



**HAL**  
open science

# Degree Bounds for Putinar's Positivstellensatz on the hypercube

Lorenzo Baldi, Lucas Slot

► **To cite this version:**

Lorenzo Baldi, Lucas Slot. Degree Bounds for Putinar's Positivstellensatz on the hypercube. 2023. hal-04003633v1

**HAL Id: hal-04003633**

**<https://hal.science/hal-04003633v1>**

Preprint submitted on 24 Feb 2023 (v1), last revised 22 Sep 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# DEGREE BOUNDS FOR PUTINAR’S POSITIVSTELLENSATZ ON THE HYPERCUBE

LORENZO BALDI\* AND LUCAS SLOT†

ABSTRACT. The Positivstellensätze of Putinar and Schmüdgen show that any polynomial  $f$  positive on a compact semialgebraic set can be represented using sums of squares. Recently, there has been large interest in proving effective versions of these results, namely to show bounds on the required degree of the sums of squares in such representations. These *effective* Positivstellensätze have direct implications for the convergence rate of the celebrated moment-SOS hierarchy in polynomial optimization. In this paper, we restrict to the fundamental case of the hypercube  $B^n = [-1, 1]^n$ . We show an upper degree bound for Putinar-type representations on  $B^n$  of the order  $O(f_{\max}/f_{\min})$ , where  $f_{\max}$ ,  $f_{\min}$  are the maximum and minimum of  $f$  on  $B^n$ , respectively. Previously, specialized results of this kind were available only for Schmüdgen-type representations and not for Putinar-type ones. Complementing this upper degree bound, we show a lower degree bound in  $\Omega(\sqrt[s]{f_{\max}/f_{\min}})$ . This is the first lower bound for Putinar-type representations on a semialgebraic set with nonempty interior described by a standard set of inequalities.

## CONTENTS

1. Introduction	2
1.1. Our contributions	3
2. Related works and applications	4
2.1. General effective Positivstellensätze	4
2.2. Specialized effective Positivstellensätze	5
2.3. Effective Positivstellensätze for the hypercube	5
2.4. Lower degree bounds	6
2.5. Applications to polynomial optimization	7
3. Preliminaries	8
3.1. Notations	8
3.2. The Markov Brothers’ inequality	9
3.3. Schmüdgen’s Positivstellensatz for scaled hypercubes	10
4. Proof of the upper degree bound	11
4.1. Overview of the proof	11
4.2. Proof of Theorem 3	12
5. Proof of the lower degree bound	17
6. Discussion	19
References	22
Appendix A. Explicit expressions for degree shifts	24

---

\*SORBONNE UNIVERSITÉ, CNRS, F-75005, PARIS, FRANCE

†ETH ZÜRICH, ZÜRICH, SWITZERLAND

*E-mail address:* Lorenzo.Baldi@lip6.fr, lucas.slot@inf.ethz.ch.

*Date:* February 24, 2023.

## 1. INTRODUCTION

Let  $\mathcal{S}(\mathbf{g}) \subseteq \mathbb{R}^n$  be a (basic, closed) semialgebraic set, defined in terms of the tuple of polynomials  $\mathbf{g} = (g_1, g_2, \dots, g_m)$  as:

$$\mathcal{S}(\mathbf{g}) = \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}.$$

Consider the problem of determining whether a given polynomial  $f$  belongs to the cone  $\mathcal{P}_{\geq 0}(\mathcal{S}(\mathbf{g}))$  of polynomials nonnegative on  $\mathcal{S}(\mathbf{g})$ . In general, this is a hard problem. In the unconstrained case, a straightforward way of certifying nonnegativity of  $f$  on  $\mathbb{R}^n$  is to write

$$f(\mathbf{x}) = p_1(\mathbf{x})^2 + p_2(\mathbf{x})^2 + \dots + p_\ell(\mathbf{x})^2,$$

i.e., to write  $f$  as a *sum of squares* of polynomials. Indeed, the cone  $\Sigma[\mathbf{x}]$  of such polynomials is clearly contained in  $\mathcal{P}_{\geq 0}(\mathbb{R}^n)$ . This idea extends to the constrained case by considering the *quadratic module*  $\mathcal{Q}(\mathbf{g})$  and *preordering*  $\mathcal{T}(\mathbf{g})$  of  $\mathbf{g}$ , given respectively by:

$$\mathcal{Q}(\mathbf{g}) = \left\{ \sum_{i=0}^m \sigma_i g_i : \sigma_i \in \Sigma[\mathbf{x}], \quad i = 0, 1, \dots, m \right\}, \quad (1)$$

$$\mathcal{T}(\mathbf{g}) = \left\{ \sum_{I \subseteq [m]} \sigma_I g_I : \sigma_I \in \Sigma[\mathbf{x}], \quad I \subseteq [m] \right\}. \quad (2)$$

Here,  $g_I := \prod_{i \in I} g_i$  for  $I \subseteq [m] = \{1, 2, \dots, m\}$ , and we have adopted the convention that  $g_0 = g_\emptyset = 1$ . Note that the quadratic module generated by  $\mathbf{g}$  is contained in the preordering, and that they are both contained in  $\mathcal{P}_{\geq 0}(\mathcal{S}(\mathbf{g}))$ . General representations for nonnegative polynomials on semialgebraic sets have been provided by Krivine [9] and Stengle [36]: these representations use *ratios* of polynomials in the preordering, and extend Artin's solution to Hilbert's seventeenth problem [1].

A natural question is then whether all nonnegative polynomials on  $\mathcal{S}(\mathbf{g})$  admit a denominator-free representation, i.e., whether they lie in  $\mathcal{T}(\mathbf{g})$  or even in  $\mathcal{Q}(\mathbf{g})$ . While this is not true in general, the Positivstellensätze of Putinar (under mild conditions on  $\mathbf{g}$ ) and Schmüdgen show that this is the case if one restricts to the cone  $\mathcal{P}_{> 0}(\mathcal{S}(\mathbf{g}))$  of *strictly positive* polynomials on a compact semialgebraic set  $\mathcal{S}(\mathbf{g})$ .

**Theorem 1** (Schmüdgen's Positivstellensatz [31]). *Assume that  $\mathcal{S}(\mathbf{g}) \subseteq \mathbb{R}^n$  is compact. We then have:*

$$\mathcal{P}_{> 0}(\mathcal{S}(\mathbf{g})) \subseteq \mathcal{T}(\mathbf{g}).$$

**Theorem 2** (Putinar's Positivstellensatz [25]). *Assume that  $\mathcal{Q}(\mathbf{g})$  is Archimedean, i.e., that  $R - x_1^2 - \dots - x_n^2 \in \mathcal{Q}(\mathbf{g})$  for some  $R \geq 0$ . We then have:*

$$\mathcal{P}_{> 0}(\mathcal{S}(\mathbf{g})) \subseteq \mathcal{Q}(\mathbf{g}).$$

Clearly, semialgebraic sets associated with Archimedean quadratic modules are compact, but this condition is not equivalent to compactness: there exist non-Archimedean quadratic modules that define compact semialgebraic sets, see e.g. [24, Ex. 6.3.1]. On the other hand, Theorem 1 shows that a preordering  $\mathcal{T}(\mathbf{g})$  is Archimedean if and only if the semialgebraic set  $\mathcal{S}(\mathbf{g})$  is compact.

Recently, there has been a substantial interest in proving *effective* versions of the theorems above. This means to show bounds on the minimum degree  $r$  so that

a positive polynomial  $f$  lies in the *truncated* quadratic module or preordering, that are defined, using the convention that  $g_0 = g_\emptyset = 1$ , as:

$$\mathcal{Q}(\mathbf{g})_r = \left\{ \sum_{i=0}^m \sigma_i g_i : \sigma_i \in \Sigma[\mathbf{x}], \deg(\sigma_i g_i) \leq r, \quad i = 0, 1, \dots, m \right\}, \quad (3)$$

$$\mathcal{T}(\mathbf{g})_r = \left\{ \sum_{I \subseteq [m]} \sigma_I g_I : \sigma_I \in \Sigma[\mathbf{x}], \deg(\sigma_I g_I) \leq r, \quad I \subseteq [m] \right\}. \quad (4)$$

Such bounds have immediate implications for the convergence rate of the celebrated moment-SOS hierarchy [12, 22] for polynomial optimization (see also Section 2.5). The Putinar-type representations are of particular interest, as their corresponding hierarchy leads to bounds which may be computed by solving a semidefinite program of polynomial size in the number of variables  $n$  and the number of constraints  $m$ .

**1.1. Our contributions.** In this paper, we consider the fundamental special case of the hypercube  $[-1, 1]^n$ , which can be defined as a semialgebraic set by the inequalities  $g_i(\mathbf{x}) = 1 - x_i^2 \geq 0$ ,  $i = 1, 2, \dots, n$ . The associated quadratic module  $\mathcal{Q}(\mathbf{g}) = \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)$  is Archimedean, as it contains  $n - x_1^2 - \dots - x_n^2$ , and Putinar's Positivstellensatz thus applies in this setting. In this paper we prove an upper bound and a lower bound on the degree required for representations of positive polynomials on the hypercube as elements of the quadratic module  $\mathcal{Q}(\mathbf{g})$ .

**Theorem 3** (Upper degree bound). *Let  $f \in \mathcal{P}_{>0}([-1, 1]^n)$  be a polynomial of degree  $d$ . Denote by  $f_{\max}$ ,  $f_{\min}$  the maximum and the minimum of  $f$  on  $[-1, 1]^n$ , respectively. Then there exists an absolute constant  $\mathfrak{c} > 0$  such that:*

$$f \in \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{rn} \quad \text{whenever} \quad r \geq 4\mathfrak{c} \cdot d^2(\log n) \cdot \frac{f_{\max}}{f_{\min}} + O\left(\frac{f_{\max}}{f_{\min}}\right)^{1/2}.$$

We give a precise expression for the term  $O(\sqrt{f_{\max}/f_{\min}})$  in Theorem 11. See also Section 6 for a related discussion.

**Theorem 4** (Lower degree bound). *Let  $n \geq 2$ . For any  $\varepsilon > 0$  and  $r \in \mathbb{N}$ , we have:*

$$(1 - x_1^2)(1 - x_2^2) + \varepsilon \in \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_r \implies r = \Omega(1/\sqrt[\varepsilon]{\varepsilon}).$$

Note that the function  $f(\mathbf{x}) = (1 - x_1^2)(1 - x_2^2) + \varepsilon$  satisfies  $f_{\min} = \varepsilon$  and  $f_{\max} = 1 + \varepsilon$ . We could therefore replace  $\Omega(1/\sqrt[\varepsilon]{\varepsilon})$  in Theorem 4 by  $\Omega(\sqrt[\varepsilon]{f_{\max}/f_{\min}})$ .

