# LEAST MULTIVARIATE CHEBYSHEV POLYNOMIALS ON DIAGONALLY DETERMINED DOMAINS

M. DRESSLER, S. FOUCART, M. JOLDES, E. DE KLERK, J. B. LASSERRE, AND Y. XU

ABSTRACT. We consider a new multivariate generalization of the classical monic (univariate) Chebyshev polynomial that minimizes the uniform norm on the interval $[-1, 1]$. Let $\Pi_n^*$ be the subset of polynomials of degree at most $n$ in $d$ variables, whose homogeneous part of degree $n$ has coefficients summing up to 1. The problem is determining a polynomial in $\Pi_n^*$ with the smallest uniform norm on a domain $\Omega$, which we call a least Chebyshev polynomial (associated with $\Omega$). Our main result solves the problem for $\Omega$ belonging to a non-trivial class of domains, defined by a property of its diagonal, and establishes the remarkable result that a least Chebyshev polynomial can be given via the classical, univariate, Chebyshev polynomial. In particular, the solution can be independent of the dimension. The result is valid for fairly general domains that can be non-convex and highly irregular.

## 1. Introduction

Among its numerous properties, the Chebyshev polynomial $T_n(x) = \cos(n \arccos x)$ provides a solution for the best approximation to the monomial $x^n$ on the interval $[-1, 1]$ in the uniform norm. More precisely, the polynomial

$$q_n^*(x) = x^n - 2^{1-n} T_n(x)$$

of degree $n - 1$ is the best polynomial of approximation to $x^n$ on $[-1, 1]$; that is

$$q_n^* = \arg \min_{q \in \Pi_{n-1}} \sup_{x \in [-1,1]} |x^n - q(x)|, \tag{1.1}$$

with $\Pi_{n-1}$ being the vector space of univariate polynomials of degree at most $n - 1$. In other words, the monic Chebyshev polynomial $x^n - q_n^*$ is the *least polynomial* in the sense that it has the least uniform norm among all monic polynomials of degree $n$.

There have been multiple extensions of Chebyshev polynomials to multivariate settings from different angles. From the point of view of approximation, an immediate generalization is finding the best approximation to monomials [1, 2, 3, 4, 7, 6, 8, 12, 13, 14]. Namely, for $d > 1$, $\alpha \in \mathbb{N}^d$ and $|\alpha| \coloneqq \sum_{i=1}^d \alpha_i = n$, we consider the problem

$$\min_{q \in \Pi_{n-1}^d} \sup_{\mathbf{x} \in \Omega} |\mathbf{x}^\alpha - q(\mathbf{x})|, \tag{1.2}$$

where $\Pi_{n-1}^d$ denotes the real polynomials of total degree at most $n-1$ in the variables $\mathbf{x} = (x_1, \ldots, x_d)$, and where we define the monomial $\mathbf{x}^\alpha \coloneqq x_1^{\alpha_1} \ldots x_d^{\alpha_d}$. Since we will deal with monomials of degree $n$ throughout the paper, we introduce the notation

$$\mathbb{N}_n^d \coloneqq \left\{ \alpha \in (\mathbb{N}_0)^d \ : \ |\alpha| = n \right\},$$

2020 *Mathematics Subject Classification.* Primary: 41A10, 41A63.

*Key words and phrases.* Least polynomial, Chebyshev polynomial, best approximation, extremal signature.

where $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ denotes the nonnegative integers. Thus $\mathbb{N}_n^d$ indexes all monomials of degree *exactly* $n$ in $d$ variables.

Problem (1.2) can be regarded as a natural multivariate generalization of (1.1), where $\Omega \subset \mathbb{R}^d$ is a subset of $\mathbb{R}^d$. While the interval $[-1, 1]$ is a prototype of a compact connected set of the real line, there is no 'prototype' in higher dimensions for such a set. In the literature, this problem has been studied primarily on a few special regular domains. In two variables, the problem (1.2) is solved for the square, the disk, and the isosceles right triangle [3, 4, 6, 8, 12]. While the solution on the square can be extended to the cube for $d > 2$, the problem is solved only for a few cases, mostly monomials of lower degrees, on the ball and the simplex [1, 2, 13]. Moreover, the existing examples indicate an increasing complexity, so much so that it does not appear possible to find an analytic solution even for these regular domains.

Recently, in [5] we have proposed to investigate (1.2) for various choices of $\Omega \subset \mathbb{R}^d$ and $\alpha \in \mathbb{N}_n^d$ by combining analytical tools with numerical tools from optimization (and notably the so-called moment-SOS hierarchy). During this study, we have encountered an optimization problem that has initiated a change of view: Namely, instead of studying the best polynomial of approximation to monomials, we can study the least polynomial instead. While the two concepts are identical in one variable, they can be quite different in higher dimensions, as seen from the definition below.

