



# Uniqueness of solutions in multivariate Chebyshev approximation problems

Vera Roshchina<sup>1</sup> · Nadezda Sukhorukova<sup>2</sup>  · Julien Ugon<sup>3</sup>

Received: 10 November 2022 / Accepted: 19 July 2023 / Published online: 4 August 2023  
© The Author(s) 2023

## Abstract

We study the solution set to multivariate Chebyshev approximation problem, focusing on the ill-posed case when the uniqueness of solutions can not be established via strict polynomial separation. We obtain an upper bound on the dimension of the solution set and show that nonuniqueness is generic for ill-posed problems on discrete domains. Moreover, given a prescribed set of points of minimal and maximal deviation we construct a function for which the dimension of the set of best approximating polynomials is maximal for any choice of domain. We also present several examples that illustrate the aforementioned phenomena, demonstrate practical application of our results and propose a number of open questions.

**Keywords** Chebyshev approximation · Uniqueness of solutions · Multivariate polynomial approximation

**Mathematics Subject Classification** 41A10 · 41A52 · 41A63 · 49K30 · 49K35

---

✉ Nadezda Sukhorukova  
nsukhorukova@swin.edu.au

Vera Roshchina  
v.roshchina@unsw.edu.au

Julien Ugon  
julien.ugon@deakin.edu.au

<sup>1</sup> UNSW Sydney, High St., Kensington, NSW 2052, Australia

<sup>2</sup> Swinburne University of Technology, John Street, Hawthorn, VIC 3122, Australia

<sup>3</sup> Deakin University, 221 Burwood Highway, Burwood, VIC 3125, Australia

# 1 Introduction

## 1.1 Background

The classical Chebyshev approximation problem is to construct a polynomial of a given degree that has the smallest possible absolute deviation from some continuous function on a given interval. For univariate polynomials of degree  $d \geq 0$  the solution is unique and satisfies an elegant alternation condition: there exist  $d + 2$  points of alternating minimal and maximal deviation of the function from approximating polynomial [7] (see Fig. 1).

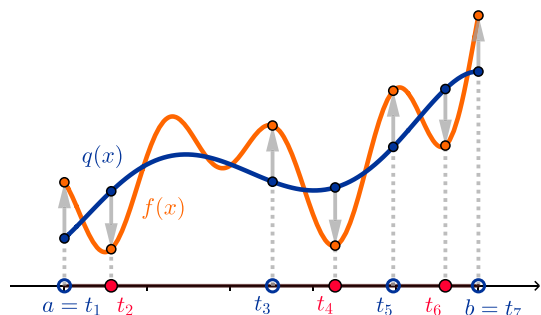
Once we depart from the classical case and consider approximating a continuous function on a compact subset  $X$  of  $\mathbb{R}^n$  by multivariate polynomials, the uniqueness is lost: the result of Mairhuber [10] demonstrates that a multivariate Chebyshev approximation problem has a unique solution generically (for all continuous functions on a given compact subset of  $\mathbb{R}^n$ ) if and only if the underlying set  $X$  is homeomorphic to a closed subset of a circle. In particular, if  $X \subset \mathbb{R}^n$  contains an interior point, then there is no Haar (Chebyshev) space of dimension  $n \geq 2$  for  $X$  (i.e. there is no finite system of continuous functions such that every continuous function on  $X$  has a unique Chebyshev approximation in the span of this system). An example of such nonunique approximation is shown in Fig. 2.

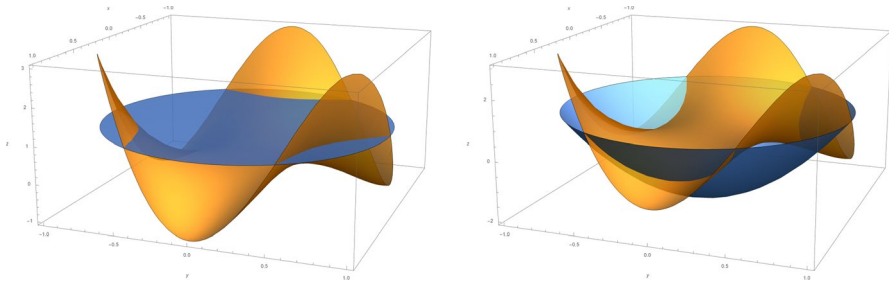
Even though the uniqueness of solutions is lost in the multivariate case, the alternation result holds in the form of algebraic separation. It was first shown in [15] that a polynomial approximation of degree  $d$  is optimal if the sets of points of minimal and maximal deviation can not be separated by a polynomial of degree at most  $d$ . This result can be reproduced using the standard tools of modern convex analysis, as demonstrated in [19]. Another approach to generalise the notion of alternation to multivariate problems is based on the alternating signs of certain determinants [8].

## 1.2 Motivation

The classical alternation result was obtained by Chebyshev in 1854 [7], but little is known about the shape of the solutions of a more general multivariate problem. In particular, related work [4] that studies a version of this problem for polynomials with integer coefficients, mentions that the multivariate problem is ‘virtually

**Fig. 1** A typical distribution of the points of minimal and maximal deviation of a continuous function ( $f$ , shown in orange) from its best Chebyshev approximation by a polynomial of degree at most 5 (denoted by  $q$ , shown in blue) on a bounded interval  $[a, b]$





**Fig. 2** The function  $f(x, y) = x^6 + y^6 + 3x^4y^2 + 3x^2y^4 + 6xy^2 - 2x^3$  has several best quadratic approximations on the disk  $x^2 + y^2 \leq 1$ . The plot of the function in orange colour is shown together with two different best approximations in blue:  $q_0(x, y) = 1$  (on the left) and  $q_1(x, y) = 3x^2 + 3y^2 - 2$  (on the right)

untouched'. Even though the solutions to the multivariate problem satisfy a form of an alternation condition, the structure of the solutions and the location of points of maximal and minimal deviation are more complex compared to the univariate case, which results in many interesting challenges.

From the point of view of classical approximation theory multivariate polynomial approximation is relatively inefficient: for a range of key applications some other approaches such as the radial basis functions [6] provide superior results. However modern optimisation is increasingly fusing with computational algebraic geometry, successfully tackling problems that were insurmountable in the past, and polynomial approximation emerges in this context as valuable not only for solving computationally challenging problems, but also as an analytic tool that together with Gröbner basis methods may lead to algorithmic solutions for finding extrema in nonconvex problems. Another potential application is a generalisation of trust-region methods, where instead of local quadratic approximations to the function locally more versatile higher order polynomial approximations may be used.

It is also important to mention rational approximation [1]. Rational functions are able to approximate nonsmooth and abruptly changing functions and there are a number of efficient methods for univariate rational approximation [3, 12]. Some of these methods have been extended to multivariate function approximation [2], but the choice is not so extensive. Rational functions are ratios of two polynomials and therefore their advances require a better understanding of polynomial approximation. All these motivate us to study polynomial approximation in details.

Consider the space  $\mathbb{P}_d(\mathbb{R}^n)$  of real polynomials in  $n$  variables of degree at most  $d$ . Let  $f : X \rightarrow \mathbb{R}$  be a continuous function defined on a compact set  $X \subset \mathbb{R}^n$ . A polynomial  $q^* \in \mathbb{P}_d(\mathbb{R}^n)$  solves the multivariate Chebyshev approximation problem for  $f$  on  $X$  if

$$\max_{x \in X} |f(x) - q^*(x)| \leq \max_{x \in X} |f(x) - q(x)| \quad \forall q \in \mathbb{P}_d(\mathbb{R}^n).$$

We are interested in the set  $Q \subset \mathbb{P}_d(\mathbb{R}^n)$  of all such solutions. In some special cases the solution to the multivariate Chebyshev approximation problem is known explicitly. For instance, the best approximation by monomials on a unit cube is obtained

from the products of classical Chebyshev polynomials (see [21] and a more recent overview [22]); this is related to another generalisation of Chebyshev's results, when the problem of a best approximation of zero with polynomials having a fixed highest degree coefficient is considered: in some special cases, solutions on the unit cube are known from [17]; solutions for the unit ball were obtained in [13].

