A note on total degree polynomial optimization by Chebyshev grids*

Federico Piazzon[1] and Marco Vianello[1]

June 13, 2017

Abstract

Using the approximation theory notions of polynomial mesh and Dubiner distance in a compact set, we derive error estimates for total degree polynomial optimization on Chebyshev grids of the hypercube.

2010 AMS subject classification: 41A17, 65K05, 90C26.

Keywords: multivariate polynomial optimization, polynomial meshes, Dubiner distance, hypercube, Chebyshev grids.

1 Introduction

In this note we apply, in the framework of polynomial optimization, a notion that has been playing an emerging role over the last decade in the theory of multivariate polynomial inequalities and multivariate polynomial approximation: the notion of *polynomial mesh* of a compact set $K \subset \mathbb{R}^d$.

In what follows we denote by $\mathbb{P}_n^d$ the subspace of d-variate real polynomials of total degree not exceeding n, and by $N = \dim(\mathbb{P}_n^d) = \binom{n+d}{d}$ its dimension. Moreover, $\|f\|_X$ will denote the sup-norm of a bounded real function on a discrete or continuous compact set $X \subset \mathbb{R}^d$.

We recall that a compact set $K \subset \mathbb{R}^d$ is termed $\mathbb{P}_n^d$-determining if the only polynomial in $\mathbb{P}_n^d$ vanishing on the whole K is the null polynomial. In such a case we have that $\dim(\mathbb{P}_n^d(K)) = \dim(\mathbb{P}_n^d)$. The compact set is termed *polynomial determining* if it is $\mathbb{P}_n^d$-determining for every n (a sufficient condition being, for example, that K has interior points). We can give now the following:

Definition 1 *A polynomial mesh on a polynomial determining compact set $K \subset \mathbb{R}^d$ is a sequence of finite norming subsets $\mathcal{A}_n \subset K$ such that the polynomial inequality*

$$\|p\|_K \leq c\,\|p\|_{\mathcal{A}_n}\ ,\ \forall p \in \mathbb{P}_n^d\ , \tag{1}$$

holds for some $c > 1$ independent of p and n, where card$(\mathcal{A}_n) = \mathcal{O}(n^s)$, $s \geq d$.

*Work partially supported by the DOR funds and the biennial projects CPDA143275 and BIRD163015 of the University of Padova, and by the GNCS-INdAM.

[1]Department of Mathematics, University of Padova, Italy
corresponding author: marcov@math.unipd.it

Observe that $\mathcal{A}_n$ is automatically $\mathbb{P}_n^d$-determining, and thus $\mathrm{card}(\mathcal{A}_n) \geq N = \dim(\mathbb{P}_n^d) \sim n^d/d!$ (d fixed, $n \to \infty$). A polynomial mesh is termed *optimal* when $s = d$. All these notions can be given for $K \subset \mathbb{C}^d$ but we restrict here to the real case.

The notion of polynomial mesh was introduced in the seminal paper [7] and then used from both the theoretical and the computational point of view. Indeed, polynomial meshes have interesting computational features, for example they: are affinely invariant and are stable under small perturbations; can be extended by algebraic transforms, finite union and product; contain computable near optimal interpolation sets of Fekete type (maximal Vandermonde determinant); are near optimal for uniform Least Squares approximation. Concerning the theory, computation and application of polynomial meshes we refer the reader, e.g., to [3, 6, 7, 18, 19, 20, 21] and the references therein.

