show that $x^* \in \Delta(m,n)$, where the denominator m is bounded by

$$m \leq (4\bar{q})^{n-\ell-1}.$$

The proof uses the same argument as in Vavasis [18] and is omitted here. We only state this bound to make clear that the best-known upper bounds on m are exponential in n in general. This means that Theorem 2.2 does not yield a PTAS for standard quadratic optimization, but our interest here is in the dependence of the error bound on the parameter r.

The proof of Theorem 2.2 uses the following easy fact (whose proof is omitted).

LEMMA 2.3. *Let $r, k, m \geq 1$ be integers such that $(k-1)m < r \leq km$. Then,*

$$\frac{km - r}{km - 1} \leq \frac{m}{r}.$$

Proof of Theorem 2.2. Let $k \geq 1$ be an integer such that $(k-1)m < r \leq km$. We apply Theorem 2.1 to r and km (instead of m) and obtain that

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

Now, observe that $f_{\Delta(n,km)} = f_{\Delta(n,m)} = \underline{f}$, since $x^* \in \Delta(n,m) \subseteq \Delta(n,km) \subseteq \Delta_n$, and use the inequality from Lemma 2.3. □

As a direct application of Theorem 2.2, we see that the rate of convergence of the sequence $\rho_r(f)$ in (1.4) is in the order $O(1/r^2)$, where the constant depends only on the denominator of a rational global minimizer.

COROLLARY 2.4. *For any quadratic polynomial $f \in \mathcal{H}_{n,2}$, $\rho_r(f) = O(1/r^2)$.*

Moreover, the results of Theorems 2.1 and 2.2 refine the known error estimate from Theorem 1.1, which shows that $\rho_r(f) \leq \frac{1}{r}$. To see it, use Theorem 2.1 and the fact that $\frac{m-r}{r(m-1)} \leq \frac{1}{r}$ if $1 \leq r \leq m$, and use Theorem 2.2 and the inequality $\frac{m}{r^2} \leq \frac{1}{r}$ in the case $r \geq m$.

The following example shows that the inequality in Theorem 2.1 can be tight.

Example 2.5 (see [8, Example 2]). Consider the quadratic polynomial $f = \sum_{i=1}^n x_i^2$. Since f is convex, one can easily check $\overline{f} = 1$ (attained at any standard unit vector) and $\underline{f} = \frac{1}{n}$ (attained at $x = \frac{1}{n}e$, with denominator $m = n$). Moreover, for any integer $r \leq n$, we have $f_{\Delta(n,r)} = \frac{1}{r}$. Thus, we have

$$f_{\Delta(n,r)} - \underline{f} = \frac{n-r}{r(n-1)}(\overline{f} - \underline{f}) = \frac{m-r}{r(m-1)}(\overline{f} - \underline{f}).$$

Hence, for this example, the result in Theorem 2.1 is tight, while the result in Theorem 1.1 is not tight.

3. Cubic and square-free polynomial optimizations over the simplex. For the minimization of cubic and square-free polynomials over the standard simplex, the following results from [8] refine Theorem 1.2.

THEOREM 3.1.

(i) [8, Corollary 2] *For any polynomial $f \in \mathcal{H}_{n,3}$ and $r \geq 2$, one has*

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

(ii) [8, Corollary 3] *For any square-free polynomial $f \in \mathcal{H}_{n,d}$ and $r \geq 1$, one has*

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

We can show the following analogue of Theorem 2.1 for cubic and square-free polynomials. We delay the proof to Appendix A, since the details are similar to the quadratic case (but more technical).

THEOREM 3.2.

(i) *Let $f \in \mathcal{H}_{n,3}$. Given integers r, m satisfying $1 \leq r \leq m$ and $m \geq 3$, one has*

$$f_{\Delta(n,r)} - f_{\Delta(n,m)} \leq \frac{(m-r)(4mr - 2m - 2r)}{r^2(m-1)(m-2)}\left(\overline{f} - \underline{f}\right).$$

(ii) *Let $f \in \mathcal{H}_{n,d}$ be a square-free polynomial. Given integers r, m satisfying $1 \leq r \leq m$ and $m \geq d$, one has*

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

When problem (1.1) admits a rational global minimizer, then one can show that Theorem 3.2(ii) implies Theorem 3.1(ii) and that Theorem 3.2(i) implies Theorem 3.1(i) for $r \geq 1 + \frac{m-1}{\sqrt{2m-1}}$. We give the proofs for these statements in Appendix B.