There is a different approach to generalising Chebyshev polynomials, based on extending the relation  $T_k(\cos x) = \cos kx$  to the multivariate case. In [11, 16] more general functions  $h : \mathbb{R}^n \rightarrow \mathbb{R}^n$  periodic with respect to fundamental domains of affine Weyl groups are considered, and the aforementioned relation is replaced by  $P_k(h(x)) = h(kx)$ . Such generalised Chebyshev polynomials are in fact systems of polynomials, as  $P_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . We note here that the aforementioned work, as well as other approximation techniques based on Chebyshev polynomials (common in numerical PDEs), use nodal interpolation with Chebyshev polynomials. This is a conceptually different framework compared to our optimisation setting; in particular, this approach requires a careful choice of interpolation nodes on the domain to ensure the quality of approximation.

### 1.3 Challenges

For the univariate problem the optimal solutions to the Chebyshev approximation problem can be obtained using numerical techniques that fit in the context of linear programming and the simplex method, and exchange algorithm pioneered by Remez [14] is perhaps the most well-known technique. Even though the multivariate problem can be solved approximately by linear programming, the problem rapidly becomes intractable with the increase in the degree and number of variables, and hence there is much need for more efficient methods. This is another exciting research direction, as the rich structure of the problem is likely to yield specialised methods which surpass the performance of direct linear programming discretisation. The general framework for the potential generalisation of the exchange approach was laid out in [18–20]. In these papers, authors (partially) extended de la Vallée-Poussin procedure, which is the core of the Remez method. However, several implementation issues need to be resolved for a practically viable version of the method. It is also possible that some of these issues can not be extended to the case of multivariate polynomials due to the the loss of uniqueness of optimal solution, which is the target of this paper.

For any polynomial  $q$  we can define the sets of points of minimal and maximal deviation, i.e. such  $x \in X$  for which the values  $q(x) - f(x)$  and  $f(x) - q(x)$  respectively coincide with the maximum  $\max_{x' \in X} |f(x') - q(x')|$ . These sets may be different for different polynomials in the optimal set  $Q$ . We show that it is possible to identify an intrinsic pair of such subsets pertaining to all polynomials in  $Q$  (see Theorem 3); moreover the location of these points determines the maximal possible dimension of the solution set (see Lemma 1). We also show that for any prescribed arrangement of points of minimal and maximal deviation and any choice of the maximal degree there exists a continuous function and a relevant approximating polynomial for which these points are precisely the points of minimal and

maximal deviation; moreover, the set of all best approximations has the largest possible dimension, for any choice of domain  $X$  (Lemma 2). Finally, we show that the set of best Chebyshev approximations is always of the maximal possible dimension if the domain  $X$  is finite (Lemma 3). All these constructions are essential for designing computational algorithms for multivariate polynomial approximations. In particular, since the basis functions do not form a Chebyshev system in multivariate cases, the proofs of convergence are much more challenging due to the necessity to trace several possibilities.

We begin with some preliminaries and examples in Sect. 2, focussing on the well-known separation characterisation of optimality and Mairhuber’s uniqueness result. In Sect. 3 we present our new results. We then summarise our findings and present some open problems in Sect. 4.

## 2 Preliminaries and examples

### 2.1 Multivariate polynomials

A multivariate polynomial of degree  $d$  with real coefficients can be represented as

$$q(x) = \sum_{|\alpha| \leq d} a_\alpha x^\alpha,$$

where  $\alpha = (\alpha_1, \dots, \alpha_n)$  is an  $n$ -tuple of nonnegative integers,

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n},$$

$|\alpha| = |\alpha_1| + |\alpha_2| + \dots + |\alpha_n|$ , and  $a_\alpha \in \mathbb{R}$  are the coefficients. All polynomials of degree not exceeding  $d$  constitute a vector space  $\mathbb{P}_d(\mathbb{R}^n) = \text{span}\{x^\alpha \mid |\alpha| \leq d\}$  of dimension  $\binom{n+d}{d}$ .

Note that, generally speaking, we can consider any finite set of (linearly independent) polynomials in  $n$  variables,  $G = (g_1, \dots, g_N)$  and instead of the space  $\mathbb{P}_d(\mathbb{R}^n)$  consider the linear span  $V$  of  $G$ , i.e.

$$V = \text{span}\{g_i \mid i \in \{1, \dots, N\}\}. \tag{1}$$

Then the solution set  $Q \subseteq V$  to the Chebyshev approximation problem for a given continuous function  $f$  defined on a compact set  $X \subseteq \mathbb{R}^n$  is

$$Q := \text{Arg min}_{q \in V} \|f - q\|_\infty, \tag{2}$$

where

$$\|f - q\|_\infty = \max_{x \in X} |f(x) - q(x)|.$$

Fixing a continuous function  $f : X \rightarrow \mathbb{R}$ , for every polynomial  $q \in V$  we define the sets of points of minimal and maximal deviation explicitly as

$$\begin{aligned}\mathcal{N}(q) &:= \{x \in X \mid q(x) - f(x) = \|f - q\|_\infty\}, \\ \mathcal{P}(q) &:= \{x \in X \mid f(x) - q(x) = \|f - q\|_\infty\}.\end{aligned}\quad (3)$$

Observe that for any given polynomial  $q$  at least one of these sets is nonempty, and for any  $q^* \in Q$  both of them are nonempty (otherwise one can add an appropriate small constant to  $q^*$  and decrease the value of the maximal absolute deviation). Also observe that the sets  $\mathcal{N}(q)$  and  $\mathcal{P}(q)$  are disjoint unless  $q \equiv f$  on  $X$  (in this case  $\mathcal{N}(q) = \mathcal{P}(q) = X$ ).

The minimisation problem of (2) is an unconstrained convex optimisation problem: the objective function  $\|f - q\|_\infty$  can be interpreted as the maximum over two families of linear functions parametrised by the domain variable  $x \in X$ , i.e.

$$\|f - q\|_\infty = \max_{x \in X} |f(x) - q(x)| = \max_{\substack{x \in X \\ \sigma \in \{-1, 1\}}} \sigma(f(x) - q(x)). \quad (4)$$

The solution set  $Q$  is nonempty, since it represents the metric projection of  $f$  onto a finite-dimensional linear subspace  $V$  of the normed linear space of functions bounded on  $X$ . It is also easy to see from the continuity of  $f$  that this set is closed. Moreover, since a maximum function over a family of linear functions is convex,  $Q$  is convex (e.g. see [9, Proposition 2.1.2]).

**Example 1** (Solution set is unbounded) We consider a degenerate case of the problem: find the best linear approximation to  $f(x, y) = x^2$  on  $X = [-1, 1] \times \{0\}$ . Since the domain is effectively restricted to the line segment  $[-1, 1]$ , the solution reduces to the classical univariate case: there is a unique best approximation, which happens to be constant,  $\frac{1}{2}$ . Observe however that in the true two-dimensional setting any linear polynomial of the form  $q(x, y) = \frac{1}{2} + \alpha y$  is also a best approximation of  $f$  on  $X$ . This means that the solution set of best approximations is unbounded,  $Q = \{\frac{1}{2} + \alpha y, \alpha \in \mathbb{R}\}$ , even though all such optimal solutions coincide on  $X$ , and effectively—on the set  $X$ —provide the same unique best approximation.

## 2.2 Optimality conditions

**Definition 1** We say that a polynomial  $p \in V$  separates two sets  $N, P \subset \mathbb{R}^n$  if

$$p(x) \cdot p(y) \leq 0 \quad \forall x \in N, y \in P; \quad (5)$$

we say that the separation is *strict* if the inequality in (5) is strict, i.e.

$$p(x) \cdot p(y) < 0 \quad \forall x \in N, y \in P. \quad (6)$$

Recall the well-known characterisations of optimality (see [15] for the original result and [19] for modern proofs). These important results are highlighted in Sect. 1 as our preliminary background of the problem.

**Theorem 1** *Let  $X$  be a compact subset of  $\mathbb{R}^n$ , and assume that  $f : X \rightarrow \mathbb{R}$  is a continuous function. A polynomial  $q \in V$  is an optimal solution to the Chebyshev approximation problem (2) if and only if there exists no  $p \in V$  that strictly separates the sets  $\mathcal{N}(q)$  and  $\mathcal{P}(q)$ .*

**Example 2** (Best quadratic approximation is not unique) We focus on the function  $f(x, y) = x^6 + y^6 + 3x^4y^2 + 3x^2y^4 + 6xy^2 - 2x^3$  discussed in the Introduction and demonstrate that it does indeed have multiple best quadratic approximations on the disk  $x^2 + y^2 \leq 1$  (see Fig. 2).