In the sequel we shall use the fact that Chebyshev grids of suitable cardinality form a polynomial mesh of the hypercube. To this respect, it is worth recalling the notion of Dubiner distance in a multivariate compact set. Such a distance is defined for $x, y \in K$ as

$$\mathrm{dist}_D(x,y) = \sup_{deg(p)\geq 1,\, \|p\|_K \leq 1} \left\{ \frac{1}{deg(p)} |\cos^{-1}(p(x)) - \cos^{-1}(p(y))| \right\} . \quad (2)$$

Introduced by M. Dubiner in the seminal paper [14], it belongs to a family of three distances (the other two are the Markov distance and the Baran distance) that play an important role in multivariate polynomial approximation and have deep connections with multivariate polynomial inequalities. As far as we know, until now the Dubiner distance is known analytically only in very few instances: the interval (where it coincides with the usual distance $|\cos^{-1}(x) - \cos^{-1}(y)|$), more generally the hypercube, and the sphere, the simplex, and the ball. We may recall that the Dubiner distance in the *hypercube* is

$$\mathrm{dist}_D(x,y) = \max \left\{ |\cos^{-1}(x_1) - \cos^{-1}(y_1)|, \ldots, |\cos^{-1}(x_d) - \cos^{-1}(y_d)| \right\} , \quad (3)$$

whereas the Dubiner distance on the *sphere* coincides with the *geodesic distance*. We refer the readers, e.g., to [4, 5] and to the references therein for other properties and results.

A simple connection of the Dubiner distance with the theory of polynomial meshes is given by the following:

Proposition 1 *Let X be a compact subset of a compact set $K \subset \mathbb{R}^d$ whose covering radius $\rho(X)$ with respect to the Dubiner distance does not exceed θ/n, where $\theta \in (0, \pi/2)$ and $n \geq 1$, i.e.*

$$\rho(X) = \max_{x \in K} dist_D(x, X) = \max_{x \in K} \min_{y \in X} dist_D(x,y) \leq \frac{\theta}{n} . \quad (4)$$

Then, the following inequality holds

$$\|p\|_K \leq \frac{1}{\cos\theta} \|p\|_X , \quad \forall p \in \mathbb{P}_n^d . \quad (5)$$

Proof. Assume that $\|p\|_K = 1$ (otherwise we consider $p/\|p\|_K$), and let $\xi \in K$ be a point such that $|p(\xi)| = 1$. Possibly multiplying p by -1, we may assume that $p(\xi) = 1$. By the definition of Dubiner distance, there is $y \in X$ such that

$$|\cos^{-1}(p(\xi)) - \cos^{-1}(p(y))| = |\cos^{-1}(p(y))| \leq \frac{\theta\, deg(p)}{n} \leq \theta \,,$$

and hence

$$\cos^{-1}(p(y)) \leq \theta < \frac{\pi}{2} \,.$$

Then, since the inverse cosine function is monotonically decreasing, we get

$$p(y) \geq \cos(\theta) > 0$$

and finally

$$\|p\|_K = 1 \leq \frac{p(y)}{\cos\theta} \leq \frac{1}{\cos\theta}\, \|p\|_X \,. \quad \square$$

Remark 1 In the one-dimensional case with $K = [-1, 1]$, taking as X the set of ℓ Chebyshev points (the zeros of $T_\ell(x) = \cos(\ell \cos^{-1}(x))$, $\ell > n$), (5) is a well-known inequality obtained by Ehlich and Zeller in 1964, where $\theta = n\pi/(2\ell)$; see [15] and [6], where the case of $\ell+1$ Chebyshev-Lobatto points is also considered.

Observe however that X need not to be discrete. For example, in [2] it has been proved via the Dubiner distance that suitable Lissajous curves are norming sets for $\mathbb{P}_n^d$ in the hypercube. On the other hand, Proposition 1 gives a tool to recognize when discrete subsets form a polynomial mesh.

2 Optimization on Chebyshev grids

Polynomial optimization on suitable grids is a well-known technique, that has been developed for example by uniform rational sampling in the simplex and the hypercube; see, e.g., [10, 12, 13] and the references therein. On the other hand, polynomial optimization on Chebyshev grids seems to have been studied essentially only via tensor-product polynomial spaces, see [16, 23] and also [22], where it is used by the functions `min2` and `max2` within the Matlab package Chebfun2 (square) (and more recently Chebfun3 for the cube, see [17]).