Theorem 3.1 shows that the ratio $\rho_r(f)$ is in the order $O(1/r)$. As an application of Theorem 3.2, we can show that the ratio $\rho_r(f)$ is in the order $O(1/r^2)$ for cubic polynomials admitting a rational global minimizer over the simplex (see Corollary 3.3, whose proof is given in Appendix C). The same holds for square-free polynomials, as we will see in the next section.

COROLLARY 3.3. *Let $f \in \mathcal{H}_{n,3}$ and assume that f has a rational global minimizer in Δ_n. Then, $\rho_r(f) = O(1/r^2)$.*

4. General fixed-degree polynomial optimization over the simplex. In this section, we study the general fixed-degree polynomial optimization problem over the standard simplex. We first upper bound the range $f_{\Delta(n,r)} - f_{\Delta(n,m)}$ in terms of $\overline{f} - \underline{f}$.

THEOREM 4.1. *Let $f \in \mathcal{H}_{n,d}$. For any integers r, m satisfying $1 \leq r \leq m$ and $m \geq d$, one has*

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

Note that when f is square-free, we have proved a better bound in Theorem 3.2 (ii).

For the proof of Theorem 4.1, we will use the Vandermonde–Chu identity

$$\left(\sum_{i=1}^n x_i\right)^{\underline{d}} = \sum_{\alpha\in I(n,d)} \frac{d!}{\alpha!} x^{\underline{\alpha}} \quad \forall x \in \mathbb{R}^n \tag{4.1}$$

(see [14]), as well as the multinomial theorem (1.8). We will also need the following two lemmas about the Stirling numbers of the second kind.

LEMMA 4.2 (e.g., [8, Lemma 3]). *For any positive integer d and $r \geq 1$, one has*

$$\sum_{k=1}^{d-1} r^{\underline{k}} S(d,k) = r^d - r^{\underline{d}}.$$

LEMMA 4.3 (e.g., [8, Lemma 4]). *Given $\alpha \in I(n,k)$ and $d > k$, one has*

$$S(d,k) = \frac{\alpha!}{k!} \sum_{\beta\in I(n,d)} \frac{d!}{\beta!} \prod_{i=1}^n S(\beta_i, \alpha_i).$$

Furthermore, we will use the following technical result.

LEMMA 4.4. *Given $\beta \in I(n,d)$, for any integers r, m with $1 \leq r \leq m$, $m \geq d$ and integers m_i ($i \in [n]$) with $\sum_{i=1}^n m_i = m$, one has*

$$A_\beta := r^{\underline{d}} \left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i}\right) + \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta,\alpha\neq\beta} \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i) \geq 0, \tag{4.2}$$

$$\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} A_\beta = r^d m^{\underline{d}} - r^{\underline{d}} m^d. \tag{4.3}$$

Proof. We first prove (4.2). For any $\alpha \in \mathbb{N}^n$ with $|\alpha| \leq d$, one can easily check that $\frac{r^{\underline{|\alpha|}}}{r^{\underline{d}}} \geq \frac{m^{\underline{|\alpha|}}}{m^{\underline{d}}}$, that is, $r^{\underline{d}} \leq \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}}$. Hence, one has

$$\begin{aligned} A_\beta &= r^{\underline{d}} \left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i}\right) + \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta,\alpha\neq\beta} \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i) \\ &\geq r^{\underline{d}} \underbrace{\left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i} + \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta,\alpha\neq\beta} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i)\right)}_{:=B_\beta} = r^{\underline{d}} B_\beta. \end{aligned}$$

Then we consider the quantity B_β and show that $B_\beta = 0$. As $S(\beta_i, \beta_i) = 1$, one can rewrite B_β as

$$B_\beta = \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i) - \prod_{i=1}^n m_i^{\beta_i}.$$

Applying Lemma 4.2 (with (m_i, β_i) in place of (r, d)), we have

$$m_i^{\beta_i} = \sum_{\alpha_i=0}^{\beta_i} m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i), \quad \text{implying} \quad \prod_{i=1}^n m_i^{\beta_i} = \sum_{\alpha\in\mathbb{N}^n:\alpha\leq\beta} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i),$$

which shows that $B_\beta = 0$, and thus $A_\beta \geq 0$, which concludes the proof of (4.2).