For two different polynomials  $q_0(x, y) = 1$  and  $q_1(x, y) = 3x^2 + 3y^2 - 2$  the points of maximal negative and positive deviation of  $f$  from these polynomials are

$$\mathcal{N}(q_0) = \{z_1, z_3, z_5\}, \quad \mathcal{N}(q_1) = \mathcal{N}(q_0) \cup \{z_0\}, \quad \mathcal{P}(q_0) = \mathcal{P}(q_1) = \{z_2, z_4, z_6\},$$

where

$$z_0 = (0, 0), \quad z_1 = (1, 0), \quad z_2 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad z_3 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right),$$

$$z_4 = (-1, 0), \quad z_5 = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right), \quad z_6 = \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right). \tag{7}$$

This is not difficult to verify using standard calculus techniques (see appendix).

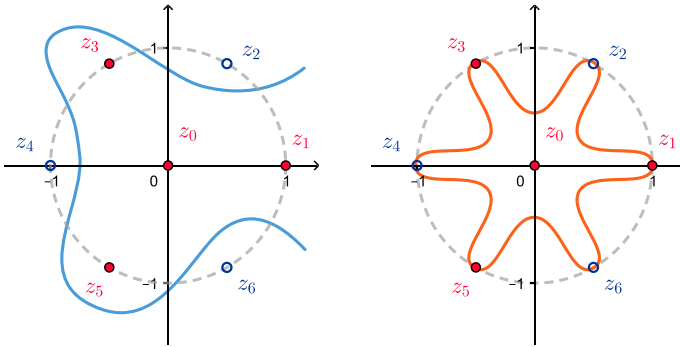
### 2.3 Location of maximal and minimal deviation points

Observe that the points  $z_1, z_2, \dots, z_6$  lie on the unit circle. By the Bézout theorem, this circle can have at most 4 intersections with any other quadratic curve. However if we could find a quadratic polynomial that strictly separates the points of maximal and minimal deviation, the relevant curve would intersect the circle in at least six points, as shown in Fig. 3.

Hence such separation is impossible, so both  $q_0$  and  $q_1$  are optimal.

We conclude this section with the well-known result of Mairhuber [10] (generalised to compact Hausdorff spaces by Brown [5]). These results contributed to the motivation for this paper, since the uniqueness of the optimal solution is lost (Sect. 1).

**Theorem 2** (Mairhuber) *A compact subset  $X$  of  $\mathbb{R}^n$  containing at least  $k \geq 2$  points may serve as the domain of definition of a set of real continuous functions  $f_1(x), \dots, f_k(x)$  that provide a unique Chebyshev approximation to any continuous function  $f$  on the set  $X$ , if and only if  $X$  is homeomorphic to a closed subset of the circumference of a circle.*



**Fig. 3** On the left: the intersection of two quadratic curves at six points contradicts the Bézout theorem; on the right: a subset of the unit disk homeomorphic to a circle

With relation to our setting, Mairhuber's result is a necessary condition for uniqueness, since our choice of the system of functions is restricted to multivariate polynomials. Hence it is possible to identify a compact set  $X$  homeomorphic to a circle and a set of polynomials linearly independent on  $X$  that do not provide a unique multivariate approximation to a continuous function on  $X$ .

**Example 3** Observe that any best approximation to  $f$  from Example 2 on the disk is also the best approximation to  $f$  on any subset of the disk that contains the sets  $\mathcal{N}(q_0)$  and  $\mathcal{P}(q_0)$ . Even though the two different best approximations  $q_0$  and  $q_1$  coincide on the boundary of the disk, they take different values everywhere in the interior, and hence we can choose another subset of the unit disk that is homeomorphic to a circle (like the one shown in Fig. 3 on the right) to obtain two different optimal solutions. This does not contradict Mairhuber's theorem, since in this case we have restricted ourselves to a very specific choice of the basic functions.

### 3 Structure of the solution set

#### 3.1 The location of maximal and minimal deviation points for different optimal solutions

The key technical result of this section is the following theorem that establishes the existence of uniquely defined subsets of points of maximal and minimal deviation across all optimal solutions. This means that the points of maximal and minimal deviation do not wander around the domain  $X$  as we move from one optimal solution to another.

**Theorem 3** *Let  $f : X \rightarrow \mathbb{R}$  be a continuous function defined on a compact set  $X \subset \mathbb{R}^n$ , let  $V$  be a subspace of multivariate polynomials in  $n$  variables (1), and suppose that  $Q$  is the set of optimal solutions to the relevant optimisation problem, as in (2). Then*



- (i)  $\mathcal{N}(q) = \mathcal{N}(p), \mathcal{P}(q) = \mathcal{P}(p) \forall p, q \in \text{ri } Q;$
- (ii)  $\mathcal{N}(q) \subseteq \mathcal{N}(p), \mathcal{P}(q) \subseteq \mathcal{P}(p) \forall q \in \text{ri } Q, p \in Q.$

Here the relative interior is considered with respect to the convex sets of the coefficients in the representation of the solutions as linear combinations of polynomials in  $V$ .

For the proof of this theorem, we will need the following elementary result about max-type convex functions. In particular, we prove the following proposition.

**Proposition 1** *Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}$  be a pointwise maximum over a family of linear functions,*

$$F(v) = \max_{t \in T} F_t(v), \quad F_t : \mathbb{R}^n \rightarrow \mathbb{R} \text{ linear } \forall t \in T.$$

*Let  $I(v) = \{t \mid F_t(v) = F(v)\}, Q := \text{Arg min}_{v \in \mathbb{R}^n} F(v)$ . If  $Q \neq \emptyset$ , then*

$$I(v) \subseteq I(u) \quad \forall v \in \text{ri } Q, u \in Q.$$

**Proof** Let  $v \in \text{ri } Q, u \in Q$ . Assume that there exists  $t \in T$  such that  $t \in I(v) \setminus I(u)$ . Then  $F(v) = F(u) = F_t(v) > F_t(u)$ , and since  $F_t$  is linear, we then have

$$F(v - \alpha(u - v)) \geq F_t(v - \alpha(u - v)) = F_t(v) - \alpha(F_t(u) - F_t(v)) > F(v) \quad \forall \alpha > 0,$$

hence,  $v - \alpha(u - v) \notin Q$  for  $\alpha > 0$ , while  $u = v + (u - v) \in Q$ , which means  $v \notin \text{ri } Q$ . □

**Proof of Theorem 3** Recall that our objective function can be represented as the maximum over a family of linear functions, as in (4). For every polynomial  $q \in V$  define the set of active indices

$$I(q) = \{(x, \sigma) \in X \times \{-1, 1\} \mid \sigma(f(x) - q(x)) = \|f - q\|_\infty\}.$$

It is evident from the definition (3) of  $\mathcal{N}(q)$  and  $\mathcal{P}(q)$  that

$$x \in \mathcal{N}(q) \Leftrightarrow (x, -1) \in I(q); \quad x \in \mathcal{P}(q) \Leftrightarrow (x, 1) \in I(q).$$

The result now follows from Proposition 1. □

The following corollary of Theorem 3 characterises the structure of the location of maximal deviation points corresponding to different optimal solutions.

**Corollary 1** *The sets of points of minimal and maximal deviation remain constant if the optimal solutions belong to the relative interior of the solution set. Additional maximal and minimal deviation points can only occur if an optimal solution is on the relative boundary.*

For any given continuous function  $f$  defined on a compact set  $X$  we can hence define the *minimal* or *essential* sets of points of minimal and maximal deviation,

$$\mathcal{P} = \mathcal{P}(q), \mathcal{N} = \mathcal{N}(q), q \in \text{ri } Q,$$

where  $\mathcal{P}(q)$  and  $\mathcal{N}(q)$  are defined in the standard way, as in (3). For instance, in Example 2 we have  $\mathcal{N} = \{z_1, z_3, z_5\}$  and  $\mathcal{P} = \{z_2, z_4, z_6\}$ , while  $\mathcal{N}(q_1)$  contains an additional point  $z_0$ .

### 3.2 Dimension of the solution set

We next focus on the relation between the family of separating polynomials and the dimension of solution set.

For a fixed continuous function  $f : X \rightarrow \mathbb{R}$  and a polynomial  $q \in V$  consider the set of all polynomials in  $V$  that separate the points of minimal and maximal deviation,

$$S(q) = \{s \in V \mid s(x) \cdot s(y) \leq 0 \forall x \in \mathcal{P}(q), y \in \mathcal{N}(q)\}.$$

Notice that the zero polynomial is always in  $S(q)$ , and for the polynomials in the optimal solution set we may have a nontrivial set of separating functions. This happens in particular when all points of minimal and maximal deviation are located on an algebraic variety of a subset of  $V$ .