The same asymptotic results of Theorem 3 and Theorem 4 hold if we use  $1 \pm x_i$ ,  $i = 1, \dots, n$  (another set of standard inequalities defining  $[-1, 1]^n$ ) instead of  $1 - x_i^2$ , see Section 6.

**Outline.** The paper is structured as follows. In Section 2, we discuss the existing literature on effective Archimedean Positivstellensätze and their applications to polynomial optimization. We give detailed versions of our main results and explain their relations to prior works. In Section 3, we cover some preliminaries, particularly on approximation theory. In Section 4, we prove our upper degree bound, Theorem 3. In Section 5, we prove the lower degree bound, Theorem 4. We conclude in Section 6 by discussing possible future research directions. Finally, Appendix A is dedicated to the presentation of explicit polynomial identities exploited in Section 4.

## 2. RELATED WORKS AND APPLICATIONS

In this section, we explain the relation of our main results to the existing literature and their applications. In particular, we focus on existing effective Archimedean Positivstellensätze, for general  $\mathbf{g}$  and specific for the hypercube. Degree bounds for these theorems are usually stated in terms of a parameter of the form  $\|f\|/f_{\min, \mathcal{S}(\mathbf{g})}$ , whose inverse intuitively measures how close  $f$  is to having a zero on  $\mathcal{S}(\mathbf{g})$ . Here,  $f_{\min, \mathcal{S}(\mathbf{g})} = \min_{\mathbf{x} \in \mathcal{S}(\mathbf{g})} f(\mathbf{x})$ , and  $\|\cdot\|$  is a norm on  $\mathbb{R}[\mathbf{x}]_{\leq d}$ . Common choices include the supremum norm on  $\mathcal{S}(\mathbf{g})$  (or on a compact domain containing  $\mathcal{S}(\mathbf{g})$ ), denoted  $f_{\max, \mathcal{S}(\mathbf{g})}$ , and the coefficient norm  $\|\cdot\|_{\text{coef}}$ , defined in terms of the monomial expansion  $f(\mathbf{x}) = \sum_{\alpha} f_{\alpha} \mathbf{x}^{\alpha}$  as  $\|f\|_{\text{coef}} = \max_{\alpha} |f_{\alpha}| \cdot \frac{\prod_i (\alpha_i!)}{(\sum_i \alpha_i)!}$ . For fixed number of variables  $n$  and degree  $d$  of  $f$ , these choices are equivalent.

**2.1. General effective Positivstellensätze.** For general constraints  $\mathbf{g}$  that define a compact semialgebraic set  $\mathcal{S}(\mathbf{g})$ , Schweighofer [32] showed in a seminal work that any positive polynomial on  $\mathcal{S}(\mathbf{g})$  has a representation in the preordering  $\mathcal{T}(\mathbf{g})_r$  truncated at degree

$$r \geq O\left(\frac{\|f\|_{\text{coef}}}{f_{\min, \mathcal{S}(\mathbf{g})}}\right)^c,$$

where  $c > 0$  is a (possibly large) constant depending on  $\mathbf{g}$ . For the quadratic module, Nie & Schweighofer [20] showed a degree bound for Archimedean  $\mathcal{Q}(\mathbf{g})$  with exponential dependence on  $\|f\|_{\text{coef}}/f_{\min}$ . This result was only recently improved in [3, 4] to match Schweighofer's polynomial bound for the preordering (although the exponent  $c$  may differ).

**Theorem 5** ([4, Cor. 3.3]). *Let  $\mathcal{Q}(\mathbf{g})$  be an Archimedean quadratic module, and let  $f$  be a polynomial of degree  $d$  positive on  $\mathcal{S}(\mathbf{g})$ . Then we have, for fixed  $n$  and  $d$ ,*

$$f \in \mathcal{Q}(\mathbf{g})_r \quad \text{for} \quad r \geq O\left(\frac{f_{\max, \mathcal{D}}}{f_{\min, \mathcal{S}(\mathbf{g})}}\right)^{7L+3}$$

where  $\mathcal{D}$  is a scaled simplex containing  $\mathcal{S}(\mathbf{g})$  and  $L = L(\mathbf{g})$  is a constant (called Lojasiewicz exponent) depending only on  $\mathbf{g}$ .

The Lojasiewicz exponent can be large even when the number of variables  $n$  and the degrees  $\deg g_1, \dots, \deg g_m$  of the constraints are fixed, see [3, 4]. However, in regular cases, namely when the constraints  $\mathbf{g}$  satisfy the Constraint Qualification Conditions (CQC), one has  $L = 1$ .

**Definition 6** (CQC). *We say that a tuple of polynomials  $\mathbf{g}$  satisfies the constraint qualification conditions if, for every  $\mathbf{x} \in \mathcal{S}(\mathbf{g})$ , the gradients of the active constraints at  $\mathbf{x}$ :*

$$\{\nabla g(\mathbf{x}) : g \in \mathbf{g}, g(\mathbf{x}) = 0\}$$

are linearly independent (in particular, nonzero).

**Corollary 7** ([3, Thm. 2.11], [4, Thm. 2.10 and Cor. 3.4]). *Let  $\mathcal{Q}(\mathbf{g})$  be an Archimedean quadratic module, and let  $f$  be a polynomial of degree  $d$  positive on  $\mathcal{S}(\mathbf{g})$ . Assume that  $\mathbf{g}$  satisfies the CQC. Then we can take  $L = 1$  in Theorem 5, and thus, for fixed  $n$  and  $d$ ,*

$$f \in \mathcal{Q}(\mathbf{g})_r \quad \text{for} \quad r \geq O\left(\frac{f_{\max, \mathcal{D}}}{f_{\min, \mathcal{S}(\mathbf{g})}}\right)^{10}$$

where  $D$  is a scaled simplex containing  $\mathcal{S}(\mathbf{g})$ .

**2.2. Specialized effective Positivstellensätze.** If we restrict to certain fundamental special cases, stronger bounds are known. When  $\mathcal{S}(\mathbf{g})$  is the hypersphere [7], the hypercube [15], the unit ball [35], or the standard simplex [35], we have representations of degree  $r = O(\sqrt{f_{\max, \mathcal{S}(\mathbf{g})}/f_{\min, \mathcal{S}(\mathbf{g})}})$  in the preordering. For the hypersphere and unit ball, this bound carries over to the quadratic module (which, in those cases, is equal to the preordering). However, despite the research effort (see, e.g. Theorem 10 below), no specialized bounds on the minimum degree required for a representation in the quadratic module are known for the hypercube and the standard simplex. In this paper we start filling this gap, providing the first dedicated analysis for the quadratic module of the hypercube.

**2.3. Effective Positivstellensätze for the hypercube.** The unit hypercube  $B^n := [-1, 1]^n$  is a compact semialgebraic set that is naturally defined as:

$$B^n = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \geq 0, 1 \leq i \leq n\}, \quad g_i(\mathbf{x}) := 1 - x_i^2. \quad (5)$$

Throughout the article, we abuse notation and refer to the quadratic module and preordering generated by  $1 - x_1^2, \dots, 1 - x_n^2$  as:

$$\begin{aligned} \mathcal{Q}(B^n) &:= \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2), \\ \mathcal{T}(B^n) &:= \mathcal{T}(1 - x_1^2, \dots, 1 - x_n^2), \end{aligned}$$

and we denote their truncations (see (3) and (4)) as  $\mathcal{Q}(B^n)_r$  and  $\mathcal{T}(B^n)_r$ , respectively.

Despite its simplicity, the best available effective version of Putinar's Positivstellensatz for  $B^n$  is the general result of [4]. Indeed, since the constraints  $\mathbf{g}$  in (5) satisfy the CQC, Corollary 7 gives a degree bound of the order  $O((f_{\max}/f_{\min})^{10})$ . On the contrary, for Schmüdgen Positivstellensatz, specialized results are available, and a much stronger bound of the order  $O(\sqrt{f_{\max}/f_{\min}})$  is known.

**Theorem 8** ([15, Cor. 3]). *Let  $f \in \mathcal{P}_{>0}(B^n)$  be a polynomial of degree  $d$ , and let  $f_{\min}, f_{\max} > 0$  be the minimum and maximum of  $f$  on  $B^n$ , respectively. Then:*

$$f \in \mathcal{T}(B^n)_{(r+1)n}, \quad \text{for } r \geq \max \left\{ \left( C(n, d) \cdot \frac{f_{\max}}{f_{\min}} \right)^{1/2}, \pi d \sqrt{2n} \right\}.$$

Here,  $C(n, d)$  is a constant depending polynomially on  $n$  (for fixed  $d$ ), and polynomially on  $d$  (for fixed  $n$ ).

For ease of exposition, we stated the bound in Theorem 8 in a (slightly) weaker form than the one of [15, Cor. 3]. Theorem 8 improves upon an earlier analysis due to de Klerk & Laurent [6], who established a bound in  $O(f_{\max}/f_{\min})$ .

In the same work<sup>1</sup>, the authors propose the following conjecture (which remains open):

**Conjecture 9** (de Klerk & Laurent, 2010). *For  $n \in \mathbb{N}$  even, we have:*

$$(1 - x_1^2)(1 - x_2^2) \dots (1 - x_n^2) + \frac{1}{n(n+2)} \in \mathcal{Q}(B^n)_n.$$

<sup>1</sup>In fact, they consider there the cube  $[0, 1]^n$  defined by the constraints  $x_i \geq 0, 1 - x_i \geq 0, i \in [n]$ , but all statements carry over after a change of variables. See also Section 6.

Assuming Conjecture 9, one may prove effective versions of Putinar's Positivstellensatz for  $B^n$  starting from an effective version of Schmüdgen's Positivstellensatz. In the original paper [6], de Klerk & Laurent do so only for  $d = 2$ . Magron [16] performs an analysis in the general case.

**Theorem 10** ([16, Thm. 4]). *Let  $f \in \mathcal{P}_{>0}(B^n)$  be a polynomial of degree  $d$ . Assuming Conjecture 9 holds, we have:*

$$f \in \mathcal{Q}(B^n)_r, \quad \text{for } r \geq \exp\left(\frac{d^2 n^{d+1} \cdot \|f\|_{\text{coef}}}{f_{\min}}\right),$$

where, writing  $f(x) = \sum_{\alpha} f_{\alpha} x^{\alpha}$ , we set  $\|f\|_{\text{coef}} := \max_{\alpha} |f_{\alpha}| \cdot \frac{\prod_i (\alpha_i!)}{(\sum_i \alpha_i)!}$ .

We note that the bound of Theorem 10 is asymptotically weaker than the general result of Baldi & Mourrain (Theorem 5), but it predates it, and its dependence on  $n, d$  is more explicit.