We now show (4.3). By the definition (4.2), one has

$$\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} A_\beta = \underbrace{\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} r^{\underline{d}} \left(\prod_{i=1}^n m_i^{\underline{\beta_i}} - \prod_{i=1}^n m_i^{\beta_i}\right)}_{:=C_1} + \underbrace{\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} \sum_{\alpha\in\mathbb{N}^n:\alpha\le\beta,\alpha\neq\beta} \frac{r^{\underline{|\alpha|}} m^{\underline{d}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i)}_{:=C_2}.$$

On the one hand, using the Vandermonde–Chu identity (4.1), the multinomial theorem (1.8), and the identity $\sum_{i=1}^n m_i = m$, we find

$$C_1 = r^{\underline{d}}(m^{\underline{d}} - m^d).$$

On the other hand, we may exchange the summations in the definition of C_2 by recalling that $S(\beta_i,\alpha_i) = 0$ if $\alpha_i > \beta_i$ and noting that $\alpha \le \beta$, $\alpha \neq \beta$, and $\beta \in I(n,d)$ imply that $\alpha \in I(n,k)$ for some $k < d$. This allows us to remove the conditions $\alpha \le \beta$ and $\alpha \neq \beta$ in the summation, and we obtain

$$\begin{aligned} C_2 &= m^{\underline{d}} \sum_{k=1}^{d-1} \sum_{\alpha\in I(n,k)} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} \left(\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} \prod_{i=1}^n S(\beta_i,\alpha_i)\right) \\ &= m^{\underline{d}} \sum_{k=1}^{d-1} \frac{r^{\underline{k}}}{m^{\underline{k}}} S(d,k) \left(\sum_{\alpha\in I(n,k)} \frac{k!}{\alpha!} \prod_{i=1}^n m_i^{\underline{\alpha_i}} \right) && \text{(using Lemma 4.3)} \\ &= m^{\underline{d}} \sum_{k=1}^{d-1} r^{\underline{k}} S(d,k) && \text{(using Vandermonde–Chu identity (4.1))} \\ &= m^{\underline{d}} (r^d - r^{\underline{d}}) && \text{(using Lemma 4.2)} \end{aligned}$$

We can now conclude that $\sum_{\beta\in I(n,d)} \frac{d!}{\beta!} A_\beta = C_1 + C_2 = r^d m^{\underline{d}} - r^{\underline{d}} m^d$. ∎

Now we are ready to prove Theorem 4.1.

Proof of Theorem 4.1. Let $x^* \in \Delta(n,m)$ be a minimizer of f over $\Delta(n,m)$, i.e., $f(x^*) = f_{\Delta(n,m)}$. Set $m_i = mx_i^*$ for $i \in [n]$. Let the random variables X_i be defined as in (1.5) and (1.6), so that the random variable $X = (X_1, X_2, \ldots, X_n)$ takes its values in $\Delta(n,r)$. By Corollary 1.4 we have, for $\beta \in I(n,d)$,

$$\mathbb{E}[X^\beta] = \frac{1}{r^d} \sum_{\alpha\in\mathbb{N}^n:\alpha\le\beta} \frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}} \prod_{i=1}^n m_i^{\underline{\alpha_i}} S(\beta_i,\alpha_i).$$

Then, as $S(\beta_i,\beta_i)=1$, we can rewrite

$$\begin{aligned}\mathbb{E}[X^\beta] &= \frac{1}{r^d}\frac{r^{\underline{d}}}{m^{\underline{d}}}\prod_{i=1}^n m_i^{\underline{\beta_i}} + \underbrace{\frac{1}{r^d}\sum_{\alpha\in\mathbb{N}^n:\alpha\le\beta,\alpha\neq\beta}\frac{r^{\underline{|\alpha|}}}{m^{\underline{|\alpha|}}}\prod_{i=1}^n m_i^{\underline{\alpha_i}}S(\beta_i,\alpha_i)}_{:=D_\beta}\\
&= \prod_{i=1}^n\left(\frac{m_i}{m}\right)^{\beta_i}\left[\frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}+\frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\left(\prod_{i=1}^n\frac{m_i^{\underline{\beta_i}}}{m_i^{\beta_i}}-1\right)\right]+D_\beta\\
&= \underbrace{\prod_{i=1}^n\left(\frac{m_i}{m}\right)^{\beta_i}\frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}}_{:=T_1}+\underbrace{\frac{r^{\underline{d}}}{r^d m^{\underline{d}}}\left(\prod_{i=1}^n m_i^{\underline{\beta_i}}-\prod_{i=1}^n m_i^{\beta_i}\right)+D_\beta}_{:=T_2}\\
&= T_1+T_2=(x^*)^\beta\frac{r^{\underline{d}}m^d}{r^d m^{\underline{d}}}+\frac{A_\beta}{r^d m^{\underline{d}}}.\end{aligned}$$