Since the pair of sets of minimal and maximal deviation is minimal on the relative interior of  $Q$ , and such minimal pair is unique according to Theorem 3, we can define the *maximal* set of separating polynomials as  $S = S(q)$  for  $q \in \text{ri } Q$ .

For the rest of the section, we work with an arbitrary fixed continuous real-valued function  $f$  defined on a compact set  $X \subset \mathbb{R}^n$ , so we do not repeat this assumption in each statement, and simply refer to the solution set  $Q$  of the corresponding Chebyshev approximation problem.

**Lemma 1** *For the solution set  $Q$  we have  $\dim Q \leq \dim S$ ; moreover, for any  $q \in \text{ri } Q$ ,  $p \in Q$  we have  $p - q \in S(p) \subseteq S$ .*

**Proof** Observe that it is enough to show that for any  $q \in \text{ri } Q$  and any  $p \in Q$  we have  $p - q \in S$ . It then follows that  $\text{aff } Q \subseteq S + q$ , and hence  $\dim Q \leq \dim S$ .

Let  $q \in \text{ri } Q$  and assume  $p \in Q$ . Then  $\|f - q\|_\infty = \|f - p\|_\infty$ . By Theorem 3 we have  $\mathcal{N}(q) \subseteq \mathcal{N}(p)$ ,  $\mathcal{P}(q) \subseteq \mathcal{P}(p)$ , therefore:

- if  $u \in \mathcal{N}(q)$ , then  $q(u) - f(u) = p(u) - f(u) = \|f - p\|_\infty$ ;
- if  $u \in \mathcal{N}(p) \setminus \mathcal{N}(q)$ , then  $q(u) - f(u) < \|f - p\|_\infty = p(u) - f(u)$ .

And a similar relation, with inverse inequalities apply for  $u \in \mathcal{P}(p)$ . Therefore:

$$\begin{aligned} f(u) - p(u) &\leq f(u) - q(u) \quad \forall u \in \mathcal{N}(p), \\ f(u) - p(u) &\geq f(u) - q(u) \quad \forall u \in \mathcal{P}(p). \end{aligned}$$

Let  $s(x) = p(x) - q(x)$ . We have

$$s(u) = p(u) - q(u) \geq 0 \quad \forall u \in \mathcal{N}(p), \quad s(u) = p(u) - q(u) \leq 0 \quad \forall u \in \mathcal{P}(p),$$

and so  $s(u) \in S(p) \subseteq S(q)$ . □

**Corollary 2** *If for the solution set  $Q$  we have  $\dim Q > 0$ , then all essential points of minimal and maximal deviation lie on an algebraic variety of some nontrivial polynomial  $s \in V$ .*

**Proof** This follows directly from a modification of the proof of Lemma 1: if  $Q$  is of dimension 1 or higher, then there exist two different polynomials  $q \in \text{ri } Q$  and  $p \in Q$ . We have

$$\begin{aligned} f(u) - p(u) &= f(u) - q(u) \quad \forall u \in \mathcal{N} = \mathcal{N}(q), \\ f(u) - p(u) &= f(u) - q(u) \quad \forall u \in \mathcal{P} = \mathcal{P}(q). \end{aligned}$$

Hence,  $s(u) = p(u) - q(u) = 0 \quad \forall u \in \mathcal{N} \cup \mathcal{P}$ . □

The next corollary is a well-known uniqueness result.

**Corollary 3** *If the set  $S$  is trivial, then the optimal solution is unique.*

**Proof** If  $S = \{0\}$ , then  $\dim S = 0$ , and by Lemma 1 we have  $\dim Q = 0$ . □

### 3.3 Uniqueness and the location of maximal deviation points

It may happen that the dimensions of  $Q$  and  $S$  do not coincide. Consider the following example.

**Example 4** Let  $f(x, y) = (x^2 - \frac{1}{2})(1 - y^2)$  and consider the problem of finding a best linear approximation of this function on the square  $X = [-1, 1] \times [-1, 1]$ .

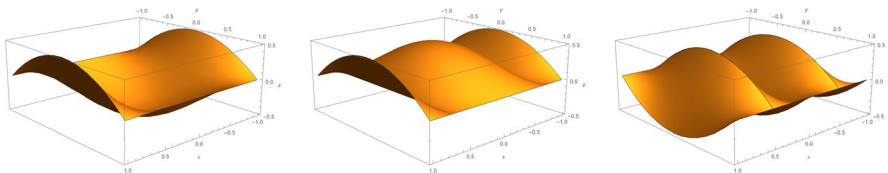
It is not difficult to verify that the constant function  $q_0(x, y) \equiv 0$  is an optimal solution: the points of maximal deviation are the maxima of  $f(x, y)$  on the square, attained at  $\mathcal{P}(q_0) = \{(1, 0), (-1, 0)\}$ ; the set of points of minimal deviation is a singleton  $\mathcal{N}(q_0) = \{(0, 0)\}$  (we provide technical details in the appendix).

Since these three alternating points of maximal and minimal deviation lie on a straight line  $y = 0$ , there is no strict linear separator between them (see the left image in Fig. 5), hence this constant solution must be optimal by Theorem 1. Also notice that taking any point out of either  $\mathcal{N}(q_0)$  or  $\mathcal{P}(q_0)$  ruins the optimality condition (in fact, our configuration of the points of minimal and maximal deviation is also known as *critical set* in the notation of [15]). Hence we must have  $\mathcal{N} = \mathcal{N}(q_0)$  and  $\mathcal{P} = \mathcal{P}(q_0)$ , so these are the essential sets of the points of minimal and maximal deviation. These three points can be separated non-strictly by the linear functions of the form  $l(x, y) = \alpha y, \alpha \in \mathbb{R}$ . We therefore have

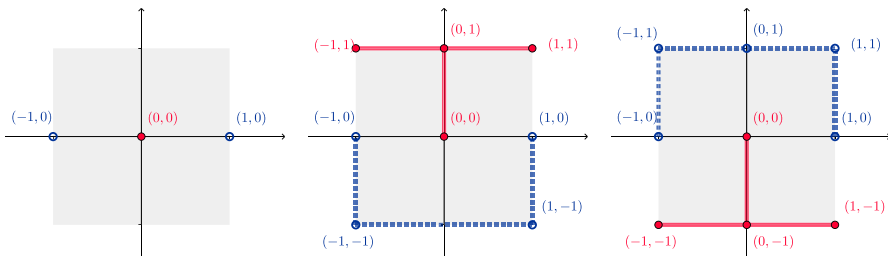
$$S = \{\alpha y \mid \alpha \in \mathbb{R}\}.$$

Even though  $\dim S = 1$ , the best linear approximation is unique. It follows from Lemma 1 that  $Q \subseteq S$ , and hence any best linear approximation should have the form  $q_\alpha(x, y) = \alpha y$  for some  $\alpha \in \mathbb{R}$ . When  $x = \pm 1$ , we have the deviation  $d_\alpha(x, y) = f(x, y) - q_\alpha(x, y) = \frac{1-y^2}{2} - \alpha y$ . The maximum of  $d_\alpha(x, y)$  is attained at  $y = -\alpha$ , with the value  $d_\alpha(\pm 1, -\alpha) = \frac{1}{2} + \frac{\alpha^2}{2} > \frac{1}{2}$  for  $\alpha \neq 0$ , which means that there are no optimal solutions in the neighbourhood of  $q_0(x, y) \equiv 0$ , and hence, due to the convexity of  $Q$ , the best approximation is unique.