Our main result improves exponentially upon Magron's bound, with explicit constants, and without assuming Conjecture 9. Compared to Corollary 7, it improves the dependence on  $f_{\max}/f_{\min}$  by a power of 10. With respect to Theorem 8, the degree bound is quadratically weaker, but it applies to representations in the quadratic module rather than the preordering.

**Theorem 11** (Theorem 3 with explicit constants). *Let  $f \in \mathcal{P}_{>0}(B^n)$  be a polynomial of degree  $d$  and denote  $f_{\max}, f_{\min}$  the maximum and the minimum of  $f$  on  $B^n$ , respectively. Then we have  $f \in \mathcal{Q}(B^n)_{rn}$  whenever*

$$r \geq 4c \cdot d^2 (\log n) \cdot \frac{f_{\max}}{f_{\min}} + \max \left\{ \pi d \sqrt{2n}, \left( 2c \cdot \frac{f_{\max}}{f_{\min}} \cdot C(n, d) \right)^{1/2} \right\},$$

where  $c > 0$  is the absolute constant given in Lemma 21 and  $C(n, d)$  is the constant of Theorem 8.

**2.4. Lower degree bounds.** To contextualize the positive results on the strength of sum-of-squares representations discussed above, it would be nice to have complementing *negative* results, i.e., *lower* bounds on the degree  $r$  required to represent positive polynomials. Remarkably, such results are rather rare in the literature<sup>2</sup> For non-finite semialgebraic sets, the authors are only aware of the following result of Stengle [37], which shows a lower degree bound already in the case  $n = 1$  if one uses a nonstandard representation of the interval  $B^1 = [-1, 1] \subseteq \mathbb{R}$ .

**Theorem 12** ([37, Thm. 4]). *For any  $\varepsilon > 0$  and  $r \in \mathbb{N}$ , we have:*

$$(1 - x^2) + \varepsilon \in \mathcal{T}((1 - x^2)^3)_r \implies r = \Omega(1/\sqrt{\varepsilon}).$$

Notably, the lower bound of Theorem 12 matches the best-known upper bound of Theorem 8 for the preordering of  $B^n$  (with the standard description). In Section 5, we prove the following lower degree bound for the quadratic module:

**Proposition 13.** *For any  $\varepsilon > 0$  and  $r \in \mathbb{N}$ , we have*

$$(1 - x^2)(1 - y^2) + \varepsilon \in \mathcal{Q}(B^2)_r \implies r = \Omega(1/\sqrt[8]{\varepsilon}).$$

<sup>2</sup>The exception is the case where  $\mathcal{S}(\mathbf{g}) \subseteq \mathbb{R}^n$  is a finite set, in which case *every nonnegative* polynomial on  $\mathcal{S}(\mathbf{g})$  has a representation in  $\mathcal{Q}(\mathbf{g})_N$  for some fixed  $N = N(\mathbf{g}) \in \mathbb{N}$ . There is a large body of research in that setting, particularly when  $\mathcal{S}(\mathbf{g}) \subseteq \{-1, 1\}^n$ , see, e.g., [11] and references therein.

Proposition 13 differs from Stengle's result in three important ways. First, it applies to a standard description of the hypercube  $B^n$ , while Theorem 12 does not (see Section 6 for a more detailed discussion). In particular, this description meets the constraint qualification conditions, see Definition 6, while the description that Stengle uses does not. Second, our result not only separates the quadratic module from the cone of positive polynomials, but also from the preordering. As far as we are aware, this is the first result quantifying the asymptotic gap between the quadratic module and the preordering. Third, the bound shown in our result is much weaker than Stengle's bound (it is of the order  $1/\sqrt[8]{\varepsilon}$  compared to  $1/\sqrt{\varepsilon}$ ). In fact, Stengle [37] shows his bound is best-possible up to log-factors, whereas we have no reason to believe our bound is close to optimal asymptotically (the upper bound of Theorem 3 is of the order  $1/\varepsilon$ ).

Proposition 13 generalizes to the setting  $n > 2$  in a straightforward way, yielding an immediate implication for Conjecture 9:

**Corollary 14.** *Let  $n \in \mathbb{N}$ . For any  $\varepsilon > 0$  and  $r \in \mathbb{N}$ , we have:*

$$(1 - x_1^2)(1 - x_2^2) \dots (1 - x_n^2) + \varepsilon \in \mathcal{Q}(B^n)_r \implies r = \Omega(1/\sqrt[8]{\varepsilon}).$$

*In particular, we have:*

$$(1 - x_1^2)(1 - x_2^2) \dots (1 - x_n^2) + \varepsilon \in \mathcal{Q}(B^n)_n \implies \varepsilon = \Omega(1/n^8),$$

*for every  $n \in \mathbb{N}$ .*

*Proof.* Suppose that  $n \geq 2$  and we have a representation:

$$\prod_{i=1}^n (1 - x_i^2) + \varepsilon = \sigma_0(\mathbf{x}) + \sum_{i=1}^n (1 - x_i^2) \sigma_i(\mathbf{x}) \in \mathcal{Q}(B^n)_r.$$

Then setting  $x_i = 0$  for all  $i > 2$  yields a representation:

$$\prod_{i=1,2} (1 - x_i^2) + \varepsilon = \sigma_0(\mathbf{x}) + \sum_{i=1,2} (1 - x_i^2) \sigma_i(x_1, x_2, \mathbf{0}) + \sum_{i=3}^n \sigma_i(x_1, x_2, \mathbf{0}) \in \mathcal{Q}(B^2)_r,$$

and so the lower bound  $r = \Omega(1/\sqrt[8]{\varepsilon})$  of Proposition 13 applies here as well.  $\square$

In a more abstract direction, the existence of lower degree bounds for Putinar's and Schmüdgen's Positivstellensätze is deeply related to the *non*-stability property for  $\mathcal{Q}(\mathbf{g})$  and  $\mathcal{T}(\mathbf{g})$ . This connection is hardly found in the literature (with the exception of [28]). In Section 6, we therefore recall the notion of stability, give an overview of the related results and propose some research directions.

**2.5. Applications to polynomial optimization.** A polynomial optimization problem (POP) asks to minimize a given polynomial  $p$  over a (compact) semi-algebraic set  $\mathcal{S}(\mathbf{g})$ , that is, to compute:

$$p_{\min} := \min_{\mathbf{x} \in \mathcal{S}(\mathbf{g})} p(\mathbf{x}). \quad (\text{POP})$$

Problems of the form (POP) are generally hard and have broad applications [13, 14]. The simple case of the minimization of a polynomial on the unit hypercube is of particular interest. For example, the stability number of a graph  $G = (V, E)$  equals (see for instance [21, Eq. (17)])

$$\alpha(G) = \min_{\mathbf{x} \in [-1, 1]^V} \frac{1}{2} \sum_{i \in V} (1 - x_i) - \frac{1}{4} \sum_{\{i, j\} \in E} (1 - x_i)(1 - x_j).$$



The *moment-SOS hierarchy* [12, 22] provides a series of tractable *lower bounds* on  $p_{\min}$ . Namely, for  $r \in \mathbb{N}$ , we set:

$$p_{(r)} := \max_{\lambda \in \mathbb{R}} \{ \lambda : p - \lambda \in \mathcal{Q}(\mathbf{g})_r \} \leq p_{\min}. \quad (6)$$

For fixed  $r \in \mathbb{N}$ , the bound  $p_{(r)}$  may be computed by solving a semidefinite program of size polynomial in the number of variables  $n$  and the number of constraints  $m$  defining  $\mathcal{S}(\mathbf{g})$ . If  $\mathcal{Q}(\mathbf{g})$  is Archimedean, Putinar's Positivstellensatz tells us that  $\lim_{r \rightarrow \infty} p_{(r)} = p_{\min}$ , i.e., that the hierarchy converges. In this light, effective versions of Putinar's Positivstellensatz can be thought of as bounds on the *rate* of this convergence. In this direction, our upper bound Theorem 3 and our lower bound Theorem 4 imply the following.

**Corollary 15.** *Let  $p \in \mathbb{R}[\mathbf{x}]$  be a polynomial to be minimized over the hypercube  $B^n$ , defined by  $g_i = 1 - x_i^2$  for  $i = 1, \dots, n$ , and let  $p_{(r)} \leq p_{\min}$  be the lower bound of (6). Then we have:*

$$p_{\min} - p_{(r)} = O(1/r) \quad (r \rightarrow \infty).$$

**Corollary 16.** *For each  $2 \leq n \in \mathbb{N}$ , there exists a polynomial  $p$  of degree 4 to be minimized over the hypercube  $B^n$ , defined by  $g_i = 1 - x_i^2$  for  $i = 1, \dots, n$ , with  $p_{\min} = 0$ ,  $p_{\max} = 1$ , and for which the bound of (6) satisfies:*

$$p_{\min} - p_{(r)} = \Omega(1/r^8) \quad (r \rightarrow \infty).$$

In principle, one could define a (tighter) lower bound of the form (6) using the *preordering*  $\mathcal{T}(\mathbf{g})$  instead of the quadratic module  $\mathcal{Q}(\mathbf{g})$ . The analysis with the preordering is performed in [15] (see also Theorem 8) where the authors deduce a convergence rate of  $O(1/r^2)$ . On the other hand, Corollary 15 shows weaker a degree bound in  $O(1/r)$  for case of the quadratic module. But computing the bound using the preordering would require solving a semidefinite program that is not of polynomial size in the number of constraints  $m$ , while the bound using the quadratic module has linear size in  $m$ . For this reason, the bound of Corollary 15 in  $O(1/r)$  is more relevant in practice, and its implications for polynomial optimization are arguably greater.

We notice also that the same asymptotic bounds hold true if we describe the hypercube  $B^n$  using the other standard set of inequalities, namely  $1 \pm x_i$  for  $i \in [n]$ , as explained in Section 6.

### 3. PRELIMINARIES

**3.1. Notations.** Throughout the article:

- $[n] = \{1, 2, \dots, n\}$  for  $n \in \mathbb{N}$ ;
- $x, t \in \mathbb{R}$  and  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  denote real variables;
- $\mathbb{R}[\mathbf{x}] = \mathbb{R}[x_1, \dots, x_n]$  denotes the polynomial ring in  $n$  variables;
- $\Sigma[\mathbf{x}] \subseteq \mathbb{R}[\mathbf{x}]$  denotes the convex cone of sums of squares;
- $\mathcal{Q}(B^n)_r = \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_r$  denotes the truncated quadratic module at degree  $r$  associated to the unit hypercube  $B^n = \mathcal{S}(1 - x_1^2, \dots, 1 - x_n^2)$ ;
- $f \in \mathbb{R}[\mathbf{x}]$  is a polynomial of degree  $d$ ;
- $f_{\min}, f_{\max}$  are the minimum and maximum of  $f$  on  $B^n$ , respectively;
- for  $k \in \mathbb{N}$  and  $\mathbf{x} \in \mathbb{R}^n$ ,  $\|\mathbf{x}\|_k = (\sum_{i=1}^n x_i^k)^{1/k}$  denotes the  $L^k$ -norm of  $\mathbf{x}$ , and  $\|\mathbf{x}\|_{\infty} = \max_{i=1, \dots, n} |x_i|$  denotes its  $L^{\infty}$ -norm.