For the above last equality, note that, since $x_i^*=\frac{m_i}{m}$, one has

$$T_1=(x^*)^\beta\frac{r^{\underline{d}}m^d}{r^d m^{\underline{d}}}$$

and using the definition of A_β in (4.2), we can write

$$T_2=\frac{A_\beta}{r^d m^{\underline{d}}}.$$

Thus we obtain

$$\begin{aligned}\mathbb{E}[f(X)] &= \mathbb{E}\left[\sum_{\beta\in I(n,d)}f_\beta X^\beta\right]=\sum_{\beta\in I(n,d)}f_\beta\mathbb{E}[X^\beta]\\
&= \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}f(x^*)+\frac{1}{r^d m^{\underline{d}}}\sum_{\beta\in I(n,d)}f_\beta A_\beta.\end{aligned}$$

Therefore, we have

$$r^d m^{\underline{d}}(\mathbb{E}[f(X)]-f(x^*))=(r^{\underline{d}}m^d-r^d m^{\underline{d}})f(x^*)+\sum_{\beta\in I(n,d)}f_\beta A_\beta.$$

We now upper bound the two terms $(r^{\underline{d}}m^d-r^d m^{\underline{d}})f(x^*)$ and $\sum_{\beta\in I(n,d)}f_\beta A_\beta$.

First, since $r^{\underline{d}}m^d-r^d m^{\underline{d}}<0$ and $f(x^*)\ge\min_{\beta\in I(n,d)}f_\beta$ (see (1.9)), one obtains

$$(r^{\underline{d}}m^d-r^d m^{\underline{d}})f(x^*)\le(r^{\underline{d}}m^d-r^d m^{\underline{d}})\left(\min_{\beta\in I(n,d)}f_\beta\frac{\beta!}{d!}\right). \tag{4.4}$$

Second, using the fact that $A_\beta\ge 0$ (by Lemma 4.4), one obtains

$$\sum_{\beta\in I(n,d)}f_\beta A_\beta\le\left(\max_{\beta\in I(n,d)}f_\beta\frac{\beta!}{d!}\right)\sum_{\beta\in I(n,d)}\frac{d!}{\beta!}A_\beta.$$

Using the identity $\sum_{\beta\in I(n,d)}\frac{d!}{\beta!}A_\beta=r^d m^{\underline{d}}-r^{\underline{d}}m^d$ (see (4.3)), one can obtain

$$\sum_{\beta\in I(n,d)}f_\beta A_\beta\le\left(\max_{\beta\in I(n,d)}f_\beta\frac{\beta!}{d!}\right)(r^d m^{\underline{d}}-r^{\underline{d}}m^d).$$

Combining with (4.4), this implies

$$r^d m^{\underline{d}}(\mathbb{E}[f(X)] - f(x^*)) \le (r^d m^{\underline{d}} - r^{\underline{d}} m^d)\left(\max_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!} - \min_{\beta\in I(n,d)} f_\beta \frac{\beta!}{d!}\right).$$

Using Theorem 1.6, Lemma 1.5, and the fact that $f(x^*) = f_{\Delta(n,m)}$, we finally obtain

$$r^d m^{\underline{d}}(f_{\Delta(n,r)} - f_{\Delta(n,m)}) \le r^d m^{\underline{d}}(\mathbb{E}[f(X)] - f(x^*)) \le (r^d m^{\underline{d}} - r^{\underline{d}} m^d)\binom{2d-1}{d} d^d(\overline{f} - \underline{f}),$$

which concludes the proof of Theorem 4.1. ∎

In what follows we now assume that $f \in \mathcal{H}_{n,d}$ has a rational global minimizer x^* with denominator m, i.e., $x^* \in \Delta(n,m)$, so that $\underline{f} = f_{\Delta(n,m)}$.

First, observe that Theorem 4.1 refines the result from Theorem 1.2 (which follows from the fact that $1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} \le 1 - \frac{r^{\underline{d}}}{r^d}$ for any $k \ge 1$).

Next, we show as an application of Theorem 4.1 that the ratio $\rho_r(f)$ is in the order $O(1/r^2)$.