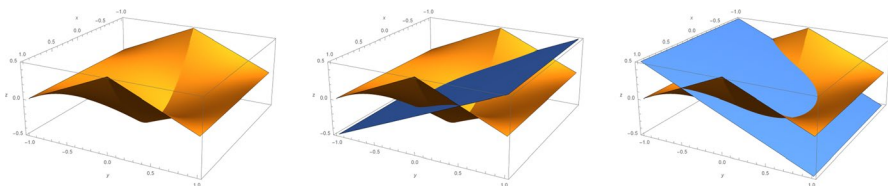
Now consider a modified example: let  $h(x, y) = (x^2 - \frac{1}{2})(1 - |y|)$  (see Fig. 6, left hand side). The same trivial constant function  $q_0(x, y) \equiv 0$  is a best linear approximation to  $h$ , with the same sets of points of minimal and maximal deviation (see Fig. 4, right). However, this best approximation is not unique: any function  $q_\alpha(x, y) = \alpha y$  for  $\alpha \in [-\frac{1}{2}, \frac{1}{2}]$  is also a best linear approximation of  $f$  on the square  $X$



**Fig. 4** The function  $f(x, y) = (x^2 - \frac{1}{2})(1 - y^2)$  (on the left), the absolute deviation of  $f$  from the constant  $q_0(x, y) \equiv 0$ ,  $|d_0(x, y)| = |f(x, y) - q_0(x, y)|$  (middle), and the function  $g(x, y) = (\min\{|2x|, 2 - |2x|\} - 1/2)(1 - y^2)$ , (on the right)



**Fig. 5** The points of minimal and maximal deviation for different cases: on the left for  $h$  (and also  $f$ ) approximated by  $q_0$ ; in the middle for  $h$  and  $q_{\frac{1}{2}}(x, y) = \frac{y}{2}$ ; on the right for  $h$  and  $q_{-\frac{1}{2}}(x, y) = -\frac{y}{2}$



**Fig. 6** The function  $h(x, y) = (x^2 - \frac{1}{2})(1 - |y|)$  on the left, and the same function shown together with two different best approximations:  $q_{\frac{1}{2}}(x, y) = \frac{y}{2}$  and  $q_{-\frac{1}{2}}(x, y) = -\frac{y}{2}$

(see appendix for technical computations). Moreover, the sets of points of maximal and minimal deviation are different at the endpoints of the optimal interval, i.e. for  $\alpha = \pm \frac{1}{2}$ , see Fig. 5 (the technical computations are presented in appendix).

### 3.4 Uniqueness and smoothness

Finally, we would like to point out that smoothness of the function that we are approximating is not necessary for the uniqueness of a best approximation, as one may be tempted to conclude from the study of the functions  $f$  and  $h$ . Note that for yet another modification,

$$g(x, y) := (\min\{|2x|, 2 - |2x|\} - 1/2)(1 - y^2),$$

the function  $q_0(x, y) \equiv 0$  is a unique best approximation, while the points of maximal and minimal deviation are distributed in a similar fashion, along the line  $y = 0$ , potentially allowing for nonuniqueness. Notice that the function  $g(x, y)$  is nondifferentiable at the points of minimal and maximal deviation. This function is however smooth in  $y$  for every fixed  $x$ . This observation is related to the problem of relating the specific (partial) smoothness properties of the function we are approximating with the solution set. We discuss this open question in some detail in the conclusions section.

We have seen from the preceding example that whether the Chebyshev approximation problem has a solution is determined not only by the location of points of maximal and minimal deviation, but also by the properties of the function that is being approximated; in particular the smoothness of the function at the points of minimal and maximal deviation appears to be a decisive factor.

**Example 5** For the distribution of points of maximal and minimal deviation from Example 2, i.e.  $N = \{z_1, z_3, z_5\}$ ,  $P = \{z_2, z_4, z_6\}$ , where  $z_1, z_2, \dots, z_6$  are defined by (7), we construct a nonsmooth continuous function

$$f(x) = f_1(x) - f_2(x),$$

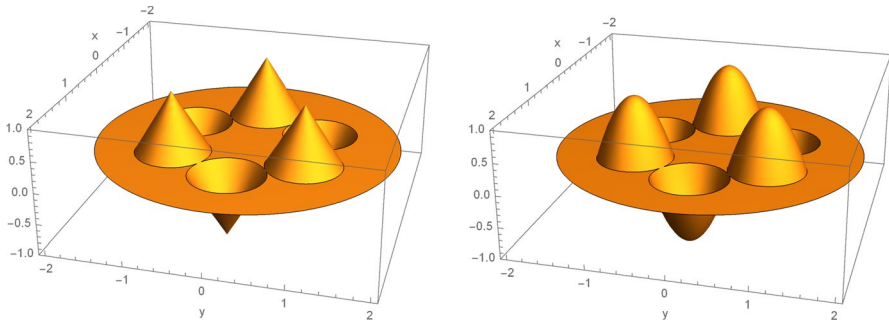
where

$$\begin{aligned} f_1(x) &= \min\{2\|x - z_1\|, 2\|x - z_3\|, 2\|x - z_5\|, 1\}, \\ f_2(x) &= \min\{2\|x - z_2\|, 2\|x - z_4\|, 2\|x - z_6\|, 1\}, \end{aligned}$$

shown in Fig. 7 on the left.

The function  $g(x, y) = 0$  is an optimal solution to the quadratic approximation problem for the function  $f$  on  $X = \{x \mid \|x\| \leq 2\}$  (since this is exactly the same pattern of points of minimal and maximal deviation as discussed in one of the two cases in Example 2). Moreover, the polynomial

$$q_\alpha(x, y) = \alpha(x^2 + y^2 - 1)$$



**Fig. 7** The functions  $f$  and  $h$  in Example 5

is also a best approximation of  $f$  for sufficiently small values of  $\alpha$  (this may be already evident to the reader from the plot; the mathematically rigorous reasons for this will be laid out in the proof of Lemma 2).

Modifying the ‘bump’ that defines each of the peaks that correspond to the points of minimal and maximal deviation so that the function  $f$  is smooth around these points, results in the uniqueness of the approximation  $q_0$ . Indeed, let

$$h(x) = h_1(x) - h_2(x),$$

where

$$\begin{aligned} h_1(x) &= \min\{4\|x - z_1\|^2, 4\|x - z_3\|^2, 4\|x - z_5\|^2, 1\}, \\ h_2(x) &= \min\{4\|x - z_2\|^2, 4\|x - z_4\|^2, 4\|x - z_6\|^2, 1\}, \end{aligned}$$

this function is shown in Fig. 7 on the right.

The same constant polynomial  $q_0(x, y) = 0$  is optimal for  $h$ , however, this time the solution is unique: indeed, suppose that another polynomial in  $S$  provides a best approximation. This polynomial must be of the form  $p_\alpha(x, y) = \alpha(x^2 + y^2 - 1)$  for some  $\alpha \neq 0$ . By convexity of the solution set,  $p_{\alpha'}$  should also be optimal for any  $\alpha'$  between 0 and  $\alpha$ .

In the neighbourhood of the point  $z_1$  we have  $h(x, y) = 4\|x - z_1\|^2 - 1 = 4[(x - 1)^2 + y^2] - 1$ . Then for a sufficiently small  $|\alpha'|$

$$h\left(\frac{4}{4 - \alpha'}, 0\right) - p_{\alpha'}\left(\frac{4}{4 - \alpha'}, 0\right) = -1 - \frac{(\alpha')^2}{4 - \alpha'} < -1,$$

hence this is not a solution.

The next result provides a more general justification for the non-uniqueness of the approximation to a nonsmooth function  $f$  that we have just considered.

**Lemma 2** *Let  $V$  be as in (1), and let  $N$  and  $P$  be two disjoint compact subsets of  $\mathbb{R}^n$  such that they can not be separated strictly by a polynomial in  $V$ . Let*

$$S = \{s \in V \mid s(x) \cdot s(y) \leq 0 \forall x \in P, y \in N\}.$$

*There exists a continuous function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  such that for any compact  $X \in \mathbb{R}^n$  with  $N, P \subseteq X$ , the optimal solution set  $Q$  to the relevant optimisation problem satisfies  $\dim Q = \dim S$ . Moreover, there exists  $q_0 \in Q$  such that  $\mathcal{P}(q_0) = P, \mathcal{N}(q_0) = N$ .*

**Proof** Let

$$f(x) := \max_{u \in P} \varphi_u(x) - \max_{v \in N} \varphi_v(x),$$

where

$$\varphi_u(x) = \max \left\{ 1 - \frac{2}{d} \|x - u\|, 0 \right\}, \quad d = \min_{\substack{u \in P \\ v \in N}} \|u - v\|.$$