**3.2. The Markov Brothers' inequality.** A key technical tool in the proofs of Section 4 and Section 5 is the Markov Brothers' inequality [17, 18], see [33] for a modern account. In its general form, it bounds the norm of (higher-order) derivatives of a polynomial of given degree in terms of its supremum norm on an appropriate unit ball. It is applied by Stengle [37] in his proof of Theorem 12. To state the theorem, we first need to introduce Chebyshev polynomials.

**Definition 17** (see, e.g., [38]). *For  $d \in \mathbb{N}$ , the Chebyshev polynomial  $T_d \in \mathbb{R}[x]$  of degree  $d$  is defined as:*

$$T_d(x) = \begin{cases} \cos(d \arccos x) & |x| \leq 1, \\ \frac{1}{2} \left( (x + \sqrt{x^2 - 1})^d + (x - \sqrt{x^2 - 1})^d \right) & |x| \geq 1. \end{cases} \quad (7)$$

We recall that  $|T_d(x)| \leq T_d(1) = 1$  for  $x \in [-1, 1]$ , that  $T_d(x) = (-1)^d \cdot T_d(-x)$  for all  $x \in \mathbb{R}$ , and finally that  $T_d(x)$  is monotonely increasing in  $x$  for  $x \geq 1$ .

**Theorem 18** (special case of [34, Thm. 2], see also [8, Thm. 1]). *Let  $\|\cdot\|$  be any norm on  $\mathbb{R}^n$ . Let  $p \in \mathbb{R}[\mathbf{x}]$  be a polynomial of degree  $d$ , and write  $\|p\|_\infty = \max_{\|\mathbf{x}\| \leq 1} |p(\mathbf{x})|$ . Then for all  $k \geq 0$  and  $\mathbf{y} \in \mathbb{R}^n$  with  $\|\mathbf{y}\| \leq 1$ , we have:*

$$\left| \frac{d^k}{dt^k} p(\mathbf{x} + t\mathbf{y}) \Big|_{t=0} \right| \leq \begin{cases} \|p\|_\infty \cdot T_d^{(k)}(1) & \|\mathbf{x}\| \leq 1, \\ \|p\|_\infty \cdot T_d^{(k)}(\|\mathbf{x}\|) & \|\mathbf{x}\| \geq 1. \end{cases} \quad (8)$$

*In particular, setting  $k = 0$ , we have:*

$$|p(\mathbf{x})| \leq \|p\|_\infty \cdot T_d(\|\mathbf{x}\|) \quad \text{for } \|\mathbf{x}\| \geq 1. \quad (9)$$

We will apply Theorem 18 for the norm  $\|\mathbf{x}\|_\infty = \max_{i=1, \dots, n} |x_i|$ , whose unit ball  $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_\infty \leq 1\}$  is the hypercube  $B^n$ . The following lemma allows us to relate the supremum norm of polynomials on scaled unit balls (i.e., scaled hypercubes), which will be convenient in the proofs of our main results.

**Lemma 19** (cf. [37, Eq. (3)]). *Let  $\|\cdot\|$  be any norm on  $\mathbb{R}^n$ , and let  $p \in \mathbb{R}[\mathbf{x}]$  be a polynomial of degree  $d$ . Then for any  $\delta \in (0, 1)$ , we have:*

$$\max_{\|\mathbf{x}\|^2 \leq \frac{1}{1-\delta}} |p(\mathbf{x})| \leq T_d\left(\frac{1}{1-\delta}\right) \cdot \max_{\|\mathbf{x}\|^2 \leq 1-\delta} |p(\mathbf{x})|.$$

*Proof.* Using (9), we find that:

$$\begin{aligned} \max_{\|\mathbf{x}\|^2 \leq \frac{1}{1-\delta}} |p(\mathbf{x})| &= \max_{\|\mathbf{y}\| \leq \frac{1}{\sqrt{1-\delta}}} |p(\mathbf{y} \cdot \sqrt{1-\delta})| \\ &\leq T_d\left(\frac{1}{1-\delta}\right) \cdot \max_{\|\mathbf{y}\| \leq 1} |p(\mathbf{y} \cdot \sqrt{1-\delta})| \\ &= T_d\left(\frac{1}{1-\delta}\right) \cdot \max_{\|\mathbf{x}\|^2 \leq 1-\delta} |p(\mathbf{x})|. \end{aligned}$$

To obtain the first equality, we simply change variables  $\mathbf{y} = \mathbf{x}/\sqrt{1-\delta}$ . Then, to get the inequality, we apply (9) to the polynomial  $\mathbf{y} \rightarrow p(\mathbf{y} \cdot \sqrt{1-\delta})$ , noting that  $\max_{|x| \leq \frac{1}{1-\delta}} T_d(x) = T_d\left(\frac{1}{1-\delta}\right)$ . Finally, we change variables again to conclude.  $\square$

In order to apply the inequalities stated above, we need the following facts on Chebyshev polynomials. These are known in the literature, but we restate them for ease of reference and completeness.

**Lemma 20** (see e.g. [33]). *For any  $x \in \mathbb{R}$ , and  $k \geq 0$ , we have:*

$$|T_d^{(k)}(x)| \leq \frac{d^2(d^2-1^2)\dots(d^2-(k-1)^2)}{1 \cdot 3 \cdot \dots \cdot (2k-1)} \cdot |T_d(x)| \leq d^{2k} \cdot |T_d(x)|.$$

**Lemma 21** (cf. [37, pf. of Thm. 4]). *Let  $1 > \delta > 0$ . Then if  $d = O(1/\sqrt{\delta})$ , we have:*

$$1 \leq T_d\left(\frac{1}{1-\delta}\right) = O(1) \quad (\delta \rightarrow 0).$$

*Furthermore, there exists an absolute constant  $1 \leq \mathfrak{c} \leq e^5$  such that for any  $d \geq 2$  and  $\delta \leq 1/d^2$ , we have  $T_d\left(\frac{1}{1-\delta}\right) \leq \mathfrak{c}$ .*

*Proof.* From (7), we find that for any  $x \geq 1$ :

$$T_d(x) \leq (x + \sqrt{x^2 - 1})^d. \quad (10)$$

As  $\frac{1}{1-\delta} = 1 + \delta + O(\delta^2)$ , we may use (10) to get:

$$\begin{aligned} T_d\left(\frac{1}{1-\delta}\right) &\leq (1 + \delta + O(\delta^2) + \sqrt{1 + 2\delta + O(\delta^2) - 1})^d \\ &\leq (1 + O(\sqrt{\delta}))^d. \end{aligned}$$

It follows that  $T_d\left(\frac{1}{1-\delta}\right) = O(1)$  if  $d = O(1/\sqrt{\delta})$ . Now, if  $d \geq 2$  and  $\delta \leq 1/d^2$ , we have  $\frac{1}{1-\delta} \leq 1 + 2\delta$ , and so by (10) we get:

$$\begin{aligned} T_d\left(\frac{1}{1-\delta}\right) &\leq (1 + 2\delta + \sqrt{(1 + 2\delta)^2 - 1})^d \\ &\leq (1 + 2\delta + \sqrt{4\delta + 4\delta^2})^d \\ &\leq (1 + 5\sqrt{\delta})^d \leq (1 + 5/d)^d \leq e^5. \quad \square \end{aligned}$$

**3.3. Schmüdgen's Positivstellensatz for scaled hypercubes.** For our arguments in Section 4, we need an effective version of Schmüdgen's Positivstellensatz for scaled hypercubes  $[-\eta, \eta]^n$ , with  $\eta > 0$ . Theorem 8 carries over to this setting in a straightforward way.

**Corollary 22.** *For  $\eta > 0$ , write  $D = [-\eta, \eta]^n$ . Let  $f \in \mathbb{R}[\mathbf{x}]$  be a polynomial of degree  $d$ , and let  $f_{\min, D}, f_{\max, D} > 0$  be the minimum and maximum of  $f$  on  $D$ , respectively. Then we have:*

$$f \in \mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_{(r+1)n}, \quad \text{for } r \geq \max \left\{ \left( C(n, d) \cdot \frac{f_{\max, D}}{f_{\min, D}} \right)^{1/2}, \pi d \sqrt{2n} \right\}.$$

*Here, the constant  $C(n, d)$  is the same as in Theorem 8.*

*Proof.* Consider the polynomial  $g(\mathbf{x}) = f(\eta\mathbf{x})$ , which is of degree  $d$ , and satisfies  $g_{\min, [-1, 1]^n} = f_{\min, [-\eta, \eta]^n}$  and  $g_{\max, [-1, 1]^n} = f_{\max, [-\eta, \eta]^n}$ . We can apply Theorem 8 to write

$$f(\eta\mathbf{x}) = g(\mathbf{x}) = \sum_{I \subseteq [n]} \sigma_I(\mathbf{x}) \prod_{i \in I} (1 - x_i^2)$$

with appropriate degree bounds on the sums of squares  $\sigma_I$ . But then,

$$\begin{aligned} f(\mathbf{x}) &= \sum_{I \subseteq [n]} \sigma_I(\mathbf{x}/\eta) \prod_{i \in I} (1 - (x_i/\eta)^2) \\ &= \sum_{I \subseteq [n]} \sigma_I(\mathbf{x}/\eta) \prod_{i \in I} \frac{1}{\eta^2} (\eta^2 - x_i^2) \\ &= \sum_{I \subseteq [n]} \eta^{-2|I|} \cdot \sigma_I(\mathbf{x}/\eta) \prod_{i \in I} (\eta^2 - x_i^2), \end{aligned}$$

which is a decomposition of  $f$  in  $\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$  of the desired degree.  $\square$

#### 4. PROOF OF THE UPPER DEGREE BOUND

This section is dedicated to the proof of Theorem 3 and Theorem 11.