COROLLARY 4.5. *Let $f \in \mathcal{H}_{n,d}$ and assume that there exists a rational global minimizer for problem* 1.1. *Then, $\rho_r(f) = O(1/r^2)$.*

For the proof of Corollary 4.5, we need the following notation. Consider the univariate polynomial $(x-1)(x-2)\cdots(x-d+1)$ (in the variable x), which can be written as

$$\begin{aligned}(x-1)(x-2)\cdots(x-d+1) &= x^{d-1} - a_{d-2}x^{d-2} + a_{d-3}x^{d-3} + \cdots + (-1)^{d-1}a_0 \\ &= x^{d-1} + p(x),\end{aligned} \tag{4.5}$$

setting

$$p(x) = \sum_{i=0}^{d-2} (-1)^{d-1-i} a_i x^i, \tag{4.6}$$

where a_i are positive integers depending only on d for any $i \in \{0,1,\dots,d-2\}$. We also need the following lemma.

LEMMA 4.6. *Let r, m, and k be integers satisfying $m \ge d$, $k \ge 1$, and $(k-1)m < r \le km$. Then one has*

$$1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} \le \frac{m}{r^2} c_d$$

for some constant c_d depending only on d.

Proof. Based on (4.6), one can write

$$1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} = \underbrace{\frac{(km)^{d-1}}{(km-1)(km-2)\cdots(km-d+1)}}_{:=\sigma_0(r,km)} \underbrace{\left[\frac{p(km)}{(km)^{d-1}} - \frac{p(r)}{r^{d-1}}\right]}_{:=\sigma_1(r,km)}.$$

First we consider the term $\sigma_0(r,km)$. For any integer $i \in \{1,\dots,d-1\}$, as $k \ge 1$ and $m \ge d$, we have that $km(d-1) \ge id$, which implies $\frac{km}{km-i} \le d$. Hence, one has $\sigma_0(r,km) \le d^{d-1}$. Next we consider the term $\sigma_1(r,km)$. Recalling (4.5), we can write

$\sigma_1(r,km)$ as $\sigma_1(r,km) = \sum_{i=0}^{d-2}(-1)^{d-1-i}a_i\left(\frac{1}{(km)^{d-1-i}} - \frac{1}{r^{d-1-i}}\right)$. Since $r \le km$, then $\frac{1}{km} \le \frac{1}{r}$ and $\frac{1}{(km)^{d-1-i}} \le \frac{1}{r^{d-1-i}}$ for any $i \in \{0,1,\ldots,d-2\}$. This gives

$$\sigma_1(r,km) \le \sum_{i=0}^{d-2} a_i\left(\frac{1}{r^{d-1-i}} - \frac{1}{(km)^{d-1-i}}\right). \tag{4.7}$$

Then, we consider the term $\frac{1}{r^{d-1-i}} - \frac{1}{(km)^{d-1-i}}$ (for any $i \in \{0,1,\ldots,d-2\}$) in (4.7). For any integer $s \in [d-1]$, we have

$$\frac{1}{r^s} - \frac{1}{(km)^s} = \frac{(km)^s - r^s}{r^s(km)^s} = D_1 \cdot D_2,$$

setting

$$D_1 = \frac{km-r}{kmr},$$
$$D_2 = \frac{(km)^{s-1} + (km)^{s-2}r + \cdots + r^{s-1}}{r^{s-1}(km)^{s-1}}.$$

On the one hand, one has $D_1 \le \frac{km-r}{r(km-1)} \le \frac{m}{r^2}$, where the second inequality follows by Lemma 2.3. On the other hand, observe that for any $i,j \in \{0,1,\ldots,s-1\}$ with $i+j = s-1$, one has $(km)^i r^j \le (km)^{s-1}r^{s-1}$. Hence, $D_2 \le s \le d-1$. That is,

$$\frac{1}{r^s} - \frac{1}{(km)^s} \le \frac{m(d-1)}{r^2}.$$

Using this in (4.7), we find that $\sigma_1(r,km) \le \frac{m(d-1)}{r^2}\sum_{i=0}^{d-2} a_i$. From (4.5) and (4.6), we know that the term $(d-1)(\sum_{i=0}^{d-2} a_i)$ is a constant c_d that depends only on d. This concludes the proof. ∎

We can now prove Corollary 4.5.