Fix a compact set  $X \subset \mathbb{R}^n$  such that  $P \cup N \subseteq X$ . First observe that  $q_0(x, y) \equiv 0$  is an optimal solution to the Chebyshev approximation problem: the deviation  $f - q_0$  coincides with the function  $f$ , and we have for all  $x \in X$

$$\begin{aligned} f(x) &= \max_{u \in P} \varphi_u(x) - \max_{v \in N} \varphi_v(x) \\ &\leq \max_{u \in P} \varphi_u(x) \\ &= \max_{u \in P} \max \left\{ 1 - \frac{2}{d} \|x - u\|, 0 \right\} \\ &\leq \max_{u \in P} \max \left\{ 1 - \frac{2}{d} \min_{u \in P} \|x - u\|, 0 \right\} \\ &= \max \left\{ 1 - \frac{2}{d} \min_{u \in P} \|x - u\|, 0 \right\} \\ &= 1 - \frac{2}{d} \min \left\{ \min_{u \in P} \|x - u\|, \frac{d}{2} \right\} \leq 1; \end{aligned} \tag{8}$$

likewise

$$f(x) \geq -1 + \frac{2}{d} \min \left\{ \min_{v \in N} \|x - v\|, \frac{d}{2} \right\} \geq -1 \quad \forall x \in X. \tag{9}$$

Moreover, for  $x \in P$  we have  $f(x) = 1$ , for  $x \in N$  we have  $f(x) = -1$ , and it follows from (8) and (9) that for  $x \notin P \cap N$  we have  $-1 < f(x) < 1$ , hence,  $N = \mathcal{N}(q_0)$  and  $P = \mathcal{P}(q_0)$ , so  $q_0$  satisfies the very last statement of the lemma. We have assumed that  $N$  and  $P$  can not be strictly separated by a polynomial in  $V$ , hence we deduce that  $q_0 \equiv 0$  is a best Chebyshev approximation of  $f$  on  $X$ .

We will next show that for any direction  $p \in S$  such that  $p(N) \leq 0$  and  $p(P) \geq 0$  there exists a sufficiently small  $\alpha > 0$  such that  $\alpha p$  is another best Chebyshev approximation of  $f$  on  $X$ . Note that this guarantees that for any set of linearly independent

vectors in  $S$  we can produce a simplex with vertices at zero and at nonzero vectors along these linearly independent vectors. This yields  $\dim Q = \dim S$ .

Since  $p \in S$  is a polynomial, and the set  $X$  is compact,  $p$  is Lipschitz on  $X$  with some constant  $L$ , and its absolute value is bounded by some  $M > 0$  on  $X$ . Let  $\alpha := \min \left\{ \frac{1}{M}, \frac{2}{dL} \right\}$ , then for  $q = \alpha p$  we have

$$|q(x)| = |\alpha p(x)| \leq \alpha \cdot M \leq 1 \quad \forall x \in X,$$

$$|q(x) - q(y)| = |\alpha p(x) - \alpha p(y)| = \alpha |p(x) - p(y)| \leq \alpha L \|x - y\| \leq \frac{2}{d} \|x - y\| \quad \forall x, y \in X;$$

$$q(y) - \frac{2}{d} \|x - y\| \leq q(x) \leq q(y) + \frac{2}{d} \|x - y\| \quad \forall x, y \in X;$$

From  $q(N) \leq 0$  and  $p(P) \geq 0$  we have for all  $x \in X$

$$\frac{2}{d} \min_{y \in P} \|x - y\| \leq \max_{y \in P} \left( q(y) - \frac{2}{d} \|x - y\| \right) \leq q(x) \leq \min_{y \in N} \left( q(y) + \frac{2}{d} \|x - y\| \right) \leq \frac{2}{d} \min_{y \in N} \|x - y\|.$$

Hence,

$$\max \left\{ \frac{2}{d} \min_{y \in P} \|x - y\|, -1 \right\} \leq q(x) \leq \min \left\{ \frac{2}{d} \min_{y \in N} \|x - y\|, 1 \right\}.$$

We hence have for every  $x \in X$

$$-1 \leq f(x) - q(x) \leq 1,$$

therefore  $q$  is a best Chebyshev approximation of  $f$  on  $X$ .  $\square$

### 3.5 Uniqueness and the domain geometry

Finally, we turn our attention to the relation between the uniqueness of best Chebyshev approximation and the geometry of the domain. We show that on finite domains the best approximation is nonunique whenever the dimension of  $S$  allows for this (that is,  $\dim S > 0$ ).

**Lemma 3** *If  $X \subset \mathbb{R}^n$  is finite, then for any  $f : X \rightarrow \mathbb{R}$  we have  $\dim Q = \dim S$ .*

**Proof** If  $\dim S = 0$ , the result follows directly from Corollary 3. For the rest of the proof, assume  $\dim S > 0$ .

Let  $q \in \text{ri } Q$ ,  $s \in S$ . Then

$$s(x) \cdot s(y) \leq 0 \quad \forall x \in \mathcal{P}, y \in \mathcal{N}.$$

Let

$$q_t := q + ts.$$

Without loss of generality, assume that  $s(x) \geq 0$  for  $x \in \mathcal{P}$  and  $s(x) \leq 0$  for  $x \in \mathcal{N}$  (otherwise consider  $-s$ ).



Let

$$\alpha := \|f - q\|_\infty - \max_{x \in X \setminus (\mathcal{N} \cup \mathcal{P})} |f(x) - q(x)|,$$

where we use the standard convention that the maximum over an empty set equals  $-\infty$ , so  $\alpha = +\infty$  in the case when  $X = \mathcal{N} \cup \mathcal{P}$ . Since  $X$  is finite,  $\alpha > 0$ .

Let

$$\beta := \max_{x \in X} |s(x)|.$$

We have for all  $t \geq 0$  and  $x \in \mathcal{N}$

$$\|f - q\|_\infty = q(x) - f(x) \geq q(x) - f(x) + ts(x) \geq q(x) - f(x) - t\beta = \|f - q\|_\infty - t\beta;$$

for  $t \geq 0$  and  $x \in \mathcal{P}$

$$\|f - q\|_\infty = f(x) - q(x) \geq f(x) - q(x) - ts(x) \geq f(x) - q(x) - t\beta = \|f - q\|_\infty - t\beta;$$

So, whenever  $t\beta \leq \|f - q\|_\infty$ , we find that for  $x \in \mathcal{N} \cup \mathcal{P}$ ,

$$|f(x) - q_t(x)| \leq \|f - q\|_\infty.$$

For  $x \in X \setminus (\mathcal{N} \cup \mathcal{P})$  and all  $t \geq 0$

$$|f(x) - q_t(x)| \leq |f(x) - q(x)| + t|s(x)| \leq \|f - q\|_\infty - \alpha + t\beta.$$

Note that  $\alpha = +\infty$  only for the case when  $X = \mathcal{N} \cup \mathcal{P}$ .

Therefore, for  $t$  such that  $t\beta \leq \min\{\alpha, \|f - q\|_\infty\}$  we have

$$\|q_t - f\|_\infty \leq \|f - q\|_\infty,$$

and hence  $q_t \in Q$  for some positive  $t$ .

It remains to pick a maximal linearly independent system  $\{s_1, s_2, \dots, s_d\} \subset S$ , and observe that the convex hull  $\text{co}\{q, q + t_1s_1, \dots, q + t_d s_d\} \subseteq Q$  for some nonzero  $t_1, \dots, t_d$ . Therefore,  $\dim Q \geq \dim S$ . By Lemma 1 the converse is true, and we are done. □

It follows from the previous lemma that the uniqueness of solutions depends not only on the function itself, but also on the domain of its definition. In particular, it may happen that a function defined on a continuous domain has a unique best approximation, but a discretisation of this domain would lead to nonuniqueness of best approximation. This observation is crucial, since most numerical methods work on finite grids rather than with continuous functions directly. Therefore, they do require a certain level of discretisation. In this case there is a potential danger of finding an optimal solution to the discretised problem, while it is not relevant to the original one.

## 4 Conclusions

We have identified and discussed in detail key structural properties pertaining to the solution set of the multivariate Chebyshev approximation problem. We have clarified the relations between the points of maximal and minimal deviation for different optimal solutions, related the set of optimal solutions to the set of separating polynomials, and elucidated the relations between the geometry of the domain and smoothness of the function and uniqueness of the solutions.

However many questions remain unanswered, some of them pertinent to the potential algorithmic solutions, and more remains to be done to fully understand the relation between the uniqueness of the solutions and structure of the problem. Namely, the following questions are of paramount importance.