More precisely, in the present implementation of Chebfun2 given a bivariate "chebfun" of degree n_1 in the first and n_2 in the second variable (roughly, "chebfuns" are polynomials representing smooth functions up to machine precision), when this is not of rank 1 (i.e., the product of two univariate polynomials), it is evaluated at a Chebyshev $n_1 \times n_2$ grid, and the discrete optimum used as a starting guess for a superlinearly convergent constrained trust region method, based on [8]; see [22] for a more detailed discussion.

Here we exploit the general fact that if in (1) we can let $c \to 1$ (for fixed n), then we get a kind of "brute force" method for total degree polynomial optimization by polynomial meshes. We focus on the case of the hypercube $K = [-1, 1]^d$, where Chebyshev grids of suitable cardinality turn out to be polynomial meshes with $c \to 1$. This is summarized in the following:

Lemma 1 *Let $K = [-1, 1]^d$, $m, n \in \mathbb{N}$, $m \geq 2$, and $X_{mn} = (\mathcal{C}_{mn})^d$, where $\mathcal{C}_s = \{\cos(j\pi/s)\,,\, 0 \leq j \leq s\}$ denote the univariate Chebyshev-Lobatto points of degree s. Then for every $p \in \mathbb{P}_n^d$*

$$\|p\|_K - \|p\|_{X_{mn}} \leq \varepsilon_m\, \|p\|_{X_{mn}} \leq \varepsilon_m\, \|p\|_K \,,$$

$$\varepsilon_m = \frac{1-\cos\left(\frac{\pi}{2m}\right)}{\cos\left(\frac{\pi}{2m}\right)} \leq \frac{\sigma_m}{1-\sigma_m} \;, \tag{6}$$

where $\sigma_m = \frac{\pi^2}{8m^2} \approx \frac{1.23}{m^2}$.

Proof. Consider the Dubiner distance (3) in the hypercube. The covering radius $\rho(X_{mn})$ in such a distance, is clearly half of the separation distance that is $\min_{y,z\in X_{mn}} \text{dist}_D(y,z) = \pi/(mn) = 2\rho(X_{mn})$.

Then in Proposition 1 we have $\theta = n\rho(X_{mn}) = \pi/(2m)$ and

$$\|p\|_K \leq \frac{1}{\cos\left(\frac{\pi}{2m}\right)} \|p\|_{X_{mn}} \;, \tag{7}$$

from which (6) immediately follows by subtracting $\|p\|_{X_{mn}}$ on both sides, since $1-\cos\phi \leq \phi^2/2$ fo every $\phi \in \mathbb{R}$ and hence

$$\frac{1}{\cos\left(\frac{\pi}{2m}\right)} - 1 = \frac{1-\cos\left(\frac{\pi}{2m}\right)}{\cos\left(\frac{\pi}{2m}\right)} \leq \frac{\sigma_m}{1-\sigma_m} \;. \quad \square$$

Remark 2 We recall that the notion of polynomial mesh is affinely invariant. Thus, Lemma 1 holds true in any box $K = [a_1,b_1]\times\cdots\times[a_d,b_d]$, where we have to take as grid X_{mn} the product of the Chebyshev-Lobatto nodes of each interval. Notice that ε_m is *independent* of p, n and d, and even of the box size. Since $m \geq 2$ and hence $\sigma_m \leq \pi^2/32$, we can immediately write the following rough estimate

$$\varepsilon_m < \frac{2}{m^2} \;, \tag{8}$$

which shows the size of the relative error in terms of $1/m^2$.

We can now approximate both, $\max_K p$ and $\min_K p$, as stated below.