**Definition 1.1.** *Let $\Pi_n^d$ denote the space of polynomials of total degree at most $n$ in $d$ variables, and $\Pi_n^*$ the subset of $\Pi_n^d$ that consists of polynomials of the form*

$$\mathbf{x} \mapsto P(\mathbf{x}) := \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha \, \mathbf{x}^\alpha + Q(\mathbf{x}) \quad \textit{with} \qquad \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha = 1 \textit{ and } \; Q \in \Pi_{n-1}^d.$$

*Let $\Omega$ be a domain in $\mathbb{R}^d$. We consider the optimization problem*

$$\inf_{P \in \Pi_n^*} \|P\|_\Omega, \quad \textit{where} \quad \|P\|_\Omega := \sup_{\mathbf{x} \in \Omega} |P(\mathbf{x})| \,. \tag{1.3}$$

*If it exists, we call a minimizer $P^* \in \Pi_n^*$ of (1.3) a* least Chebyshev polynomial *of degree $n$ on the domain $\Omega$.*

For $d = 1$, there is only one monomial of degree $n$. In the case $d > 1$, every element of $\Pi_n^*$ is 'monic' and the monomial $\mathbf{x}^\alpha$ in (1.1) is only one among many possible choices in $\Pi_n^*$. However, rather than approximating a fixed monomial by polynomials of lower degree, the problem (1.3) requires finding a polynomial that has the least norm among all polynomials in $\Pi_n^*$. As far as we are aware, this polynomial has not been considered in the literature.

**Contribution.** The main purpose of this paper is to report our findings on the optimization problem (1.3). It turns out, much to our surprise, that the problem (1.3) can be solved analytically for a fairly general family of domains $\Omega$ in $\mathbb{R}^d$ for all $d \geq 2$. This family of domains will be referred to as *diagonally determined*.

**Organization of this paper.** The paper is organized as follows. We start by defining and describing several examples of diagonally determined domains in Section 2. Our main results are presented in Section 3, where we describe the least polynomial for a diagonally determined domain. We discuss the dual problem of (1.3) in Section 4 and show that it too has a closed-form solution in the case of diagonally determined sets. We also rephrase this result for the dual problem in the framework of *extremal signatures*.

## 2. Diagonally determined domains

We define the *diagonal* of a domain $\Omega \subset \mathbb{R}^d$ as the set $\operatorname{diag}(\Omega) := \{t \in \mathbb{R} : t\mathbf{1} \in \Omega\}$, where $\mathbf{1}$ denotes the all-ones vector in $\mathbb{R}^d$. For two vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, we use $\langle \mathbf{x}, \mathbf{y} \rangle$ for their standard inner product.

**Definition 2.1.** *We call a set $\Omega \subset \mathbb{R}^d$* diagonally determined *if the following two conditions hold:*

*(1) The diagonal of $\Omega$ is an interval, say $\operatorname{diag}(\Omega) = [a, b]$, and*
*(2) there exists $\mathbf{v} \in \mathbb{R}^d$ such that $\langle \mathbf{v}, \mathbf{1} \rangle = 1$ and $\langle \mathbf{v}, \mathbf{x} \rangle \in [a, b]$ for all $\mathbf{x} \in \Omega$.*

Importantly, Definition 2.1 covers domains that could be non-convex, non-compact, and even highly irregular. Figure 1 gives an example of a non-convex diagonally determined set in $\mathbb{R}^2$.

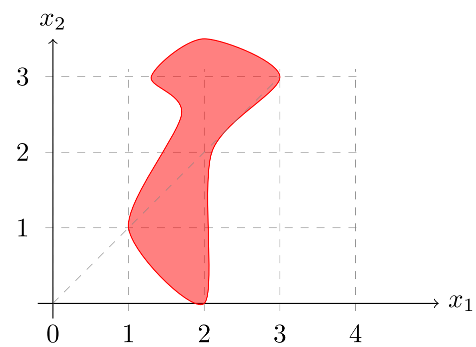


Figure 1. Example of a diagonally determined, non-convex set. Here, $\operatorname{diag}(\Omega) = [1, 3]$, and $\mathbf{v} = (1, 0)$.

As a first observation, a dilation of a diagonally determined set is again diagonally determined.

**Lemma 2.2.** *If $\Omega \subset \mathbb{R}^d$ is diagonally determined with $\operatorname{diag}(\Omega) = [a, b]$ and vector $\mathbf{v}$, then, for any $r > 0$, the dilation*

$$r\Omega := \{r\mathbf{x} \ : \ \mathbf{x} \in \Omega\}$$

*is also diagonally determined with $\operatorname{diag}(r\Omega) = [ra, rb]$ and vector $\mathbf{v}$.*

*Proof.* The proof is an immediate consequence of Definition 2.1. □

We will now show that balls in $\mathbb{R}^d$ (in any norm) are examples of diagonally determined sets.

**Proposition 2.3.** *Consider a ball of radius $r$ in $\mathbb{R}^d$ centered at the origin:*

$$\Omega := \{\mathbf{x} \in \mathbb{R}^d \ : \ \|\mathbf{x}\| \leq r\},$$

*where $\|\cdot\|$ denotes any norm on $\mathbb{R}^d$. Then $\Omega$ is a diagonally determined set.*

*Proof.* We prove the statement for unit balls; the required result then follows from Lemma 2.2.