1. Can we refine Mairhuber's theorem for the case of multivariate Chebyshev approximation by polynomials of degree at most  $d$ ? Example 2 indicates that to have a unique approximation of any continuous function on a given domain by a system of multivariate polynomials, it may not be enough to restrict the domain to a set homeomorphic to a subset of a circle. Perhaps a more algebraic condition would work, for instance, restricting the domain to sets with one-dimensional Zariski closure.
2. What are the sufficient conditions for the uniqueness of the best Chebyshev approximation in terms of the function  $f$  only? Can we guarantee that for a given set of points of maximal and minimal deviation there exists a domain  $X$  that contains them and a function  $f$  for which an optimal solution is unique and has specifically this distribution of points of minimal and maximal deviation?
3. Can we bridge the gap between Lemmas 1 and 3 and show that given a distribution of points of minimal and maximal deviation, for any  $d \in \{0, \dots, \dim S\}$  there exists a function  $f$  and domain  $X$  with  $\dim Q = d$ ? This question is closely related to our discussion at the end of Example 4, where smoothness appears to be important only with relation to the orthogonal direction to the varieties separating the points of maximal and minimal deviation.

## Appendix

### Technical computations for example 2

Consider the polynomial  $q_\alpha(x, y) = 3\alpha(x^2 + y^2 - 1) + 1$ ,  $\alpha \in [0, 1]$ , of which the polynomials  $q_0$  and  $q_1$  are special cases. Explicitly our deviation  $d_\alpha(x, y) = f(x, y) - q_\alpha(x, y)$  has the form

$$d_\alpha(x, y) = x^6 + y^6 + 3x^4y^2 + 3x^2y^4 + 6xy^2 - 2x^3 - 1 - 3\alpha(x^2 + y^2 - 1).$$

The points of maximal and minimal deviation are the global extrema of  $d_\alpha$  on the unit disk. To obtain all such extrema, we first find the global minima and maxima of  $d_\alpha$  on the boundary of the disk, using the method of Lagrange multipliers, and then study the behaviour of  $d_\alpha$  on the interior of the disk.

Our Lagrangian function is  $L_\alpha(x, y, \lambda) = d_\alpha(x, y) + 6\lambda(x^2 + y^2 - 1)$  (where we have multiplied the constraint by 6 for convenience), and the necessary condition for the constrained global stationary points on the unit circle is

$$\nabla L_\alpha(x, y, \lambda) = 6 \begin{bmatrix} (y^2 - x^2) + x[(x^2 + y^2)^2 - \alpha + 2\lambda] \\ 2xy + y[(x^2 + y^2)^2 - \alpha + 2\lambda] \\ x^2 + y^2 - 1 \end{bmatrix} = 0_3.$$

Multiplying the first line by  $y$ , and the second line by  $x$ , and subtracting, we obtain the consequence of the first two equations in the Lagrangian system:  $y(y^2 - 3x^2) = 0$ . Together with the constraint  $x^2 = 1 - y^2$  this yields six candidates for the stationary points on the boundary,

$$z_1 = (1, 0), \quad z_2 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad z_3 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right),$$

$$z_4 = (-1, 0), \quad z_5 = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right), \quad z_6 = \left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right).$$

It is not difficult to check that

$$d_\alpha(z_1) = d_\alpha(z_3) = d_\alpha(z_5) = -2, \quad d_\alpha(z_2) = d_\alpha(z_4) = d_\alpha(z_6) = 2.$$

Note that these values do not depend on  $\alpha$ .

It remains to study the behaviour of the deviation  $d_\alpha$  on the interior of the disk. If  $d_\alpha$  attains a global minimum or maximum in an interior point of the disk, then such extrema must satisfy the unconstrained optimality condition  $\nabla d_\alpha(x, y) = 0_2$ . We have explicitly

$$\nabla d_\alpha(x, y) = 6 \begin{bmatrix} (y^2 - x^2) + x[(x^2 + y^2)^2 - \alpha] \\ 2xy + y[(x^2 + y^2)^2 - \alpha] \end{bmatrix} = 0_2.$$

As before, premultiplying the equations by  $y$  and  $x$  and subtracting, we conclude that any stationary point must satisfy the equality  $y(y^2 - 3x^2) = 0$ . Hence any maximum or minimum must lie on one of the lines

$$y = 0, \quad y = -\sqrt{3}x, \quad y = \sqrt{3}x.$$

Observe that both our polynomial and the constraint are symmetric with respect to the rotation of the plane by  $2\pi/3$ , the restrictions of the polynomial  $d_\alpha$  to each of those lines are identical (under the relevant rotations), hence it is sufficient for us to study the behaviour of the restriction of  $d_\alpha$  to the open line segment  $(-1, 1) \times \{0\}$ . For convenience, we let

$$\varphi_\alpha(x) := d_\alpha(x, y) = x^6 - 2x^3 - 1 - 3\alpha(x^2 - 1).$$

Observe that

$$\varphi'_\alpha(x) = 6x^5 - 6x^2 - 6\alpha x = 6x(x(x^3 - 1) - 1) < 0 \quad \forall x \in (0, 1),$$

hence  $\varphi_\alpha(x)$  is strictly decreasing on  $(0, 1)$ , and can only have minima or maxima on the endpoints of  $[0, 1]$ . For the open line segment  $(-1, 0)$  and  $\alpha \in [0, 1]$  we have

$$\varphi_\alpha(x) = (x^3 - 1)^2 + 3\alpha(1 - x^2) - 2 < 1 + 3 - 2 = 2 \quad \forall x \in (-1, 0);$$

likewise

$$\varphi_\alpha(x) = (x^3 - 1)^2 + 3\alpha(1 - x^2) - 2 > 0 + 0 - 2 = -2 \quad \forall x \in (-1, 0).$$

Since  $\varphi(1) = -2$ , and  $\varphi(-1) = 2$ , this means that no global minimum or maximum can be achieved on  $(-1, 1)$ . We are hence left with the only candidate  $x = 0$ , for which we have

$$\varphi_\alpha(0) = -1 + 3\alpha \in [-1, 2) \quad \text{for } \alpha \in [0, 1),$$

and  $\varphi_1(0) = -2$ . This yields the distribution of points of minimal and maximal deviation of  $f$  from  $q_0$  and  $q_1$  as described in Example 2.

#### Computations for example 4

To find the points of maximal and minimal deviation of  $f(x, y) = (x^2 - \frac{1}{2})(1 - y^2)$  from the constant polynomial  $q_0(x, y) \equiv 0$  on the square  $[-1, 1] \times [-1, 1]$ , observe that the optimality condition on the interior of the square gives

$$\nabla f(x, y) = \begin{pmatrix} 2x(1 - y^2) \\ -y(x^2 - \frac{1}{2}) \end{pmatrix} = 0_2,$$

and out of the five solutions to  $\nabla f(x, y) = 0$

$$(0, 0), \left(-\frac{1}{\sqrt{2}}, -1\right), \left(\frac{1}{\sqrt{2}}, -1\right), \left(-\frac{1}{\sqrt{2}}, 1\right), \left(\frac{1}{\sqrt{2}}, 1\right)$$

only  $(0, 0)$  is in the interior of the square. Hence we have only one stationary point  $(0, 0)$  within the interior of the square, with deviation  $d_0(0, 0) = f(0, 0) - q_0(0, 0) = f(0, 0) = -\frac{1}{2}$ .

We now study the boundary of the square: restricting to  $x = \pm 1$ , and  $y \in [-1, 1]$ , we have the function  $\frac{1}{2}(1 - y^2)$ , which attains minima at the endpoints of the sides of the square, at  $(\pm 1, \pm 1)$  with deviation  $d_0(\pm 1, \pm 1) = 0$ , and maxima at  $(\pm 1, 0)$ , with the value  $d_0(\pm 1, 0) = f(\pm 1, 0) = \frac{1}{2}$ . For  $y = \pm 1$  the function is identically zero. We conclude that the points of maximal and minimal deviation of  $f$  from zero, on the square  $X = [-1, 1] \times [-1, 1]$ , are

$$\mathcal{P}(q) = \{(0, 0)\}, \quad \mathcal{N}(q) = \{(-1, 0), (1, 0)\}.$$

We next study the deviation of the function  $h(x, y) = (x^2 - \frac{1}{2})(1 - |y|)$  from polynomials  $q_\alpha(x, y) = \alpha y$  for  $\alpha \in [-\frac{1}{2}, \frac{1}{2}]$ . First of all, observe that for  $y = 0$  we have

$$d_\alpha(x, y) = h(x, y) - q_\alpha(x, y) = x^2 - \frac{1}{2},$$

and hence  $d_\alpha(x, 0)$  is minimal at  $(0, 0)$  with the value  $d_\alpha(0, 0) = -\frac{1}{2}$ , and maximal at  $(\pm 1, 0)$  with the value  $d_\alpha(\pm 1, 0) = \frac{1}{2}$ , independent on  $\alpha$ .