We start by recalling the technique used to prove general effective versions of Putinar's Positivstellensatz in [20] and [3, 4]. There, the authors reduce the question of representing a strictly positive polynomial  $f$  on a general compact semialgebraic set  $\mathcal{S}(\mathbf{g})$ , to the question of representing strictly positive polynomials on a simpler compact domain  $D = \mathcal{S}(\mathbf{h}) \supseteq \mathcal{S}(\mathbf{g})$ . More precisely, they construct a polynomial  $p \in \mathcal{Q}(\mathbf{g})$  in such a way that  $f - p > 0$  on  $D$ . As an effective version of Schmüdgen's Positivstellensatz is available for the set  $D$ , they then deduce that  $f - p \in \mathcal{T}(\mathbf{h})$  (with an appropriate degree bound). Using the Archimedean hypothesis, we have  $\mathcal{T}(\mathbf{h}) \subseteq \mathcal{Q}(\mathbf{g})$ , which gives the final representation  $f = (f - p) + p \in \mathcal{Q}(\mathbf{g})$ . The construction of the polynomial  $p \in \mathcal{Q}(\mathbf{g})$  and the effective Schmüdgen's Positivstellensatz on  $D$  are the key parts of the proof: the different constructions in [20] and [3, 4] lead to an exponential and polynomial degree bound for the representation of  $f \in \mathcal{Q}(\mathbf{g})$ , respectively. We refer to [3] for a more detailed list of references where this technique has been exploited.

**4.1. Overview of the proof.** Compared with the general effective Putinar's Positivstellensatz, for the investigation of the special case  $\mathcal{S}(\mathbf{g}) = B^n$  we make an important change of perspective: we consider a domain  $D$  that depends on  $f$ . Namely, we choose  $D$  to be a close enough outer approximation of  $B^n$ , so that  $f$  is not only strictly positive on  $B^n$ , say  $f \geq f_{\min} > 0$  on  $B^n$ , but also  $f \geq \frac{1}{2}f_{\min} > 0$  on  $D$ . In this way we can avoid using the perturbation polynomial  $p$ , and apply directly the representation results on the outer approximation  $D$ . Concretely, we proceed as follows (see also Figure 1).

**a. Selecting the outer domain.** We choose  $D = [-\eta, \eta]^n$  to be a scaled hypercube containing  $B^n$ , where  $\eta > 1$  will be chosen in such a way that:

$$\min_{\mathbf{x} \in D} f(\mathbf{x}) \geq \frac{1}{2} \min_{\mathbf{x} \in B^n} f(\mathbf{x}) > 0,$$

see Lemma 28.

**b. Obtaining a Schmüdgen-type representation** We then apply Corollary 22, a scaled version of Theorem 8 on  $D = \mathcal{S}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$ , to represent  $f$  as an element of the preordering  $\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$ , with appropriate degree bounds.

**c. Lifting the representation** Finally, we lift the representation of  $f$  from the preordering  $\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$  to the quadratic module  $\mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)$ .

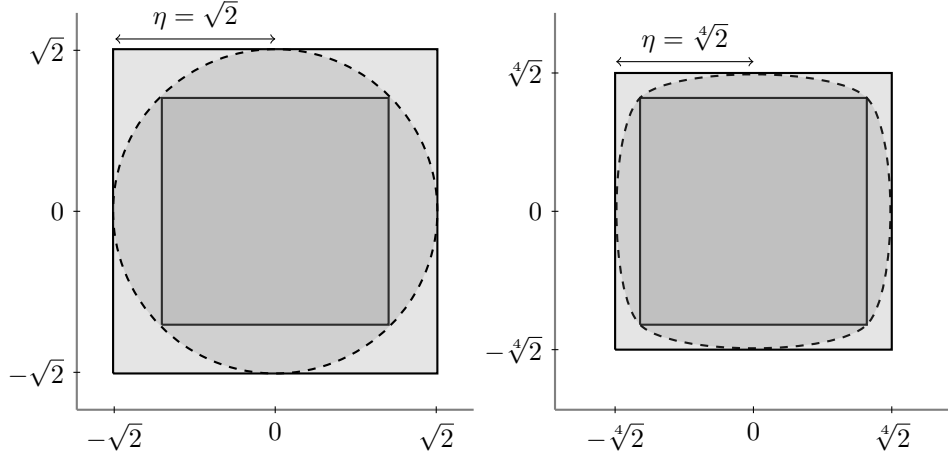


FIGURE 1. Example of the key regions involved in the proof of Theorem 3 for  $n = 2$  and  $q = 1, 2$  (on the left, right, respectively). The shaded regions depict the sets:

$$B^n = [-1, 1]^n \subseteq \mathcal{S}(n - \|\mathbf{x}\|_{2q}^{2q}) \subseteq [-\eta, \eta]^n = D.$$

For this purpose, we make use of the metric balls:

$$\{\mathbf{x} \in \mathbb{R}^n : n - \|\mathbf{x}\|_{2q}^{2q} \geq 0\}, \quad \text{where } \|\mathbf{x}\|_{2q}^{2q} = x_1^{2q} + \dots + x_n^{2q} \quad (q \in \mathbb{N}).$$

Choosing  $q \in \mathbb{N}$  large enough so that  $\eta \geq \sqrt[2q]{n}$ , we show in Lemma 26 and Theorem 27 that:

$$\begin{aligned} \mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2) &\subseteq \mathcal{T}(n - \|\mathbf{x}\|_{2q}^{2q}) \\ &= \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q}) \\ &\subseteq \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2) = \mathcal{Q}(B^n) \end{aligned}$$

with appropriate degree bounds for the truncated versions. Using the Schmüdgen-type representation obtained in the previous step, this will give us a Putinar-type representation with appropriate degree bounds:

$$f \in \mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2) \subseteq \mathcal{Q}(B^n).$$

**4.2. Proof of Theorem 3.** To present and describe our proof in a compact way, we introduce the following definition.

**Definition 23.** Let  $\mathcal{Q}(\mathbf{g})$  be a (finitely generated) quadratic module. We say that a tuple of polynomials  $\mathbf{h} = (h_1, \dots, h_s) \subseteq \mathcal{Q}(\mathbf{g})$  has degree shift  $\ell$  with respect to  $\mathbf{g}$  if  $h_i \in \mathcal{Q}(\mathbf{g})_{\deg h_i + \ell}$  for all  $i \in [s]$ .

We will make use of the following elementary lemma, that we state in general for future reference.

**Lemma 24** (Degree shift). Let  $\mathbf{h} = (h_1, \dots, h_s) \subseteq \mathcal{Q}(\mathbf{g})$  be a tuple of polynomials with degree shift  $\ell$  w.r.t.  $\mathbf{g}$ . Then, for all  $d \in \mathbb{N}$ ,

- (i)  $\mathcal{Q}(\mathbf{h})_d \subseteq \mathcal{Q}(\mathbf{g})_{d+\ell}$ ;
- (ii)  $\mathcal{T}(\mathbf{h})_d \subseteq \mathcal{Q}(g_1)_{d+s\ell}$  if  $\mathbf{g} = (g_1)$  consists of a single polynomial.

*Proof.* We set  $h_0 = g_0 = 1$  for notational convenience, and we start by proving the first part. Let  $h_i = \sum_{j=0}^m \sigma_{i,j} g_j \in \mathcal{Q}(\mathbf{g})_{\deg h_i + \ell}$ , and let  $q = \sum_{k=0}^r \bar{\sigma}_k h_k \in \mathcal{Q}(\mathbf{h})_d$ . Notice that, by definition,  $\deg(\sigma_{i,j} g_j) = \deg \sigma_{i,j} + \deg g_j \leq \deg h_i + \ell$  and  $\deg(\bar{\sigma}_k h_k) = \bar{\sigma}_k + \deg h_k \leq d$  for all  $i, j, k$ . Therefore:

$$\deg(\bar{\sigma}_k \sigma_{k,j} g_j) = \deg \bar{\sigma}_k + \deg \sigma_{k,j} + \deg g_j \leq d + \ell$$

and finally

$$q = \sum_{k=0}^r \bar{\sigma}_k h_k = \sum_{j=0}^m \left( \sum_{k=0}^r \bar{\sigma}_k \sigma_{k,j} \right) g_j \in \mathcal{Q}(\mathbf{g})_{d+\ell}$$

concluding the proof of the first part.

For the second part, we proceed in a similar way. Let  $h_i = \sigma_{0,i} + \sigma_{1,i} g_1 \in \mathcal{Q}(g_1)_{\deg h_i + \ell}$ , and let  $p = \sum_{I \subseteq [m]} \bar{\sigma}_I h_I \in \mathcal{T}(\mathbf{h})_d$ . We want to show that  $p \in \mathcal{T}(g_1)_{d+s\ell}$ . Notice that, by definition,  $\deg \bar{\sigma}_I \leq d - \deg h_I$  and

$$\prod_{i \in I} (\sigma_{0,i} + \sigma_{1,i} g_1) \in \mathcal{T}(g_1)_{\sum_{i \in I} (\deg h_i + \ell)} = \mathcal{T}(g_1)_{\deg h_I + |I|\ell} \subset \mathcal{T}(g_1)_{\deg h_I + s\ell}.$$

Therefore,

$$\bar{\sigma}_I \prod_{i \in I} (\sigma_{0,i} + \sigma_{1,i} g_1) \in \mathcal{T}(\mathbf{g})_{\deg h_I + s\ell + \deg \bar{\sigma}_I} \subseteq \mathcal{T}(g_1)_{\deg h_I + s\ell + d - \deg h_I} = \mathcal{T}(g_1)_{d+s\ell}$$

and finally:

$$p = \sum_{I \subseteq [r]} \sigma_I h_I = \sum_{I \subseteq [r]} \sigma_I \prod_{i \in I} h_i = \sum_{I \subseteq [r]} \sigma_I \prod_{i \in I} (\sigma_{0,i} + \sigma_{1,i} g_1) \in \mathcal{T}(g_1)_{d+s\ell}. \quad \square$$

In the following, we will apply Lemma 24 two times. First, to lift the representation of  $f$  from  $\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$  to  $\mathcal{T}(n - \|\mathbf{x}\|_{2q}^{2q}) = \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})$ . There, we apply the second part of Lemma 24 with  $g_1 = n - \|\mathbf{x}\|_{2q}^{2q}$  and  $h_i = \eta^2 - x_i^2$  for  $i \in [n]$ , see Lemma 26.

Second, to lift the representation of  $f$  from  $\mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})$  to  $\mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)$ . Here, we use the first part of Lemma 24 with  $\mathbf{h} = h_1 = n - \|\mathbf{x}\|_{2q}^{2q}$  and  $g_i = 1 - x_i^2$  for  $i \in [n]$ , see the proof of Theorem 27.