Let $\|\cdot\|_*$ denote the usual dual norm of $\|\cdot\|$, namely

$$\|\mathbf{u}\|_* = \sup\{\langle \mathbf{u}, \mathbf{x} \rangle \ : \ \|\mathbf{x}\| = 1\}, \quad \mathbf{u} \in \mathbb{R}^d.$$

Let $\mathbf{v} = \mathbf{u}/\|\mathbf{1}\|$, where $\mathbf{u} \in \mathbb{R}^d$ is 'dual' to $\mathbf{1}$, in the sense that $\|\mathbf{u}\|_* = 1$ and $\langle \mathbf{u}, \mathbf{1}\rangle = \|\mathbf{1}\|$. We then have $\mathrm{diag}(\Omega) = [-b, b]$ with $b = 1/\|\mathbf{1}\|$, $\langle \mathbf{u}/\|\mathbf{1}\|, \mathbf{1}\rangle = 1$, and $|\langle \mathbf{u}/\|\mathbf{1}\|, \mathbf{x}\rangle| \le (1/\|\mathbf{1}\|)\|\mathbf{u}\|_*\|\mathbf{x}\| \le 1/\|\mathbf{1}\| = b$, so that $\langle \mathbf{u}/\|\mathbf{1}\|, \mathbf{x}\rangle \in [-b, b] = [a, b]$ for all $\mathbf{x} \in \Omega$. □

By Definition 2.1, we have the following immediate corollary.

**Corollary 2.4.** *Assume $\Omega \subset \mathbb{R}^d$ is a subset of a ball in $\mathbb{R}^d$ in any norm, centered at the origin, and that the diagonal of $\Omega$ coincides with the diagonal of the ball. Then $\Omega$ is a diagonally determined set.*

To illustrate this corollary, the example in Figure 2 shows a non-convex subset of a unit (Euclidean) ball, that has the same diagonal as the ball.

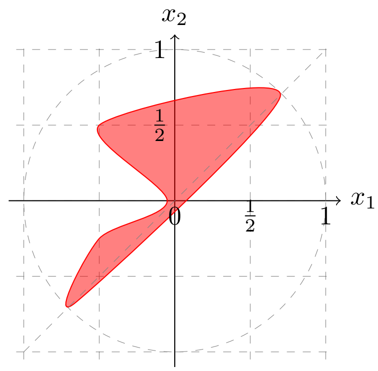


FIGURE 2. Example of a diagonally determined, non-convex set. Here, $\mathrm{diag}(\Omega) = \left[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right]$, and $\mathbf{v} = \left(\frac{1}{2}, \frac{1}{2}\right)$.

**Example 2.5.** *It is insightful to consider the specific example of the unit ball of the* owl norm *(ordered weighted $\ell_1$-norm) defined relative to weights $w_1 \ge w_2 \ge \cdots \ge w_d \ge 0$ by*

$$\|\mathbf{x}\|_{\mathsf{owl}} = \sum_{i=1}^{d} w_i x_i^*,$$

*with $(x_1^*, \ldots, x_d^*)$ being the nonincreasing rearrangement of $(|x_1|, \ldots, |x_d|)$. Note that $\|\mathbf{1}\|_{\mathsf{owl}} = \sum_{i=1}^d w_i =: W$, so that $b = 1/W$. Letting $\mathbf{w} = (w_1, \ldots, w_d)$, we set $\mathbf{v} = \mathbf{w}/W$, since $\langle \mathbf{w}/W, \mathbf{1}\rangle = (\sum_{i=1}^d w_i)/W = 1$ and, for $\|\mathbf{x}\|_{\mathsf{owl}} \le 1$,*

$$\begin{aligned}
|\langle \mathbf{w}/W, \mathbf{x}\rangle| &= \left|\sum_{i=1}^n w_i x_i\right| /W \\
&\le \left(\sum_{i=1}^n w_i |x_i|\right)/W \\
&\le \left(\sum_{i=1}^n w_i x_i^*\right)/W \\
&\le 1/W = b.
\end{aligned}$$

The intersection of certain balls in $\mathbb{R}^d$ with the nonnegative orthant in $\mathbb{R}^d$ are also diagonally determined, as the next result shows.

**Proposition 2.6.** *Assume $\|\cdot\|$ is a norm on $\mathbb{R}^d$ with dual norm $\|\cdot\|_*$, and that the following holds: there is an entrywise-nonnegative vector $\mathbf{u} \in \mathbb{R}^d_+$ such that $\|\mathbf{u}\|_* = 1$ and $\langle \mathbf{u}, \mathbf{1} \rangle = \|\mathbf{1}\|$.*

*Further assume $\Omega \subset \mathbb{R}^d$ has diagonal $\operatorname{diag}(\Omega) = [0, r/\|\mathbf{1}\|]$ for some $r > 0$, and, for all $\mathbf{x} \in \Omega$, $\mathbf{x} \geq \mathbf{0}$ and $\|\mathbf{x}\| \leq r$. Then $\Omega$ is a diagonally determined set with vector $\mathbf{v} = \mathbf{u}/\|\mathbf{1}\|$.*

*Proof.* We again prove the statement for the case $r = 1$; the result for general $r > 0$ then follows from Lemma 2.2. Setting $\mathbf{v} = \mathbf{u}/\|\mathbf{1}\|$, one has $\langle \mathbf{v}, \mathbf{1} \rangle = 1$. Moreover, for $\mathbf{x} \in \Omega$, $\langle \mathbf{v}, \mathbf{x} \rangle \geq 0$, since $\mathbf{x}, \mathbf{v} \geq 0$, and

$$\langle \mathbf{v}, \mathbf{x} \rangle = \frac{1}{\|\mathbf{1}\|} \langle \mathbf{u}, \mathbf{x} \rangle \leq \frac{1}{\|\mathbf{1}\|} \|\mathbf{u}\|_* \|\mathbf{x}\| \leq \frac{1}{\|\mathbf{1}\|},$$

so that $\langle \mathbf{v}, \mathbf{x} \rangle \in \operatorname{diag}(\Omega)$ for all $\mathbf{x} \in \Omega$. $\square$

The assumption on the norm in Proposition 2.6 is met by, for example, all $\ell^p$-norms, and the owl-norm in Example 2.5. As a simple corollary of the proposition, the simplex

$$\Omega = \left\{ \mathbf{x} \in \mathbb{R}^d \ : \ \mathbf{x} \geq \mathbf{0}, \ \sum_{i=1}^{d} x_i \leq 1 \right\},$$

is diagonally determined, since it is the intersection of the unit $\ell^1$-ball with the nonnegative orthant. See Figure 3 for an illustrative example of Proposition 2.6.