For  $y > 0$  we have  $d_\alpha(x, y) = h(x, y) - q_\alpha(x, y) = (x^2 - \frac{1}{2})(1 - y) - \alpha y$ , hence the unconstrained optimality condition gives

$$\nabla d_\alpha(x, y) = \begin{pmatrix} 2x(1 - y) \\ -x^2 + \frac{1}{2} - \alpha \end{pmatrix} = 0_2,$$

and the only case when we have solutions in the intersection of the interior of the square and  $y > 0$  is when  $\alpha = \frac{1}{2}$ ; likewise,  $\nabla d_\alpha(x, y) = 0$  gives no solutions in the interior of the square intersected with  $y < 0$  except for  $\alpha = -\frac{1}{2}$ . In both cases we have

$$\begin{aligned} d_{\frac{1}{2}}(0, y) &= -\frac{1}{2}(1 - y) - \frac{1}{2}y = -\frac{1}{2}, \quad y > 0; \\ d_{-\frac{1}{2}}(0, y) &= -\frac{1}{2}(1 + y) + \frac{1}{2}y = -\frac{1}{2}, \quad y < 0. \end{aligned}$$

For the sides of the square that correspond to  $x = \pm 1$ , and  $y \in [-1, 1]$ , we have a piecewise linear function

$$d_\alpha(\pm 1, y) = \frac{1}{2}(1 - |y|) - \alpha y = \begin{cases} \frac{1}{2} - (\alpha + \frac{1}{2})y, & y \geq 0, \\ \frac{1}{2} - (\alpha - \frac{1}{2})y, & y < 0, \end{cases}$$

hence its behaviour is completely determined by the endpoints of the relevant segments:  $(\pm 1, \pm 1), (\pm 1, 0)$ . We have

$$d_\alpha(\pm 1, -1) = \alpha; \quad d_\alpha(\pm 1, 1) = -\alpha; \quad d_\alpha(\pm 1, 0) = \frac{1}{2}. \tag{10}$$

For the remaining case of the interior of the sides,  $(-1, 1) \times \{\pm 1\}$  we have

$$d_\alpha(x, -1) = \alpha; \quad d_\alpha(x, 1) = -\alpha. \tag{11}$$

Observe that for  $\alpha = 0$  the only points of maximal and minimal deviation lie on the line  $y = 0$ , and hence the polynomial  $q_0(x, y) = 0$  is a best approximation of the function  $h$  on the square  $X$ . Also note that for  $|\alpha| > \frac{1}{2}$  the relations (11) give worse values of minimal and maximal deviation, hence,  $q_\alpha$  can not be a best approximation for  $|\alpha| > \frac{1}{2}$ . For  $|\alpha| \in (0, \frac{1}{2})$  we observe that there are no additional points of minimal and maximal deviation on top of the three alternating points on  $y = 0$  that are present for  $\alpha = 0$ . It remains to consider the values  $|\alpha| = \frac{1}{2}$ .

For  $\alpha = -\frac{1}{2}$  we have from (10) and the piecewise linear observation

$$d_{-\frac{1}{2}}(\pm 1, y) = \frac{1}{2} \quad \forall y \in [0, 1],$$

and (11) gives

$$d_{-\frac{1}{2}}(x, -1) = -\frac{1}{2}; \quad d_{-\frac{1}{2}}(x, 1) = \frac{1}{2}.$$

Likewise, for  $\alpha = \frac{1}{2}$  we obtain

$$d_{\frac{1}{2}}(\pm 1, y) = \frac{1}{2} \quad \forall y \in [-1, 0], \quad d_{\frac{1}{2}}(x, -1) = \frac{1}{2}; \quad d_{\frac{1}{2}}(x, 1) = -\frac{1}{2}.$$

**Acknowledgements** We are grateful to the Australian Research Council for supporting this work via Discovery Project DP180100602.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions.

**Data availability** Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. Achieser. Theory of Approximation. Frederick Ungar, New York, (1965)
2. Aghili, A., Sukhorukova, N., Ugon, J.: Bivariate rational approximations of the general temperature integral. *J. Math. Chem.* **59**, 2049–2062 (2021)
3. Barrodale, I., Powell, M., Roberts, F.: The differential correction algorithm for rational  $l_\infty$ -approximation. *SIAM J. Numer. Anal.* **9**(3), 493–504 (1972)
4. Borwein, P.B., Pritsker, I.E.: The multivariate integer Chebyshev problem. *Constr. Approx.* **30**(2), 299–310 (2009)
5. Brown, A.L.: An extension to Mairhuber's theorem. On metric projections and discontinuity of multivariate best uniform approximation. *J. Approx. Theory* **36**(2), 156–172 (1982)
6. Buhmann, M.D.: Radial basis functions: theory and implementations. Cambridge Monographs on Applied and Computational Mathematics, vol. 12. Cambridge University Press, Cambridge (2003)
7. Chebyshev, P.: Théorie des mécanismes connus sous le nom de parallélogrammes. Mémoires des Savants étrangers présentés à l'Académie de Saint-Petersbourg **7**, 539–586 (1854)
8. Demyanov, V., Malozemov, V.: Optimality conditions in terms of alternance: two approaches. *J. Optim. Theory Appl.* **162**, 805–820 (2014)
9. Hiriart-Urruty, J.B., Lemaréchal, C.: Fundamentals of convex analysis. Grundlehren Text Editions. Springer-Verlag, Berlin (2001)
10. Mairhuber, J.C.: On Haar's theorem concerning Chebychev approximation problems having unique solutions. *Proc. Amer. Math. Soc.* **7**, 609–615 (1956)

11. Munthe-Kaas, H.Z., Nome, M., Ryland, B.N.: Through the kaleidoscope: symmetries, groups and Chebyshev-approximations from a computational point of view. In: Foundations of computational mathematics, Budapest 2011, vol 403 of London Math. Soc. Lecture Note Ser. pp 188–229. Cambridge Univ. Press, Cambridge, (2013)
12. Nakatsukasa, Y., Sete, O., Trefethen, L.N.: The aaa algorithm for rational approximation. *SIAM J. Sci. Comput.* **40**(3), A1494–A1522 (2018)
13. Reimer, M.: On multivariate polynomials of least deviation from zero on the unit ball. *Math. Z.* **153**(1), 51–58 (1977)
14. Remes, E.: Sur une propriété extrême des polynômes de tchebychef. *Comm. Inst. Sci. math. mec. Univ. Kharkoff et de la Soc. Math. de Kharkoff (Zapiski Nauchno-issledovatel'skogo instituta matematiki i mekhaniki i Khar'kovskogo matematicheskogo obshchestva)*, 13:93–95, (1936)
15. Rice, J.R.: Tchebycheff approximation in several variables. *Trans. Amer. Math. Soc.* **109**, 444–466 (1963)
16. Ryland, B.N., Munthe-Kaas, H.Z.: On multivariate Chebyshev polynomials and spectral approximations on triangles. In: *Spectral and high order methods for partial differential equations*, vol 76 of Lect. Notes Comput. Sci. Eng. pp 19–41. Springer, Heidelberg, (2011)
17. Sloss, J.M.: Chebyshev approximation to zero. *Pacific J. Math.* **15**, 305–313 (1965)
18. Sukhorukova, N., Ugon, J.: A generalisation of de la Vallée-Poussin procedure to multivariate approximations. *Adv. Comput. Math.* **48**(1), 5 (2022)
19. Sukhorukova, N., Ugon, J., Yost, D.: Chebyshev multivariate polynomial approximation: alternance interpretation. In: *2016 MATRIX annals*, vol 1 of MATRIX Book Ser., pp 177–182. Springer, Cham, (2018)
20. Sukhorukova, N., Ugon, J., Yost, D.: Chebyshev multivariate polynomial approximation and point reduction procedure. *Constr. Approx.* **53**, 529–544 (2021)
21. Thiran, J.-P., Demaille, C.: On real and complex-valued bivariate Chebyshev polynomials. *J. Approx. Theory* **59**(3), 321–337 (1989)
22. Xu, Y.: Best approximation of monomials in several variables. *Rend. Circ. Mat. Palermo* **76**, 129–153 (2005)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.