For these applications, we will need to determine two numbers (degree shifts)  $\ell_1, \ell_2 \in \mathbb{N}$  depending on  $\eta > 0$  such that:

$$\eta^2 - x_1^2, \dots, \eta^2 - x_n^2 \in \mathcal{T}(n - \|\mathbf{x}\|_{2q}^{2q})_{2+\ell_1}, \quad (11)$$

$$n - \|\mathbf{x}\|_{2q}^{2q} \in \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{2q+\ell_2}. \quad (12)$$

To determine these degree shifts, we start by investigating the univariate case.

**Lemma 25.** *For all  $q \in \mathbb{N}$ , the degree shift of  $1 - x^2$  with respect to  $1 - x^{2q}$  is equal to  $2q - 2$ , i.e.  $1 - x^2 \in \mathcal{Q}(1 - x^{2q})_{2q}$ .*

*Proof.* For all  $1 \leq q \in \mathbb{N}$ , consider the identity:

$$1 - x^2 = \frac{(q-1) - qx^2 + x^{2q}}{q} + \frac{1 - x^{2q}}{q} \quad (13)$$

Notice that  $(q-1) - qx^2 + x^{2q} \in \mathcal{P}_{\geq 0}(\mathbb{R})$ , since the polynomial has minimum equal to 0 attained at  $\pm 1$ . Moreover, sums of squares and nonnegative polynomials coincide in one variable, and thus  $(q-1) - qx^2 + x^{2q} \in \Sigma[x]_{2q}$ . Therefore, (13) implies that  $1 - x^2 \in \Sigma[x]_{2q} + \mathbb{R}_{\geq 0}(1 - x^{2q}) = \mathcal{Q}(1 - x^{2q})_{2q}$ , concluding the proof.  $\square$

We refer to Appendix A, and in particular to (27), for a more detailed discussion of Lemma 25 and (13).

We turn our attention to the multivariate case. We investigate the degree shift of  $\eta^2 - x_1^2, \dots, \eta^2 - x_n^2$ , i.e. the polynomials defining a scaled hypercube containing  $B^n$  for  $\eta \geq 1$ , with respect to  $n - \|\mathbf{x}\|_{2q}^{2q}$ , the polynomial defining the  $L^{2q}$  unit ball. Recall that the parameter  $\eta$  will be chosen in such a way  $f \geq f_{\min} > 0$  on  $B^n$  implies  $f \geq \frac{1}{2}f_{\min} > 0$  on  $[-\eta, \eta]^n = \mathcal{S}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$ , see Lemma 28.

We prove that the degree shift of  $\eta^2 - x_1^2, \dots, \eta^2 - x_n^2$  w.r.t.  $n - \|\mathbf{x}\|_{2q}^{2q}$  coincides with the one of Lemma 25.

**Lemma 26.** *Let  $\eta = \sqrt[2q]{n}$ . Then the degree shift of  $\eta^2 - x_1^2, \dots, \eta^2 - x_n^2$  with respect to  $n - \|\mathbf{x}\|_{2q}^{2q}$  is  $2q - 2$ . In other words,  $\eta^2 - x_i^2 \in \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})_{2q}$  for all  $i \in [n]$ .*

*Proof.* First, notice that

$$1 - x_i^{2q} = \sum_{j \neq i} x_j^{2q} + 1 - \|\mathbf{x}\|_{2q}^{2q} \in \mathcal{Q}(1 - \|\mathbf{x}\|_{2q}^{2q})_{2q}$$

and thus from Lemma 25 we deduce that

$$1 - x_i^2 \in \mathcal{Q}(1 - \|\mathbf{x}\|_{2q}^{2q})_{2q}$$

for all  $i$ .

Now let  $\eta = \sqrt[2q]{n}$ . If we substitute  $x_i \mapsto x_i/\eta$  in  $1 - x_i^2$  we obtain  $\frac{\eta^2 - x_i^2}{\eta^2}$ , while if we substitute in  $1 - \|\mathbf{x}\|_{2q}^{2q}$  we obtain  $\frac{n - \|\mathbf{x}\|_{2q}^{2q}}{n}$ . Making these substitutions in the expression  $1 - x_i^2 \in \mathcal{Q}(1 - \|\mathbf{x}\|_{2q}^{2q})_{2q}$  we therefore see that

$$\eta^2 - x_i^2 \in \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})_{2q}$$

concluding the proof.  $\square$

We refer to Appendix A and in particular to (26) for a more detailed discussion of Lemma 26. We are now ready to show one of the main results of this section.

**Theorem 27.** *Let  $\eta = \sqrt[2q]{n}$ . Then for all  $k \in \mathbb{N}$ :*

$$\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_k \subseteq \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{k+n(2q-2)}$$

*Proof.* We start moving from  $\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)$  to  $\mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})$ . From Lemma 26, the degree shift is  $\ell_1 = 2q - 2$ , and from Lemma 24(ii) we have

$$\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_k \subseteq \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})_{k+n\ell_1}$$

We now move from  $\mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})$  to  $\mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)$ . Notice that, since  $1 - x_i^{2q} = (1 - x_i^2)(1 + x_i^2 + \dots + x_i^{2q-2})$ , we have:

$$n - \|\mathbf{x}\|_{2q}^{2q} = \sum_{i=1}^n 1 - x_i^{2q} \in \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{2q}$$

and thus the degree shift of  $n - \|\mathbf{x}\|_{2q}^{2q}$  with respect to  $1 - x_1^2, \dots, 1 - x_n^2$  is equal to  $\ell_2 = 0$ . From Lemma 24(i) we then deduce that

$$\mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})_{k+n\ell_1} \subseteq \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{k+n\ell_1}.$$

We therefore have the chain of inclusions:

$$\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_k \subseteq \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})_{k+n\ell_1} \subseteq \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{k+n\ell_1},$$

where  $\ell_1 = 2q - 2$ , concluding the proof.  $\square$

Theorem 27 allows to shift the representation of a polynomial in the preordering of  $[-\eta, \eta]^n$  to a representation in the quadratic module of  $B^n$ . We refer to Appendix A and in particular (28) for explicit expressions leading to this inclusion.

We now write  $\eta = 1 + \varepsilon$  and bound  $\varepsilon$  in such a way  $f \geq f_{\min} > 0$  on  $B^n$  implies  $f \geq \frac{1}{2}f_{\min} > 0$  on  $[-1 - \varepsilon, 1 + \varepsilon]^n = [-\eta, \eta]^n$ .

**Lemma 28.** *Let  $f \geq f_{\min} > 0$  on  $B^n$  be a polynomial of degree  $d$ . Let  $\mathbf{c} \geq 1$  be the absolute constant of Lemma 21. Then  $f \geq \frac{1}{2}f_{\min}$  on  $[-1 - \varepsilon, 1 + \varepsilon]^n$  whenever*

$$\varepsilon \leq \frac{f_{\min}}{2\mathbf{c} \cdot d^2 \cdot f_{\max}}. \quad (14)$$

*Proof.* Assume that  $\varepsilon$  is as in (14) and let  $\mathbf{z} \in [-1 - \varepsilon, 1 + \varepsilon]^n$  be such that  $f(\mathbf{z}) = f_{\min, [-1 - \varepsilon, 1 + \varepsilon]^n} \leq f_{\min}$  is minimal. If  $\mathbf{z} \in B^n$ , then  $f(\mathbf{z}) = f_{\min} \geq \frac{1}{2}f_{\min}$  and there is nothing to prove. So assume  $\mathbf{z} \notin B^n$  and let  $\widehat{\mathbf{z}} \in B^n$  be a point in  $B^n$  with  $0 < \|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty} \leq \varepsilon$ . Consider the univariate polynomial  $F$  given by:

$$F(u) := f(\widehat{\mathbf{z}} + u \cdot \mathbf{v}), \quad \text{where } \mathbf{v} := \frac{(\mathbf{z} - \widehat{\mathbf{z}})}{\|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty}}.$$

Note that  $F(0) = f(\widehat{\mathbf{z}})$  and  $F(\|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty}) = f(\mathbf{z})$ . We now bound the derivative  $F'(u)$  of  $F$  for all  $0 \leq u \leq \|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty}$ , so that we can obtain a bound on the difference  $|f(\mathbf{z}) - f(\widehat{\mathbf{z}})|$ . First, notice that

$$F'(u) = \frac{d}{dt} f((\widehat{\mathbf{z}} + u \cdot \mathbf{v}) + t \cdot \mathbf{v})|_{t=0}. \quad (15)$$

As  $\|\mathbf{v}\|_{\infty} \leq 1$ , we can apply (8) to the polynomial  $f$ , with  $\mathbf{x} = \widehat{\mathbf{z}} + u \cdot \mathbf{v}$  and  $\mathbf{y} = \mathbf{v}$ , to get

$$|F'(u)| = \left| \frac{d}{dt} f((\widehat{\mathbf{z}} + u \cdot \mathbf{v}) + t \cdot \mathbf{v})|_{t=0} \right| \leq T_d(\|\widehat{\mathbf{z}} + u \cdot \mathbf{v}\|_{\infty}) \cdot \max_{\|\mathbf{x}\|_{\infty} \leq 1} |f(\mathbf{x})|.$$

Notice that, since  $0 \leq u \leq \|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty} \leq \varepsilon$ , we have

$$\|\widehat{\mathbf{z}} + u \cdot \mathbf{v}\|_{\infty} \leq 1 + u \leq 1 + \varepsilon.$$

Now, using Lemma 20 and monotonicity of  $T_d$  we get

$$T_d(\|\widehat{\mathbf{z}} + u \cdot \mathbf{v}\|_{\infty}) \leq d^2 \cdot T_d(\|\widehat{\mathbf{z}} + u \cdot \mathbf{v}\|_{\infty}) \leq d^2 \cdot T_d(1 + u) \leq d^2 \cdot T_d(1 + \varepsilon).$$

Finally, noting that  $\max_{\|\mathbf{x}\|_{\infty} \leq 1} |f(\mathbf{x})| = f_{\max}$ , we may conclude that:

$$|F'(u)| \leq d^2 \cdot T_d(1 + \varepsilon) \cdot f_{\max} \quad (0 \leq u \leq \|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty}). \quad (16)$$

Assuming (14), we have  $\varepsilon \leq 1/d^2$  and thus  $T_d(1 + \varepsilon) \leq T_d(\frac{1}{1 - \varepsilon}) \leq \mathbf{c}$ , where  $\mathbf{c} \geq 1$  is the absolute constant of Lemma 21. Using (16) and the fact that  $\|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty} \leq \varepsilon$ , we thus have:

$$\begin{aligned} |f(\mathbf{z}) - f(\widehat{\mathbf{z}})| &= |F(0) - F(\|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty})| \\ &\leq \|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty} \cdot \max_{0 \leq u \leq \|\mathbf{z} - \widehat{\mathbf{z}}\|_{\infty}} |F'(u)| \\ &\leq \varepsilon \cdot d^2 \cdot T_d(1 + \varepsilon) \cdot f_{\max} \leq \varepsilon \cdot d^2 \cdot \mathbf{c} \cdot f_{\max}. \end{aligned}$$

In conclusion, if we choose  $\varepsilon$  as in (14), we have:

$$f_{\min, [-1 - \varepsilon, 1 + \varepsilon]^n} = f(\mathbf{z}) \geq f(\widehat{\mathbf{z}}) - \mathbf{c} \cdot \varepsilon \cdot d^2 \cdot f_{\max} \geq f_{\min} - \frac{1}{2}f_{\min} = \frac{1}{2}f_{\min}. \quad \square$$

We are ready to prove our main result.