Proposition 2 *Let the assumptions of Lemma 1 be satisfied. Define*

$$p_{max} = \max_{x\in K} p(x) \;,\;\; p_{min} = \min_{x\in K} p(x) \;,$$

$$\tilde{p}_{max} = \tilde{p}_{max}(m,n) = \max_{x\in X_{mn}} p(x) \;,\;\; \tilde{p}_{min} = \tilde{p}_{min}(m,n) = \min_{x\in X_{mn}} p(x) \;. \tag{9}$$

Then

$$\max\{p_{max} - \tilde{p}_{max}, \tilde{p}_{min} - p_{min}\} \leq \varepsilon_m\,(p_{max} - p_{min}) \;. \tag{10}$$

In addition, if $\|p\|_{X_{mn}} = \tilde{p}_{max}$

$$p_{max} - \tilde{p}_{max} \leq \varepsilon_m\, p_{max} \;, \tag{11}$$

whereas if $\|p\|_{X_{mn}} = |\tilde{p}_{min}|$

$$\tilde{p}_{min} - p_{min} \leq \varepsilon_m\, |p_{min}| \;. \tag{12}$$

Moreover, setting $\gamma_m = 1/\cos(\pi/(2m))$*, the following interval estimates hold*

$$\tilde{p}_{max} \leq p_{max} \leq \frac{\gamma_m+1}{2}\,\tilde{p}_{max} - \frac{\gamma_m-1}{2}\,\tilde{p}_{min} \;,$$

$$\frac{\gamma_m+1}{2}\,\tilde{p}_{min} - \frac{\gamma_m-1}{2}\,\tilde{p}_{max} \leq p_{min} \leq \tilde{p}_{min} \;. \tag{13}$$

Proof. Consider the polynomial $q(x) = p(x) - p_{max}$ which is nonpositive in K. We have that $\|q\|_K = |p_{min} - p_{max}| = p_{max} - p_{min}$, and $\|q\|_{X_{mn}} = |\tilde{p}_{min} - p_{max}| = p_{max} - \tilde{p}_{min}$. Then by Lemma 1

$$\tilde{p}_{min} - p_{min} = \|q\|_K - \|q\|_{X_{mn}} \leq \varepsilon_m \|q\|_K = \varepsilon_m (p_{max} - p_{min}) \ .$$

On the other hand, taking $q(x) = p(x) - p_{min}$ which is nonnegative in K, we have that $\|q\|_K = p_{max} - p_{min}$, and $\|q\|_{X_{mn}} = \tilde{p}_{max} - p_{min}$, so that

$$p_{max} - \tilde{p}_{max} = \|q\|_K - \|q\|_{X_{mn}} \leq \varepsilon_m \|q\|_K = \varepsilon_m (p_{max} - p_{min}) \ ,$$

and thus (10) holds.

Now, assume that $\|p\|_{X_{mn}} = \tilde{p}_{max}$. Then $p_{max} \geq \tilde{p}_{max} \geq 0$ and $p_{max} - \tilde{p}_{max} \leq \|p\|_K - \tilde{p}_{max}$ (which holds even if $\|p\|_K = |p_{min}|$). By Lemma 1

$$p_{max} - \tilde{p}_{max} \leq \|p\|_K - \tilde{p}_{max} \leq \varepsilon_m \, \tilde{p}_{max} \leq \varepsilon_m \, p_{max} \ ,$$

that is (11).

In the case where $\|p\|_{X_{mn}} = |\tilde{p}_{min}|$, the estimate (12) follows immediately applying (11) to the polynomial $-p$. For the sake of brevity, we do not report the derivation of the interval estimates (13), that have been essentially obtained in [16] (with a different value of γ_m). □

Remark 3 Observe that the error bounds (11)-(12) are classical relative error bounds, whereas (10) are relative to the range size $\max_K p - \min_K p$, a usual weight in optimization (see, e.g., [9]).

It is also worth pointing out that taking (for fixed n) any subsequence $m_k = \ell^k m_0$, with $\ell, m_0 \geq 2$ integers, then $\tilde{p}_{max}(m_k, n) \to p_{max}$ is *nondecreasing* whereas $\tilde{p}_{min}(m_k, n) \to p_{min}$ is *nonincreasing* in k (due to the Chebyshev grid property that $\mathcal{C}_s \subset \mathcal{C}_{\ell s}$ and thus $X_{m_k n} \subset X_{m_{k+1} n}$, see Lemma 1).