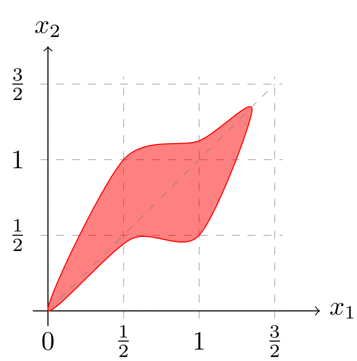


FIGURE 3. Example to illustrate Proposition 2.6. Here, $\operatorname{diag}(\Omega) = [0, b]$, with $b \approx 1.35$. Moreover, $\Omega$ is contained in the $\ell^1$-ball of radius $r = bd \approx 2.7$, intersected with the nonnegative quadrant.

## 3. Least Chebyshev polynomial for a diagonally determined domain

In this section, and with $\Omega$ being a domain in $\mathbb{R}^d$, we consider the problem of finding a least polynomial $P^* \in \Pi_n^*$ such that

$$\|P^*\|_\Omega = \inf \left\{ \|P\|_\Omega : P \in \Pi_n^* \right\},$$

for the special case where $\Omega$ is diagonally determined.

**Theorem 3.1.** *Let $\Omega$ be diagonally determined with vector $\mathbf{v}$ and $\operatorname{diag}(\Omega) = [a, b]$. Then*

$$\inf\{\|P\|_\Omega : P \in \Pi_n^*\} = \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}}, \tag{3.1}$$

*and the infimum is attained by the polynomial*

$$\mathbf{x} \mapsto P^*(\mathbf{x}) := \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}} T_n\left(-1 + 2\,\frac{\langle \mathbf{v}, \mathbf{x}\rangle - a}{b-a}\right). \tag{3.2}$$

*Proof.* Let $P \in \Pi_n^*$, say

$$P(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha \mathbf{x}^\alpha + Q(\mathbf{x}) \quad \text{with} \quad \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha = 1 \text{ and } \ Q \in \Pi_{n-1}^d.$$

Then $P(t, \ldots, t) = t^n + q_{n-1}(t)$, where $q_{n-1}(t) = Q(t, \ldots, t)$ is a univariate polynomial of degree at most $n-1$ in the single variable $t$. Hence,

$$\|P\|_\Omega \geq \max_{a \leq t \leq b} |t^n + q_{n-1}(t)| \geq \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}}, \tag{3.3}$$

where we have used a classical result in one variable, and equality in the last inequality is attained by choosing $q_{n-1}$ such that $t^n + q_{n-1}(t)$ is the rescaled Chebyshev polynomial in the right-hand side of (3.2).

Moreover, as $\Omega$ is diagonally determined, then for all $\mathbf{x} \in \Omega$, one has $\langle \mathbf{v}, \mathbf{x}\rangle \in [a, b]$, and therefore

$$\|P^*\|_\Omega = \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}} \max_{\mathbf{x} \in \Omega} \left| T_n\left(-1 + 2\frac{\langle \mathbf{v}, \mathbf{x}\rangle - a}{b-a}\right)\right| \leq \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}}.$$

Finally, we need to show that $P^* \in \Pi_n^*$. First note that $t \mapsto P^*(t\mathbf{1})$ is a monic univariate polynomial in $t$, since $\langle \mathbf{v}, \mathbf{1}\rangle = 1$. On the other hand, we may write $P^*$ in the form

$$P^*(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \mathbf{x}^\alpha + Q^*(\mathbf{x}) \ \text{ with } Q^* \in \Pi_{n-1}^d,$$

so that

$$P^*(t\mathbf{1}) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* t^n + Q^*(t\mathbf{1}).$$

Thus $\sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* = 1$, and therefore $P^* \in \Pi_n^*$, as required. □

The minimal value in (3.1) does not dependent explicitly on the dimension $d$, but the diagonal $[a, b]$ may. For example, if $\Omega$ is the unit Euclidean ball in $\mathbb{R}^d$, then $-a = b = 1/\sqrt{d}$.

**Example 3.2.** *For the cube $\Omega = [-1, 1]^d$, Theorem 3.1 shows that $\|P^*\|_\Omega = 2^{-n+1}$. In contrast, for every $\alpha \in \mathbb{N}^d$ with $|\alpha| = n$, we have*

$$\inf_{P \in \Pi_{n-1}^d} \|\mathbf{x}^\alpha - P\|_{[-1,1]^d} = 2^{-n+d},$$

*as shown in* [12]*, which depends on the dimension $d$.*

Since an optimal solution of problem (1.3) is given in terms of the univariate Chebyshev polynomial $T_n$ in Theorem 3.1, the reader may wonder if this optimal solution is in fact unique. This turns out to be not the case, as the next example shows.