**Theorem 11** (Theorem 3 with explicit constants). *Let  $f \in \mathcal{P}_{>0}(\mathbb{B}^n)$  be a polynomial of degree  $d$  and denote  $f_{\max}, f_{\min}$  the maximum and the minimum of  $f$  on  $\mathbb{B}^n$ , respectively. Then we have  $f \in \mathcal{Q}(\mathbb{B}^n)_{rn}$  whenever*

$$r \geq 4\mathfrak{c} \cdot d^2(\log n) \cdot \frac{f_{\max}}{f_{\min}} + \max \left\{ \pi d \sqrt{2n}, \left( 2\mathfrak{c} \cdot \frac{f_{\max}}{f_{\min}} \cdot C(n, d) \right)^{1/2} \right\},$$

where  $\mathfrak{c} > 0$  is the absolute constant given in Lemma 21 and  $C(n, d)$  is the constant of Theorem 8.

*Proof of Theorem 3 and Theorem 11.* Let  $0 < \varepsilon = \frac{f_{\min}}{2\mathfrak{c} \cdot d^2 \cdot f_{\max}}$  be as in (14), and let  $q \in \mathbb{N}$  be the smallest integer such that:

$$2q \geq \frac{2 \log n}{\varepsilon} = 4\mathfrak{c} \cdot (\log n) \cdot d^2 \cdot \frac{f_{\max}}{f_{\min}}. \quad (17)$$

Then, as  $\varepsilon \leq 1$ , we have  $\log(1 + \varepsilon) \geq \varepsilon - \frac{1}{2}\varepsilon^2 \geq \frac{1}{2}\varepsilon$ , and thus

$$\frac{2q \cdot \log(1 + \varepsilon)}{\log n} \geq 1,$$

or in other words, we have  $\sqrt[q]{n} \leq 1 + \varepsilon$ . Therefore, if we set  $\eta = \sqrt[q]{n} \leq 1 + \varepsilon$ , we can deduce from Lemma 28 that  $f \geq \frac{1}{2}f_{\min}$  on  $[-\eta, \eta]^n$ . From Corollary 22, we have a representation  $f \in \mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_{(\ell+1)n}$  if  $\ell$  is any integer such that

$$\ell \geq \max \left\{ \pi d \sqrt{2n}, \left( \frac{f_{\max, [-\eta, \eta]^n}}{f_{\min, [-\eta, \eta]^n}} \cdot C(n, d) \right)^{1/2} \right\}$$

We want to express the above bound using  $\frac{f_{\max}}{f_{\min}}$  instead of  $\frac{f_{\max, [-\eta, \eta]^n}}{f_{\min, [-\eta, \eta]^n}}$ . For this, recall first that  $f_{\min, [-\eta, \eta]^n} \geq \frac{1}{2}f_{\min}$  by construction. Second, since  $\varepsilon \leq 1/d^2$ , we have  $T_d(\eta) \leq T_d(1 + \varepsilon) \leq T_d(\frac{1}{1 - \varepsilon}) \leq \mathfrak{c}$  by Lemma 21 and we can use (9) to get:

$$f_{\max, [-\eta, \eta]^n} \leq T_d(\eta) \cdot f_{\max} \leq \mathfrak{c} \cdot f_{\max}.$$

Therefore, we have:

$$\left( \frac{f_{\max, [-\eta, \eta]^n}}{f_{\min, [-\eta, \eta]^n}} \cdot C(n, d) \right)^{1/2} \leq \left( 2\mathfrak{c} \cdot \frac{f_{\max}}{f_{\min}} \cdot C(n, d) \right)^{1/2}, \quad (18)$$

and we can thus choose  $\ell$  as the smallest integer such that:

$$\ell \geq \max \left\{ \pi d \sqrt{2n}, \left( 2\mathfrak{c} \cdot \frac{f_{\max}}{f_{\min}} \cdot C(n, d) \right)^{1/2} \right\}.$$

To conclude the proof, we apply Theorem 27 and deduce that:

$$\begin{aligned} f \in \mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_{(\ell+1)n} &\subseteq \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{n(\ell+1) + n(2q-2)} \\ &= \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{n(2q + \ell - 1)}. \end{aligned}$$

Since  $q$  is the smallest integer satisfying (17), we have  $f \in \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{rn}$  whenever

$$r \geq 4\mathfrak{c} \cdot (\log n) \cdot d^2 \cdot \frac{f_{\max}}{f_{\min}} + \ell \geq 2q + \ell - 1. \quad \square$$

## 5. PROOF OF THE LOWER DEGREE BOUND

In this section, we prove our lower degree bound, Theorem 4. We consider the bivariate polynomial:

$$f(x, y) = (1 - x^2)(1 - y^2).$$

Clearly,  $f$  is nonnegative on  $B^2$  and  $f \in \mathcal{T}(B^2)_4$ . On the other hand,  $f \notin \mathcal{Q}(B^2)$ . This is well known, but we give an analytical argument for this fact as a warmup to the proof of Proposition 13.

**Proposition 29.** *We have  $f(x, y) = (1 - x^2)(1 - y^2) \notin \mathcal{Q}(B^2)$ .*

*Proof.* Suppose that  $f \in \mathcal{Q}(B^2)$ . Then  $f$  can be written as:

$$f(x, y) = \sigma_0(x, y) + (1 - x^2)\sigma_1(x, y) + (1 - y^2)\sigma_2(x, y), \quad (19)$$

where the  $\sigma_i \in \Sigma[x, y]$  are sums of squares (in particular globally nonnegative). Note that  $f(1, 1) = 0$ . We can conclude immediately that  $\sigma_0(1, 1) = 0$ . In fact, we have that  $\sigma_i(1, 1) = 0$  for all  $i \in \{0, 1, 2\}$ . Indeed, suppose for instance that  $\sigma_1(1, 1) > 0$ . Then there exists an  $1 \geq \varepsilon > 0$  such that  $\sigma_1(\sqrt{1 - \varepsilon}, 1) > 0$  by continuity. But this leads to the contradiction:

$$0 = f(\sqrt{1 - \varepsilon}, 1) \geq \varepsilon \cdot \sigma_1(\sqrt{1 - \varepsilon}, 1) > 0.$$

To finish the argument, note that from the definition of  $f$ ,

$$\frac{d^2}{dt^2} f(1 + t, 1 - t)|_{t=0} < 0. \quad (20)$$

As  $\sigma_0, \sigma_1, \sigma_2$  are globally nonnegative, and since  $\sigma_i(1, 1) = 0$ , we have that:

$$\begin{aligned} \frac{d}{dt} \sigma_i(1 + t, 1 - t)|_{t=0} &= 0, \\ \frac{d^2}{dt^2} \sigma_i(1 + t, 1 - t)|_{t=0} &\geq 0. \end{aligned}$$

By (19), this would imply that  $\frac{d^2}{dt^2} f(1 + t, 1 - t)|_{t=0} \geq 0$ , contradicting (20).  $\square$

The idea for the proof of Proposition 13 (and thus of Theorem 4) is to transform the proof above into a quantitative result. This resembles the argument of Stengle [37].

**Proposition 13.** *For any  $\varepsilon > 0$  and  $r \in \mathbb{N}$ , we have*

$$(1 - x^2)(1 - y^2) + \varepsilon \in \mathcal{Q}(B^2)_r \implies r = \Omega(1/\sqrt[8]{\varepsilon}).$$

*Proof.* Let  $f(x, y) = (1 - x^2)(1 - y^2)$ , and suppose that  $f + \varepsilon \in \mathcal{Q}(\mathbf{g})_r$ , i.e. that we have a decomposition:

$$(1 - x^2)(1 - y^2) + \varepsilon = \sigma_0(x, y) + (1 - x^2)\sigma_1(x, y) + (1 - y^2)\sigma_2(x, y), \quad (21)$$

where  $\sigma_0, \sigma_1, \sigma_2$  are sums of squares of polynomials of degree  $\deg(\sigma_i) \leq r$  (more precisely, we have  $\deg(\sigma_i) \leq r - 2$  for  $i = 1, 2$ , but this will not be important). We consider the situation locally around the point  $(1, 1) \in B^2$ . We can deduce the following facts.

**Fact 1.** *We have  $\sigma_1(1, 1) \leq \frac{1}{2}\varepsilon r^2$ .*

*Proof.* Consider the univariate polynomial  $p(x) = (1-x^2)\sigma_1(x, 1)$ . By (21), we have  $0 \leq p(x) \leq \varepsilon$  for  $x \in [-1, 1]$ . By Theorem 18 and Lemma 20, we find  $|p'(x)| \leq \varepsilon r^2$  for  $x \in [-1, 1]$ . Setting  $x = 1$ , we thus have:

$$\varepsilon r^2 \geq |p'(1)| = 2\sigma_1(1, 1). \quad \square$$

**Fact 2.** For any  $1 > \delta \geq \varepsilon$ , we have:

$$\sigma_1(x, y) \leq 2 \cdot T_r\left(\frac{1}{1-\delta}\right) \quad \text{for } x^2 \leq \frac{1}{1-\delta}, y^2 \leq \frac{1}{1-\delta}.$$

In particular,

$$\max_{x, y \in [-1, 1]} \sigma_1(x, y) \leq 2 \cdot T_r\left(\frac{1}{1-\delta}\right). \quad (22)$$

*Proof.* From (21), we have:

$$(1-x^2)\sigma_1(x, y) \leq (1-x^2)(1-y^2) + \varepsilon \quad \text{for } x, y \in [-1, 1].$$

As  $\delta \geq \varepsilon$ , we thus get:

$$\sigma_1(x, y) \leq (1-y^2) + \frac{\varepsilon}{1-x^2} \leq 1 + 1 = 2 \quad \text{for } x^2 \leq 1-\delta, y^2 \leq 1-\delta.$$

In other words, we have  $\max_{\|(x, y)\|_\infty \leq 1-\delta} |\sigma_1(x, y)| \leq 2$ . We may therefore apply Lemma 19 to  $\sigma_1$  to obtain the fact.  $\square$