Proof of Corollary 4.5. Let $x^* \in \Delta(n,m)$ be a rational global minimizer of f over Δ_n. Let $r \ge d$ and let $k \ge 1$ be an integer such that $(k-1)m < r \le km$. Using Theorem 4.1 (applied to r and km (instead of m)), we obtain that

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

Combining with Lemma 4.6, one can conclude. ∎

5. Concluding remarks. As explained in the introduction, the analysis presented here is essentially a modification of the analysis in [8], in the sense that one discrete distribution on $\Delta(n,r)$ is replaced by another.

Having said that, the analysis in the current paper does not imply the PTAS results in [8] for nonquadratic f, due to the restrictive assumption of a rational global minimizer. It is not clear at this time if this assumption is an artifact of our analysis using the hypergeometric distribution or if there exist examples of problem (1.1) where all global minimizers are irrational and $\rho_r(f) = \Omega(1/r)$. This remains as an interesting question for future research.

Appendix A. We give here the proof of Theorem 3.2. As in the proof of Theorem 2.1, let $x^* \in \Delta(n,m)$ be a minimizer of f over $\Delta(n,m)$, i.e., $f(x^*) = f_{\Delta(n,m)}$, and set $m_i = mx_i^*$ for $i \in [n]$. Consider the random variables X_i defined in (1.5) and (1.6), so that $X = (X_1, X_2, \ldots, X_n)$ takes its values in $\Delta(n,r)$.

First we consider the case (i) when f is a homogeneous polynomial of degree 3. Write f as

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

By Corollary 1.4, for any $i,j,k \in [n]$, one has

$$\begin{aligned}
\mathbb{E}[X_i^3] &= \left(\frac{m_i}{m}\right)^3 \left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)}\right.\\
&\qquad \left. + (m-r)\frac{3rm_i m^2 - 3m_i m^2 + m^3 - 2rm^2}{r^2 m_i^2 (m-1)(m-2)}\right],\\
\mathbb{E}[X_i^2 X_j] &= \left(\frac{m_i}{m}\right)^2 \frac{m_j}{m} \left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)}\right.\\
&\qquad \left. + (m-r)\frac{(r-1)m^2}{r^2 m_i (m-1)(m-2)}\right],\\
\mathbb{E}[X_i X_j X_k] &= \frac{m_i}{m}\frac{m_j}{m}\frac{m_k}{m} \left[1 - (m-r)\frac{3mr - 2(m+r)}{r^2(m-1)(m-2)}\right].
\end{aligned}$$

Therefore, one obtains

$$\begin{aligned}
\mathbb{E}\left[f(X)\right] &= \textstyle\sum_i f_i \mathbb{E}[X_i^3] + \sum_{i<j} \left(f_{ij}\mathbb{E}[X_i X_j^2] + g_{ij}\mathbb{E}[X_i^2 X_j]\right)\\
&\quad \textstyle + \sum_{i<j<k} f_{ijk}\mathbb{E}[X_i X_j X_k]\\
&= f(x^*)\left[1 - (m-r)\tfrac{3mr-2(m+r)}{r^2(m-1)(m-2)}\right] + \tfrac{m-r}{r^2(m-1)(m-2)}\sigma,
\end{aligned} \tag{A.1}$$

where we set

$$\sigma := \sum_{i=1}^n f_i \frac{m_i}{m}(3m_i r - 3m_i + m - 2r) + m(r-1)\sum_{i<j}(f_{ij} + g_{ij})\frac{m_i}{m}\frac{m_j}{m}. \tag{A.2}$$

As in [7], by evaluating f at e_i and $(e_i + e_j)/2$, we obtain, respectively, the relations

$$\underline{f} \le f_i \le \overline{f}, \tag{A.3}$$

$$f_i + f_j + f_{ij} + g_{ij} \le 8\overline{f}. \tag{A.4}$$

Using (A.4), we obtain

$$\sum_{i<j}(f_{ij} + g_{ij})\frac{m_i}{m}\frac{m_j}{m} \le \sum_{i<j}(8\overline{f} - f_i - f_j)\frac{m_i}{m}\frac{m_j}{m} = 8\overline{f}\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m} - \sum_{i=1}^n f_i \frac{m_i}{m}\left(1 - \frac{m_i}{m}\right).$$