**Example 3.3.** *The simplex* $\triangle^d = \{\, \mathbf{x} \in \mathbb{R}^d \ : \ 0 \le x_1 \le x_2 \le \ldots \le x_d \le 1 \,\}$ *is diagonally determined with* $\mathbf{v} = (1, 0, \ldots, 0)$*, and* $\operatorname{diag}(\triangle^d) = [0,1]$*. Hence, by Theorem 3.1,*

$$\|P^*\|_{\triangle^d} = \inf\left\{ \|P\|_{\triangle^d} : P \in \Pi_n^* \right\} = \left(\frac{1}{2}\right)^n \frac{1}{2^{n-1}} = 2^{-2n+1}$$

*is independent of the dimension* $d$*. Moreover, one has* $P^*(\mathbf{x}) = 2^{-2n+1} T_n(2x_1 - 1)$*, which is univariate. Another valid choice for* $\mathbf{v}$ *is* $\mathbf{v} = (0, 1, 0, \ldots, 0)$ *that leads to a different least Chebyshev polynomial, namely* $\mathbf{x} \mapsto 2^{-2n+1} T_n(2x_2 - 1)$*.*

We point out that the requirement $\sum_{\alpha \in \mathbb{N}_n^d} a_\alpha = 1$, is susceptive under linear change of variables other than a translation. For instance, the triangle $\hat{\triangle}^2 = \{\, \mathbf{x} : x_1 \ge 0, x_2 \ge 0, x_1 + x_2 \le 1 \,\}$, is a mirror image of $\triangle^2$ under $x_2 \mapsto 1 - x_2$. However, under the linear change of variables $x_2 \mapsto 1 - x_2$, the leading monomial $a_0 x_1^2 + a_1 x_1 x_2 + a_2 x_2^2$ of a polynomial in $\Pi_2^*$ becomes, $a_0 x_1^2 - a_1 x_2 x_2 + a_2 x_2^2$, which is no longer an element of $\Pi_2^*$.

## 4. The dual framework and signatures

In this section we consider the dual problem of (1.3). Our goal is to show that, in the special case of diagonally determined domains, the dual problem has a closed-form solution. Moreover, we will construct a dual solution that is atomic (discrete), and supported on $n+1$ points.

### 4.1. **The dual problem.**

With $\Omega$ being a domain in $\mathbb{R}^d$, let $C(\Omega)$ be the space of continuous functions on $\Omega$, and let $C(\Omega)^*$ be the dual space of $C(\Omega)$. For compact $\Omega$, one has the following, strong duality result for problem (1.3).

**Theorem 4.1.** *Let* $\Omega \subset \mathbb{R}^d$*, and consider the dual problem of* (1.3)*, namely*

$$\gamma^* := \sup_{\substack{\gamma \in \mathbb{R} \\ L \in C(\Omega)^*}} \left\{ \gamma : \ L_{|\Pi_{n-1}^d} = 0,\ \|L\|_{C(\Omega)^*} = 1, \text{and } L(\mathbf{x} \mapsto \mathbf{x}^\alpha) = \gamma \ \text{ for all } \alpha \in \mathbb{N}_n^d \right\}, \tag{4.1}$$

*where the norm of* $L$ *is defined as*

$$\|L\|_{C(\Omega)^*} = \sup_{h \in C(\Omega)} \{ |Lh| \ : \|h\|_\Omega \le 1 \}.$$

*One has* $\gamma^* \le \inf\{ \|P\|_\Omega \ : \ P \in \Pi_n^* \}$ *(weak duality), and* $\gamma^* = \inf\{ \|P\|_\Omega \ : \ P \in \Pi_n^* \}$ *if* $\Omega$ *is compact (strong duality).*

One may derive this result directly through conic linear programming duality theory, e.g. [11, Proposition 2.9], and we will omit the proof here.

It is insightful, though, to make a link with classical Chebyshev approximation. To this end, we denote an optimal solution of problem (1.3) by

$$\mathbf{x} \mapsto P^*(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \mathbf{x}^\alpha + Q^*(\mathbf{x}) \quad \text{such that} \quad \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* = 1, \qquad Q^* \in \Pi_{n-1}^d. \tag{4.2}$$

Note that $P^* \in \Pi_n^*$. Clearly,

$$\|P^*\|_\Omega = \min_{Q \in \Pi_{n-1}^d} \left\| \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \mathbf{x}^\alpha + Q(\mathbf{x}) \right\|_\Omega ,$$

while the latter problem is the classical Chebyshev problem of approximating the homogeneous polynomial $\mathbf{x} \mapsto \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \mathbf{x}^\alpha$ from $\Pi_{n-1}^d$.

Recall the strong duality result for the classical Chebyshev approximation problem as given in the paper by Rivlin and Shapiro [10, Corollary 2].[1]

**Theorem 4.2.** *For any $f \in C(\Omega)$, with $\Omega$ compact, one has the following*

$$\min_{Q \in \Pi_{n-1}^d} \|f - Q\|_\Omega = \max_{L \in C(\Omega)^*} \left\{ L(f) :\ L_{|\Pi_{n-1}^d} = 0,\ \|L\|_{C(\Omega)^*} = 1 \right\}. \tag{4.3}$$

From this, we can immediately deduce the following relation between the two dual problems.