It is now worth comparing Proposition 2 with other error estimates for polynomial optimization by grid evaluation on the hypercube. By no loss of generality, we can concentrate on approximating the minimum. We may summarize the comparison by the following:

Proposition 3 *Consider the rational grid $Q_\nu = \{x \in K : \nu x \in \mathbb{N}^d\}$, $\nu \in \mathbb{N}$, and the Chebyshev grid X_{mn} of Proposition 2. Then, for every fixed $\varepsilon > 0$, there exist $\nu(\varepsilon)$ and $m(\varepsilon)$ such that $\min_{Q_{\nu(\varepsilon)}} p - \min_K p \leq \varepsilon \|p\|_K$ with card$(Q_{\nu(\varepsilon)}) = \mathcal{O}((n^2/\sqrt{\varepsilon})^d)$, whereas $\min_{X_{m(\varepsilon)n}} p - \min_K p \leq \varepsilon \|p\|_K$ with card$(X_{m(\varepsilon)n}) = \mathcal{O}((n/\sqrt{\varepsilon})^d)$.*

Proof. In [10] it is proved, improving a previous estimate [11], that taking the minimum of p on the rational grid $Q_\nu = \{x \in K : \nu x \in \mathbb{N}^d\}$, $\nu \in \mathbb{N}$, we have the error estimate $\min_{Q_\nu} p - \min_K p \leq C_p/\nu^2$, where $C_p = \mathcal{O}(\|\nabla^2 p\|)$, for fixed d, and card$(Q_\nu) = \mathcal{O}(\nu^d)$. Now, in view of the multivariate Markov brothers' inequality (see, e.g., [1]) we have that $\|\nabla^2 p\|_K = \mathcal{O}(n^4) \|p\|_K$ for every $p \in \mathbb{P}_n^d$, the exponent 4 being attained for example with the polynomial $p(x) = p(x_1, \ldots, x_d) = T_n(x_1) + \cdots + T_n(x_d)$. This entails that to obtain an error on the minimum not exceeding ε (relative to $\|p\|_K$), it is sufficient to take

$\nu = \nu(\varepsilon)$ proportional to $n^2/\sqrt{\varepsilon}$, and thus to evaluate p at $\mathcal{O}((n^2/\sqrt{\varepsilon})^d)$ rational nodes.

On the other hand, since in any case $\max_K p - \min_K p \leq 2\|p\|_K$, by (10) in Proposition 2 we can take $m = m(\varepsilon)$ proportional to $1/\sqrt{\varepsilon}$, so that we can evaluate p at $\mathcal{O}((n/\sqrt{\varepsilon})^d)$ Chebyshev grid nodes. $\square$

It is also worth pointing out that in [10] an alternative optimization method has been proposed, with a more sophisticated grid search, whose error is $\mathcal{O}(1/\sqrt{\nu})$ (or even $\mathcal{O}(1/\nu)$ for polynomials having a rational minimizer) by evaluating p at $\mathcal{O}(d^\nu)$ nodes.

On the other hand, in [23] optimization on Chebyshev grids is considered, assuming that the degree of the polynomial in each variable x_i, say n_i, $1 \leq i \leq d$, is known. Then following [16] the estimate $\|p\|_K \leq \gamma_{m,d}\,\|p\|_{Y_m}$ holds, where $\gamma_{m,d} = \left(\frac{1}{\cos(\pi/(2m))}\right)^d$ and $Y_m = \mathcal{C}_{mn_1} \times \cdots \times \mathcal{C}_{mn_d}$, from which bounds like (13) are derived. The underlying polynomial space is essentially $\bigotimes_{i=1}^d \mathbb{P}^1_{n_i}$.