We use this inequality together with (A.3) to upper bound the term σ from (A.2):

$$\begin{aligned}\sigma &\le \sum_{i=1}^n f_i \frac{m_i}{m}(4m_i r - 4m_i + 2m - 2r - mr) + 8m(r-1)\overline{f}\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}\\ &= 4m(r-1)\left(\sum_{i=1}^n f_i\left(\frac{m_i}{m}\right)^2 + 2\overline{f}\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}\right) + (2m-2r-mr)\sum_{i=1}^n f_i\frac{m_i}{m}\\ &\le 4m(r-1)\overline{f}\left(\sum_{i=1}^n\left(\frac{m_i}{m}\right)^2 + 2\sum_{i<j}\frac{m_i}{m}\frac{m_j}{m}\right) + 2(m-r)\sum_{i=1}^n f_i\frac{m_i}{m} - mr\sum_{i=1}^n f_i\frac{m_i}{m}\\ &\le 4m(r-1)\overline{f} + 2(m-r)\overline{f} - mr\underline{f} = (4mr - 2m - 2r)\overline{f} - mr\underline{f}.\end{aligned}$$

We can now upper bound the quantity $\mathbb{E}\left[f(X)\right]$ from (A.1) as follows:

$$\mathbb{E}\left[f(X)\right] \le f(x^*) + \frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)}(\overline{f} - \underline{f}).$$

Together with Lemma 1.5, this now concludes the proof of Theorem 3.2(i).

We now consider the case (ii) when f is a homogeneous square-free polynomial of degree d. Say $f = \sum_{I\subseteq[n],|I|=d} f_I x^I$. By Corollary 1.4, one has

$$\mathbb{E}[X^I] = \frac{r^{\underline{d}}}{r^d}\frac{\prod_{i\in I} m_i}{m^{\underline{d}}} = \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\prod_{i\in I}\frac{m_i}{m}$$

and thus

$$\mathbb{E}[f(X)] = \sum_{I\subseteq[n],|I|=d} f_I\mathbb{E}[X^I] = \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}f(x^*) = \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}f_{\Delta(n,m)}.$$

Therefore,

$$\begin{aligned}\mathbb{E}[f(X)] - f_{\Delta(n,m)} &= -\left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right)f_{\Delta(n,m)} \le -\left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right)\underline{f}\\ &\le \left(1 - \frac{r^{\underline{d}}}{r^d}\frac{m^d}{m^{\underline{d}}}\right)\left(\overline{f} - \underline{f}\right).\end{aligned}$$

Here, for the last inequality we have used the fact that $\overline{f} \ge 0$ (since $f(e_i) = 0$ for any $i \in [n]$). Together with Lemma 1.5, this concludes the proof of Theorem 3.2(ii).

Appendix B. Assume $f \in \mathcal{H}_{n,3}$ has a rational minimizer on Δ_n with denominator $m \ge 3$.

First we show how to derive Theorem 3.1(i) for $r \ge 1 + \frac{m-1}{\sqrt{2m-1}}$ from our result in Theorem 3.2(i).

When $1+\frac{m-1}{\sqrt{2m-1}} \le r \le m$, this follows directly from the fact that $\frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)} \le \frac{4}{r} - \frac{4}{r^2}$.

Assume now $r > m \ge 3$ and $(k-1)m < r \le km$ for some integer $k \ge 2$. It suffices to show the inequality $\frac{(km-r)(4kmr-2km-2r)}{r^2(km-1)(km-2)} \le \frac{4}{r} - \frac{4}{r^2}$ or, equivalently,

$$\varphi(r) := (2km-1)r^2 + (4-6km)r - k^2m^2 + 6km - 4 \ge 0.$$

One can check that the function $\varphi(r)$ is monotonically increasing for $r \geq 1+\frac{km-1}{2km-1}$ and thus for $r \geq 2$. Hence it suffices to show that $\varphi((k-1)m+1) \geq 0$. If $m \geq 3$ is fixed, then one can check that $\varphi((k-1)m+1)$, as a function of k, is monotonically increasing for $k \geq 2$. Therefore, it suffices to show that $\varphi((k-1)m+1) \geq 0$ when $k=2$ and $m \geq 3$. One can now check that $\varphi((k-1)m+1)$ with $k=2$, as a function of m, is monotonically increasing for $m \geq 3$. Finally, we can conclude that it suffices to show that $\varphi((k-1)m+1) \geq 0$ when $k=2$ and $m=3$, which can be easily checked to hold. Thus we have shown that $\varphi(r) \geq 0$ for any $r > m$.

To see that Theorem 3.2(ii) implies Theorem 3.1(ii), consider an integer $k \geq 1$ such that $(k-1)m < r \leq km$ and observe that $1 - \frac{r^{\underline{d}}(km)^d}{r^d(km)^{\underline{d}}} \leq 1 - \frac{r^{\underline{d}}}{r^d}$.