**Lemma 4.3.** *For $f(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \mathbf{x}^\alpha$, any optimal solution of problem* (4.1) *is also optimal for the dual problem in* (4.3)*.*

*Proof.* Let $L^*$ denote an optimal solution of problem (4.1). Then $L^*$ is feasible for the dual (maximization) problem in (4.3). Moreover,

$$L^*(f) = L^* \left( \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \mathbf{x}^\alpha \right) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* L^*(\mathbf{x}^\alpha) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* \gamma^* = \gamma^* = \|P^*\|_\Omega.$$

This yields the result. ☐

In the next subsection, we review the fact that the optimal dual solutions may be assumed to be atomic (discrete) without loss of generality.

4.2. **Atomic dual solutions and signatures.** We first recall a classical interpolation formula for linear functionals, as given in [10, Corollary 3].

**Theorem 4.4.** *Let $L$ be any linear functional on a finite dimensional subspace $V$ of $C(\Omega)$, with $\Omega$ compact. Then there exist points $\omega_1, \ldots, \omega_r \in \Omega$ with $r \le \dim(V)$, and non-zero scalars $\tau_1, \ldots, \tau_r$, such that, defining the point evaluation functionals*

$$L_{\omega_i}(f) = f(\omega_i) \quad i = 1, \ldots, r,$$

*one has*

$$L = \sum_{i=1}^r \tau_i L_{\omega_i}, \quad \|L\| = \sum_{i=1}^r |\tau_i|. \tag{4.4}$$

As a consequence, there exists a discrete (atomic) solution to problem (4.1). We give a proof below only for the sake of completeness and later reference — the type of argument we use is classical.

**Corollary 4.5.** *There exist points $\omega_1, \ldots, \omega_r \in \Omega$ with $r \le \dim(\Pi_n^d)$, and non-zero scalars $\tau_1, \ldots, \tau_r$ with $\sum_{i=1}^r |\tau_i| = 1$, such that an optimal solution of problem* (4.1) *is given by* (4.4)*. Moreover, the points $\omega_1, \ldots, \omega_r$ are extremal points of any optimal solution $P^*$ to problem* (1.3)*, i.e.,*

$$|P^*(\omega_i)| = \|P^*\|_\Omega \quad \textit{for all } i = 1, \ldots, r.$$

[1] In [10, Corollary 2], the objective is in fact given as $\sup |L(f)|$, but the absolute value may be omitted w.l.o.g., since $L$ is feasible for the dual problem if and only if $-L$ is feasible.

*Proof.* The first statement follows immediately from Theorem 4.4, using $V = \Pi_n^d$. It remains to show that the points $\omega_1, \dots, \omega_r$ are extremal points of any optimal solution $P^*$ to problem (1.3). Now let $L^*$ denote an optimal solution of problem (4.1), so that

$$L^* = \sum_{i=1}^{r} \tau_i L_{\omega_i}, \quad \sum_{i=1}^{r} |\tau_i| = 1. \tag{4.5}$$

By Lemma 4.3, $L^*$ is also an optimal solution for the dual problem in (4.3). Thus, for any optimal solution $P^*$ of problem (1.3), one has

$$\|P^*\|_\Omega = L^*(P^*) = \sum_{i=1}^{r} \tau_i P^*(\omega_i).$$

Since $\sum_{i=1}^r |\tau_i| = 1$, $\|P^*\|_\Omega$ is a weighted average of the values $P^*(\omega_i)$ with $\tau_i > 0$ and $-P^*(\omega_i)$ with $\tau_i < 0$. As a consequence, we have

$$\|P^*\|_\Omega = \begin{cases} P^*(\omega_i) & \text{if } \tau_i > 0, \\ -P^*(\omega_i) & \text{if } \tau_i < 0, \end{cases}$$

completing the proof. □

The extremal points in Corollary 4.5 are usually called the *support of an extremal signature*, defined along the line of [9, Section 2.2] as follows.

**Definition 4.6.** *A signature with finite support $S \subset \Omega$ is simply a (partition) function from $S$ to $\{\pm 1\}$. A signature $\sigma$ with support $S$ is said to be extremal for a subspace $V \subset C(\Omega)$ if there exist weights $\lambda_\omega > 0$, $\omega \in S$, such that $\sum_{\omega \in S} \lambda_\omega \sigma(\omega) v(\omega) = 0$ for all $v \in V$. A signature $\sigma$ with support $S$ is said to be associated with a function $g \in C(\Omega)$ if $S$ is included in the set $\{\omega \in \Omega : |g(\omega)| = \|g\|_\Omega\}$ of extremal points of $g$ and if $\sigma(\omega) = \operatorname{sgn}(g(\omega))$ for all $\omega \in S$.*

Thus, the result of Corollary 4.5 may be restated as the existence of a signature with support $S = \{\omega_1, \dots, \omega_r\}$ where $r \leq \dim(\Pi_n^d)$. This signature is extremal for $\Pi_{n-1}^d$, and associated with every optimal solution of problem (1.3). Therefore, we will simply refer to an *optimal signature* for problem (1.3). Formally we have the following result.

**Proposition 4.7.** *Any atomic solution of the dual problem* (4.1)*, say $L^*$ of the form* (4.5)*, gives rise to an optimal signature for problem* (1.3)*, by setting:*

$$S = \{\omega_1, \dots, \omega_r\}, \quad \sigma(\omega_i) = \operatorname{sgn}(\tau_i), \quad \lambda_i = |\tau_i|, \qquad i = 1, \dots, r.$$

This is essentially a reformulation of the well-known characterization of best approximation by polynomials in terms of the extremal signature ([9, Theorem 2.6]) for the dual problem.