On the contrary, our approach is valid for any polynomial, assuming that we only know a bound for its total degree, $n \geq deg(p)$, and we are able to compute its values at X_{mn}. Indeed, a polynomial is often given as a complicated expression, involving sums, products and powers of other polynomials, possibly written in a basis different from the canonical monomial basis (e.g., some orthogonal basis), or even as a black box, so that the knowledge of the individual degrees n_i is very difficult to obtain (or out of reach).

Even when the individual degrees are at hand, in some instances our approach is more efficient. For example, if $n_i = n$ for every i, given a tolerance ε by Proposition 2 we need evaluating p at $\mathcal{O}((n/\sqrt{\varepsilon})^d)$ grid nodes, whereas following [23] in that case we would evaluate p at $\mathcal{O}((n\sqrt{d}/\sqrt{\varepsilon})^d)$ grid nodes, since $\gamma_{m,d} - 1 = \mathcal{O}(d/m^2)$.

For the purpose of illustration, we have chosen a standard test function in polynomial optimization, the Styblinski-Tang 4th-degree polynomial of three variables

$$p_1(x_1, x_2, x_3) = \sum_{i=1}^{3} \left(\frac{1}{2}\,(10x_i - 5)^4 - 8(10x_i - 5)^2 + \frac{5}{2}\,(10x_i - 5)\right) , \qquad (14)$$

and a second test polynomial

$$p_2(x_1, x_2, x_3) = T_4(x_1 + x_2 + x_3) , \qquad (15)$$

with $x = (x_1, x_2, x_3) \in [0, 1]^3$, where T_4 is the 4th-degree Chebyshev polynomial of the first kind.

We have computed their minimum on the Chebyshev grids X_{mn} of Lemma 1 (see also Remark 2) with $n = 4$ and $m = 2, 4, 8, 16, 32, 64$, as well as on a uniform grid with the same size. The value of the minimum of $p = p_1$ in the cube is $p_{min} = -117.49797$, and $p_{max} = 375$ (see [10]), whereas the minimum of $p = p_2$ is $p_{min} = -1$ and its maximum is $p_{max} = 577$. Comparing the continuum and discrete maxima is not meaningful, since in both cases p_{max} is attained at a vertex of the cube, and all the vertices are nodes of the grids.

The results are collected in Table 1, where we display the relative error

$$E_m = \frac{\tilde{p}_{min} - p_{min}}{p_{max} - p_{min}} \qquad (16)$$

Table 1: Minimization of two 4th-degree polynomials in $[0,1]^3$.

	m	2	4	8	16	32	64
	ε_m	4.5e-1	8.4e-2	2.0e-2	4.8e-3	1.2e-3	3.0e-4
$\min p_1$	E_m^{cheb}	5.2e-2	5.2e-2	1.6e-3	1.6e-3	6.1e-4	6.2e-5
	E_m^{eqsp}	1.5e-2	5.6e-3	8.5e-4	4.6e-4	1.6e-5	1.6e-5
$\min p_2$	E_m^{cheb}	6.7e-6	9.3e-8	7.4e-9	1.2e-14	1.2e-14	7.2e-15
	E_m^{eqsp}	1.6e-2	3.0e-3	1.1e-3	1.3e-4	1.2e-4	4.6e-8

by discrete optimization on Chebyshev grids and equally spaced grids of the same cardinality. We also display the estimate ε_m for the Chebyshev grid optimization error. We see that the error behavior of Chebyshev grid optimization is consistent with Proposition 2 and quite satisfactory, being always below the estimate ε_m, up to several orders of magnitude for p_2.

To conclude, we may observe that though optimization on Chebyshev grids is a kind of "brute force" method, in view of its simplicity and the rigorous error estimates which depend on the sole auxiliary parameter m, it could play some role in low-dimensional polynomial optimization. For example, it could provide a basic step for more sophisticated iterative optimization methods, such as the branch and bound technique discussed in [23], or the trust region method in [8].