Appendix C. We prove Corollary 3.3. Assume $m \geq 3$ is the denominator for a rational global minimizer of f over Δ_n. If $1 \leq r \leq m$, then, by using Theorem 3.2(i), Lemma 2.1, and the inequality $\frac{4mr-2m-2r}{r(m-2)} \leq \frac{4(m-1)}{m-2}$, we deduce that

$$\rho_r(f) \leq \frac{(m-r)(4mr-2m-2r)}{r^2(m-1)(m-2)} \leq \frac{4m^2}{r^2(m-2)}.$$

Assume now $r > m$ and $(k-1)m < r \leq km$ for some integer $k \geq 2$. Then Theorem 3.2(i) implies

$$f_{\Delta(n,r)} - \underline{f} = f_{\Delta(n,r)} - f_{\Delta(n,km)} \leq \frac{(km-r)(4kmr-2km-2r)}{r^2(km-1)(km-2)} \left(\overline{f} - \underline{f}\right).$$

One can easily check that $\frac{4kmr-2km-2r}{r(km-2)} \leq 6$, which, together with Lemma 2.3, implies that $\rho_r(f) \leq \frac{6m}{r^2}$. This concludes the proof of Corollary 3.3.

REFERENCES

[1] I. M. Bomze, S. Gollowitzer, and E. A. Yildirim, *Rounding on the standard simplex: Regular grids for global optimization*, J. Global Optim., 59 (2014), pp. 243–258.

[2] I. M. Bomze and E. De Klerk, *Solving standard quadratic optimization problems via semidefinite and copositive programming*, J. Global Optim., 24 (2002), pp. 163–185.

[3] L. P. Bos, *Bounding the Lebesque function for Lagrange interpolation in a simplex*, J. Approx. Theory, 38 (1983), pp. 43–59.

[4] E. De Klerk, *The complexity of optimizing over a simplex, hypercube or sphere: A short survey*, Cent. Eur. J. Oper. Res., 16 (2008), pp. 111–125.

[5] E. De Klerk, D. Den Hertog, and G. Elabwabi, *On the complexity of optimization over the standard simplex*, European J. Oper. Res., 191 (2008), pp. 773–785.

[6] E. De Klerk and M. Laurent, *Error bounds for some semidefinite programming approaches to polynomial optimization on the hypercube*, SIAM J. Optim., 20 (2010), pp. 3104–3120.

[7] E. De Klerk, M. Laurent, and P. Parrilo, *A PTAS for the minimization of polynomials of fixed degree over the simplex*, Theoret. Comput. Sci., 361 (2006), pp. 210–225.

[8] E. De Klerk, M. Laurent, and Z. Sun, *An alternative proof of a PTAS for fixed-degree polynomial optimization over the simplex*, Math. Program., 151 (2015), pp. 433–457.

[9] N. L. Johnson, S. Kotz, and N. Balakrishnan, *Discrete Multivariate Distributions*, Wiley, New York, 1997.

[10] J.-B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim., 11 (2001), pp. 796–817.

[11] J.-B. Lasserre, *A new look at nonnegativity on closed sets and polynomial optimization*, SIAM J. Optim., 21 (2011), pp. 864–885.

[12] T. S. Motzkin and E. G. Straus, *Maxima for graphs and a new proof of a theorem of Túran*, Canad. J. Math., 17 (1965), pp. 533–540.

[13] Y. Nesterov, *Random Walk in a Simplex and Quadratic Optimization over Convex Polytopes*, Discussion paper 2003/71, CORE-UCL, Louvain-La-Neuve, 2003.

[14] V. Powers and B. Reznick, *A new bound for Pólya's theorem with applications to polynomials positive on polyhedra*, J. Pure Appl. Algebra, 164 (2001), pp. 221–229.
[15] G. Sagol and E. A. Yildirim, *Analysis of copositive optimization based linear programming bounds on standard quadratic optimization*, J. Global Optim. (2015), DOI: 10.1007/s10898-015-0269-4.
[16] Z. Sun, *A refined error analysis for fixed-degree polynomial optimization over the simplex*, J. Oper. Res. Soc. China., 2 (2014), pp. 379–393.
[17] E. A. Yildirim, *On the accuracy of uniform polyhedral approximations of the copositive cone*, Optim. Methods Softw., 27 (2012), pp. 155–173.
[18] S. A. Vavasis, *Quadratic programming is in NP*, Inform. Process. Lett., 36 (1990), pp. 73–77.