4.3. **Signatures for diagonally determined domains.** For the special case where $\Omega$ is a diagonally determined set, we may now infer information about an optimal signature from Theorem 3.1 and Proposition 4.7. The idea of the proof is to construct an atomic optimal solution for the dual problem (4.1).

**Theorem 4.8.** *Let $\Omega$ be a diagonally determined set with* $\operatorname{diag}(\Omega) = [a, b]$ *and vector* $\mathbf{v}$*. Then, an optimal signature for problem* (1.3) *is defined by a set $S$, of $n+1$ points*

$$\omega_j = a\mathbf{1} + \frac{b-a}{2}\left(1 + \cos\left(\frac{(j-1)\pi}{n}\right)\right)\mathbf{1} \quad \textit{for } j = 1, \dots, n+1, \tag{4.6}$$

*where* $\mathbf{1}$ *again denotes the all-ones vector in* $\mathbb{R}^d$*, together with the partition function* $\sigma : S \to \{-1, 1\}$*, given by*

$$\sigma(\omega_j) = \begin{cases} 1 & \text{if } j \text{ is odd} \\ -1 & \text{if } j \text{ is even,} \end{cases}$$

*as well as* $\lambda_1 = \lambda_{n+1} = \frac{1}{2}$*, and* $\lambda_i = 1$ *if* $i = 2, \dots, n$*.*

*Proof.* By Theorem 3.1, an optimal solution to problem (1.3) is given by the polynomial $P^*$ in (3.2), namely

$$P^*(\mathbf{x}) = \left(\frac{b-a}{2}\right)^n \frac{1}{2^{n-1}} T_n \left(-1 + 2\,\frac{\langle \mathbf{v}, \mathbf{x}\rangle - a}{b-a}\right).$$

Using the well-known fact that the extremal points of $T_n$ are given by the Gauss-Lobatto-Chebyshev points $\xi_j := \cos\left((j-1)\pi/n\right)$ for $j = 1, \dots, n+1$, we have that the points listed in (4.6) are extremal points of $P^*$. Indeed, using $\langle \mathbf{v}, \mathbf{1}\rangle = 1$, one has

$$\langle \mathbf{v}, \omega_j \rangle = a + \frac{b-a}{2}\left(1 + \cos\left(\frac{(j-1)\pi}{n}\right)\right) \qquad \text{for } j = 1, \dots, n+1,$$

which is the same as

$$-1 + 2\,\frac{\langle \mathbf{v}, \omega_j\rangle - a}{b-a} = \xi_j \text{ for } j = 1, \dots, n+1.$$

Next, we will construct an atomic solution of the dual problem (4.1), that will lead to the required optimal signature, by Proposition 4.7.

To this end, define the linear operator

$$L^* = \frac{1}{n+1} \sum_{i=1}^{n+1} (-1)^{i+1} \lambda_i L_{\omega_i}.$$

We claim that $L^*$ is an optimal solution of the dual problem (4.1). We first verify feasibility. We will show that $L^*$ vanishes on $\Pi^d_{n-1}$, as required, by using the Gauss-Lobatto-Chebyshev quadrature formula, which implies, for $k < n$:

$$\sum_{j=1}^{n+1} \lambda_j T_n(\xi_j) T_k(\xi_j) = 0,$$

due to the orthogonality of Chebyshev polynomials. Using $T_n(\xi_j) = (-1)^{1+j}$ yields

$$L^*(T_k) = \sum_{j=1}^{n+1} \lambda_j (-1)^{1+j} T_k(\xi_j) = 0 \text{ if } k < n. \tag{4.7}$$

Thus $L^*$ vanishes on univariate (and separable) polynomials in $\Pi^d_{n-1}$. For the general case, it suffices to consider $d = 2$ and the monomial $p(x_1, x_2) = T_{k_1}(x_1) T_{k_2}(x_2)$, where

$k_1 + k_2 < n$, and show that $L^*(p) = 0$. To this end, note that

$$\begin{aligned} L^*(p) &= \sum_{j=1}^{n+1} \lambda_j (-1)^{1+j} T_{k_1}(\xi_j) T_{k_2}(\xi_j) \\ &= \sum_{j=1}^{n+1} \lambda_j (-1)^{1+j} \frac{1}{2} [T_{|k_1-k_2|}(\xi_j) + T_{k_1+k_2}(\xi_j)] \\ &= \frac{1}{2} \sum_{j=1}^{n+1} \lambda_j (-1)^{1+j} T_{|k_1-k_2|}(\xi_j) + \frac{1}{2} \sum_{j=1}^{n+1} \lambda_j (-1)^{1+j} T_{k_1+k_2}(\xi_j) = 0, \end{aligned}$$

using (4.7). The result for general $d$ now follows by induction.