References

[1] M. Baran and L. Bialas-Ciez, Hölder continuity of the Green function and Markov brothers' inequality, Constr. Approx. 40 (2014), 121–140.

[2] L. Bos, A Simple Recipe for Modelling a d-cube by Lissajous curves, Dolomites Res. Notes Approx. DRNA 10 (2017), 1–4.

[3] L. Bos, J.-P. Calvi, N. Levenberg, A. Sommariva and M. Vianello, Geometric Weakly Admissible Meshes, Discrete Least Squares Approximation and Approximate Fekete Points, Math. Comp. 80 (2011), 1601–1621.

[4] L. Bos, N. Levenberg and S. Waldron, Metrics associated to multivariate polynomial inequalities, Advances in constructive approximation: Vanderbilt 2003, pp. 133–147, Mod. Methods Math., Nashboro Press, Brentwood, TN, 2004.

[5] L. Bos, N. Levenberg and S. Waldron, Pseudometrics, distances and multivariate polynomial inequalities J. Approx. Theory 153 (2008), 80–96.

[6] L. Bos and M. Vianello, Low cardinality admissible meshes on quadrangles, triangles and disks, Math. Inequal. Appl. 15 (2012), 229–235.

[7] J.P. Calvi and N. Levenberg, Uniform approximation by discrete least squares polynomials, J. Approx. Theory 152 (2008), 82–100.

[8] T.F. Coleman and Y. Li, An interior trust region approach for nonlinear minimization subject to bounds, SIAM J. Optim. 6 (1996), 418–445.

[9] E. de Klerk, The complexity of optimizing over a simplex, hypercube or sphere: a short survey, CEJOR Cent. Eur. J. Oper. Res. 16 (2008), 111–125.

[10] E. de Klerk, J.B. Lasserre, M. Laurent and Z. Sun, Bound-constrained polynomial optimization using only elementary calculations, Math. Oper. Res., published online on March 15, 2017.

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

[12] E. de Klerk, M. Laurent and Z. Sun, An error analysis for polynomial optimization over the simplex based on the multivariate hypergeometric distribution, SIAM J. Optim. 25 (2015), 1498–1514.

[13] E. de Klerk, M. Laurent, Z. Sun and J.C. Vera, On the convergence rate of grid search for polynomial optimization over the simplex, Optim. Lett. 11 (2017), 597–608.

[14] M. Dubiner, The theory of multidimensional polynomial approximation, J. Anal. Math. 67 (1995), 39–116.

[15] H. Ehlich and K. Zeller, Schwankung von Polynomen zwischen Gitter punkten, Math. Z. 86 (1964), 41–44.

[16] U. Gärtel, Error Estimations for Second-Order Vector-Valued Boundary Tasks, in Particular for Problems from Chemical Reaction Diffusion Theory, PhD Thesis, University of Cologne, 1987 (in German).

[17] B. Hashemi and L.N. Trefethen, Chebfun in three dimensions, SIAM J. Sci. Comp., submitted for publication (available online at: `https://people.maths.ox.ac.uk/trefethen/papers.html`).

[18] A. Kroó, On optimal polynomial meshes, J. Approx. Theory 163 (2011), 1107–1124.

[19] F. Piazzon, Optimal polynomial admissible meshes on some classes of compact subsets of $\mathbb{R}^d$, J. Approx. Theory 207 (2016), 241–264.

[20] F. Piazzon and M. Vianello, Small perturbations of polynomial meshes, Appl. Anal. 92 (2013), 1063–1073.

[21] F. Piazzon and M. Vianello, Constructing optimal polynomial meshes on planar starlike domains, Dolomites Res. Notes Approx. DRNA 7 (2014), 22–25.

[22] A. Townsend and L.N. Trefethen, An extension of Chebfun to two dimensions, SIAM J. Sci. Comput. 35 (2013), C495–C518.

[23] J.F. Zhang and C.P. Kwong, Some applications of a polynomial inequality to global optimization, J. Optim. Theory Appl. 127 (2005), 193–205.