Secondly, we verify that $L^*(\mathbf{x}^\alpha)$ is independent of $\alpha$, when $|\alpha| = n$, say $L^*(\mathbf{x}^\alpha) = \hat{\gamma}$ whenever $\alpha \in \mathbb{N}_n^d$. This follows immediately from the construction of $L^*$, since all coordinates of the vectors $\omega_i$ are equal for all $i = 1, \ldots, n+1$. Finally, to prove optimality, we note that, for $P^*$ written in the form (4.2),

$$\|P^*\|_\Omega = L^*(P^*) = \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* L^*(\mathbf{x}^\alpha) = \hat{\gamma} \sum_{\alpha \in \mathbb{N}_n^d} a_\alpha^* = \hat{\gamma},$$

so, by the weak duality theorem, $L^*$ is indeed optimal. The required result now follows from Proposition 4.7. □

*Remark* 4.1. It is interesting to note that — for diagonally determined sets — there exists an optimal signature for problem (1.3) with support size $n+1$. One can compare this to the general result for compact $\Omega$ of a signature support size of $\dim(\Pi_n^d) = \binom{n+d}{d}$. Remarkably the minimum support size is independent of $d$ in the diagonally determined case. Moreover, we did not use compactness of $\Omega$ in the proof of Theorem 4.8, whereas compactness is needed for the general bound. Finally, note that the optimal signature in Theorem 4.8 depends on $\Omega$ only through its diagonal $[a, b]$, and is independent of $\mathbf{v}$.

**Acknowledgments.** This work is the result of a collaboration made possible by the SQuaRE program at the American Institute of Mathematics (AIM). We are truly grateful to AIM for the supportive and mathematically rich environment they provided. In addition, we owe thanks to several funding agencies, as M. D. is supported by the Australian Research Council Discovery Early Career Award DE240100674, S. F. is partially supported by grants from the National Science Foundation (DMS-2053172) and from the Office of Naval Research (N00014-20-1-2787), E. de K. is supported by grants from the Dutch Research Council (NWO) (OCENW.M.23.050 and OCENW.GROOT.2019.015), J. B. L. is supported by the AI Interdisciplinary Institute through the French program "Investing for the Future PI3A" (ANR-19-PI3A-0004) and by the National Research Foundation, Singapore, through the DesCartes and Campus for Research Excellence and Technological Enterprise (CREATE) programs, and Y. X. is partially supported by the Simons Foundation (grant #849676).

## References

[1] N. N. Andreev and V. A. Yudin, Polynomials of least deviation from zero, and Chebyshev-type cubature formulas, *Tr. Mat. Inst. Steklova* **232** (2001), 45–57.
[2] N. N. Andreev and V. A. Yudin, Best approximation of polynomials on the sphere and on the ball, in *Recent progress in multivariate approximation*, 23–30; Internat. Ser. Numer. Math., **137**, Birkhäuser, Basel, 2001.

[3] W. B. Gearhart, Some Chebyshev approximations by polynomials in two variables, *J. Approximation Theory*, **8** (1973), 195–209.

[4] B. D. Bojanov, W. Haussmann, and G. P. Nikolov, Bivariate polynomials of least deviation from zero, *Canad. J. Math.* **53** (2001), 489–505.

[5] M. Dressler, S. Foucart, E. de Klerk, M. Joldes, J. B. Lasserre, Y. Xu. *Optimization-Aided Construction of Multivariate Chebyshev Polynomials*, arXiv:2405.10438 [math.OC] (2024).

[6] D. J. Newman and Y. Xu, Tchebycheff polynomials on a triangular region, *Constr. Approx.* **9** (1993), no. 4, 543–546.

[7] I. Moale and F. Peherstorfer, An explicit class of min–max polynomials on the ball and on the sphere, *J. Approx. Theory* **163** (2011), no. 6, 724–737.

[8] M. Reimer, On multivariate polynomials of least deviation from zero on the unit ball, *Math. Z*, **1**53 (1977), no. 1, 51–58.

[9] T. J. Rivlin. Chebyshev Polynomials. Second edition. Courier Dover Publications, 2020.

[10] T. J. Rivlin and H. S. Shapiro, A unified approach to certain problems of approximation and minimization, *J. Soc. Indust. Appl. Math.* **9** (1961), 670–699.

[11] A. Shapiro. On duality theory of conic linear problems, in *Semi-Infinite Programming: Recent Advances* (M,Á. Goberna and M.A. López, eds.), Springer, 135–165 (2001).

[12] J. M. Sloss, Chebyshev approximation to zero, *Pacific J. Math.* **15** (1965), 305–313.

[13] Y. Xu, On polynomials of least deviation from zero in several variables, *Experiment. Math.* **13** (2004), no. 1, 103–112.

[14] Y. Xu, Best approximation of monomials in several variables, *Rend. Cir. Mat. Palermo (2) Suppl.* (2005), no. 76, 129–155.

Mareike Dressler, School of Mathematics and Statistics, University of New South Wales, Sydney, NSW 2052, Australia.
*Email address*: `m.dressler@unsw.edu.au`

Simon Foucart, Department of Mathematics, Texas A&M University, United States.
*Email address*: `foucart@tamu.edu`

Mioara Joldes, LAAS-CNRS, France.
*Email address*: `joldes@laas.fr`

Etienne de Klerk, Department of Econometrics and Operations Research, Tilburg University, The Netherlands.
*Email address*: `E.deKlerk@tilburguniversity.edu`

Jean-Bernard Lasserre, LAAS-CNRS and Toulouse School of Economics, France.
*Email address*: `lasserre@laas.fr`

Yuan Xu, Department of Mathematics, University of Oregon, United States.
*Email address*: `yuan@uoregon.edu`