**Fact 3.** Let  $g(t) = \sigma_1(1+t, 1-t)$ . Then for any  $1 > \delta \geq \varepsilon$ , and any  $u \in [-\delta, \delta]$ , we have:

$$\frac{1}{2}|g''(u)| \leq r^4 \cdot T_r\left(\frac{1}{1-\delta}\right)^2.$$

*Proof.* Assume w.l.o.g. that  $u \geq 0$ . Note that  $\frac{1}{1-\delta} \geq 1+\delta \geq 1+u$ . Using (8), (22), and Lemma 20, we therefore have that:

$$\begin{aligned} |g''(u)| &= \frac{d^2}{dt^2}(\sigma_1(1+u+t, 1-u-t))\Big|_{t=0} \\ &\leq T_r^{(2)}(1+u) \cdot \max_{x, y \in [-1, 1]} \sigma_1(x, y) \\ &\leq r^4 \cdot T_r\left(\frac{1}{1-\delta}\right) \cdot 2T_r\left(\frac{1}{1-\delta}\right). \end{aligned} \quad \square$$

**Fact 4.** Let  $g(t) = \sigma_1(1+t, 1-t)$ . Then for any  $1 > \delta \geq \varepsilon$ , we have:

$$g'(0) \leq \frac{\varepsilon}{2\delta}r^2 + \delta r^4 \cdot T_r\left(\frac{1}{1-\delta}\right)^2.$$

*Proof.* Assume  $g'(0) \geq 0$  (otherwise the statement is trivial). Note that  $g(t) \geq 0$  for all  $t \in \mathbb{R}$ . By Taylor's theorem, there exists  $u \in [-\delta, 0]$  such that:

$$\begin{aligned} 0 &\leq g(-\delta) = g(0) - g'(0) \cdot \delta + \frac{1}{2}g''(u) \cdot \delta^2, \\ \implies g'(0) &\leq \frac{g(0)}{\delta} + \frac{1}{2}|g''(u)| \cdot \delta \leq \frac{\varepsilon r^2}{2\delta} + \frac{1}{2}|g''(u)| \cdot \delta, \end{aligned}$$

where we have used that  $g(0) = \sigma_1(1, 1) \leq \frac{1}{2}\varepsilon r^2$  by Fact 1. Now apply Fact 3 to conclude the proof.  $\square$

We are ready to conclude the argument. Let  $g(t) = \sigma_1(1+t, 1-t)$ . By Taylor's theorem, there exists a  $u \in [0, \delta]$  such that:

$$\begin{aligned} g(\delta) &= g(0) + g'(0) \cdot \delta + \frac{1}{2}g''(u) \cdot \delta^2 \\ &\leq \frac{1}{2}\varepsilon r^2 + \left( \frac{1}{2}\varepsilon r^2 + \delta^2 r^4 \cdot T_r\left(\frac{1}{1-\delta}\right)^2 \right) + \delta^2 r^4 \cdot T_r\left(\frac{1}{1-\delta}\right)^2 \\ &= \varepsilon r^2 + 2\delta^2 r^4 \cdot T_r\left(\frac{1}{1-\delta}\right)^2, \end{aligned} \quad (23)$$

where we have used Fact 1, Fact 3 and Fact 4 to get the inequality. Now set  $\delta = \sqrt{\varepsilon} \geq \varepsilon$ . In light of (21), and since  $\delta \leq 1$ , we have

$$\begin{aligned} -3\delta \cdot g(\delta) &\leq (1 - (1+\delta)^2) \cdot g(\delta) \leq f(1+\delta, 1-\delta) + \varepsilon \leq -4\delta^2 + \delta^4 + \varepsilon \leq -2\varepsilon, \\ \implies g(\sqrt{\varepsilon}) &= g(\delta) \geq \frac{2}{3}\sqrt{\varepsilon}. \end{aligned}$$

Using (23), we thus find that:

$$\frac{2}{3}\sqrt{\varepsilon} \leq \varepsilon r^2 + 2\varepsilon r^4 \cdot T_r\left(\frac{1}{1-\sqrt{\varepsilon}}\right)^2, \quad \implies \frac{1}{3\sqrt{\varepsilon}} \leq \frac{1}{2}r^2 + r^4 \cdot T_r\left(\frac{1}{1-\sqrt{\varepsilon}}\right)^2. \quad (24)$$

We may assume that  $r = O(1/\sqrt[4]{\varepsilon})$  (otherwise there is nothing to prove), in which case Lemma 21 tells us that  $T_r\left(\frac{1}{1-\sqrt{\varepsilon}}\right)^2 = O(1)$ . But then (24) implies that:

$$r = \Omega(1/\sqrt[8]{\varepsilon}). \quad \square$$

*Proof of Theorem 4.* It remains to see that Proposition 13 implies Theorem 4, which is rather straightforward. Indeed, any decomposition of  $(1-x_1^2)(1-x_2^2) + \varepsilon$  in  $\mathcal{Q}(\mathbb{B}^n)_r$ ,  $n \geq 3$ , immediately gives a decomposition of  $(1-x_1^2)(1-x_2^2) + \varepsilon$  in  $\mathcal{Q}(\mathbb{B}^2)_r$  by setting  $x_3 = \dots = x_n = 0$  (see also the proof of Corollary 16).  $\square$

## 6. DISCUSSION

We have proven an upper bound on the required degree of a Putinar-type representation of a positive polynomial on  $\mathbb{B}^n = [-1, 1]^n$ , described using the inequalities  $1-x_1^2, \dots, 1-x_n^2$ , of the order  $O(f_{\max}/f_{\min})$ , see Theorem 3. This result improves upon the previously best known bound of  $O((f_{\max}/f_{\min})^{10})$ , obtained from the general result Corollary 7. Complementing this upper bound, we have exhibited a family of polynomials  $f = f_\varepsilon$  of degree 4 with  $f_{\max} = 1+\varepsilon$ ,  $f_{\min} = \varepsilon$  whose Putinar-type representations are necessarily of degree at least  $\Omega(\sqrt[8]{f_{\max}/f_{\min}}) = \Omega(1/\sqrt[8]{\varepsilon})$ , see Theorem 4. These results have direct application in polynomial optimization, see Corollary 15 and Corollary 16.

We remark that the same asymptotic results hold true if we describe  $\mathbb{B}^n$  using the inequalities  $1 \pm x_i$  for  $i = 1, \dots, n$  instead of  $1 - x_i^2$ . This follows from the identities:

$$\begin{aligned} 1 \pm x_i &= \frac{1}{2} \left( (1 \pm x_i)^2 + 1 - x_i^2 \right) \\ 1 - x_i^2 &= \frac{1}{2} \left( (1 - x_i)^2 (1 + x_i) + (1 + x_i)^2 (1 - x_i) \right) \end{aligned}$$

Hereafter we describe more connections of these results with existing literature and propose some possible future research directions.

**Stability and lower degree bounds for the Positivstellensätze.** Recall that a quadratic module  $\mathcal{Q}(\mathbf{g})$  is *Archimedean* if there exists an  $R \in \mathbb{R}$  such that  $R - \|\mathbf{x}\|_2^2 \in \mathcal{Q}(\mathbf{g})$ . Clearly,  $\mathcal{Q}(\mathbf{B}^n)$  is Archimedean, since  $n - x_1^2 - \dots - x_n^2 \in \mathcal{Q}(\mathbf{B}^n)$ . Putinar's Positivstellensatz tells us that if  $\mathcal{Q}(\mathbf{g})$  is Archimedean and  $f > 0$  on  $\mathcal{S}(\mathbf{g})$ , then  $f \in \mathcal{Q}(\mathbf{g})$ . As we have shown in Theorem 4 and Section 5 the degree needed for the representation  $f \in \mathcal{Q}(\mathbf{B}^n)$  may go to infinity as  $f_{\max}/f_{\min}$  goes to infinity for  $n \geq 2$ , even if the degree of  $f$  is fixed.

A strictly related concept is *stability*, introduced in [23]. We say that the quadratic module  $\mathcal{Q}(\mathbf{g})$  is *stable* if for all  $d \in \mathbb{N}$  there exists a  $k \in \mathbb{N}$  such that  $\mathcal{Q}(\mathbf{g}) \cap \mathbb{R}[\mathbf{x}]_{\leq d} \subseteq \mathcal{Q}(\mathbf{g})_k$ . Theorem 4 (through Proposition 13) shows that  $\mathcal{Q}(\mathbf{B}^n)$  is non-stable for  $n \geq 2$ : indeed, the degree needed for the representation of  $(1 - x_1^2)(1 - x_2^2) + \varepsilon \in \mathcal{Q}(\mathbf{B}^n)$  depends on  $\varepsilon$  and not only on the degree  $d = 4$  and  $n$ . We can regard Theorem 4, Proposition 13 and the result of Stengle [37] as quantitative versions of the *non-stability* property. Even if clearly connected, the stability and non-stability properties have not received great attention from the community working on the effective Archimedean Positivstellensätze. Therefore, hereafter we give an overview of results relating Archimedean and stability properties, proposing directions for future investigations.

We start with the one dimensional case, i.e. quadratic modules and preorderings that are subsets of  $\mathbb{R}[x]$  (for the more general case of quadratic modules and preorderings defining semialgebraic sets on real curves, see [26, 28, 30]). Recall that in  $\mathbb{R}[x]$  every finitely generated quadratic module defining a compact semialgebraic set is an Archimedean preordering, see [27]. The result of Stengle [37] shows that there are compact, one dimensional subsets of the real line which are defined by a (finitely generated) preordering that is non-stable. This is also an example of an Archimedean quadratic module that is non-stable. This happens because the choice for the generator of the preordering is not the *natural* one, see [10, 19]. The generator also does not satisfy the constraint qualification conditions. Indeed, if the preordering defining the compact set contains the natural generators, then the preordering is stable. This follows from a direct computation as in [19, Prop. 2.7.3] or applying [28, Cor. 3.18]. The converse is not true in general: the preordering  $\mathcal{T}(-x^2)$  is stable (and Archimedean) but it does not contain the natural generators  $\pm x$  of the origin. See [19, Thm. 9.3.3] for a generalization of the idea of natural generators.

We turn our attention to the two dimensional case. Every Archimedean preordering defining a semialgebraic subset of  $\mathbb{R}^2$  with nonempty interior is non-stable, see [28, Thm. 5.4] and also [28, Ex. 5.1]. Notice that in [28, Ex. 5.1], a family of strictly positive polynomials and an interior point of the semialgebraic set is used to prove non-stability, while in Proposition 13 we use a boundary point. In particular, the results in [28] apply to both  $\mathcal{Q}(\mathbf{B}^2)$  and  $\mathcal{T}(\mathbf{B}^2)$ , which are therefore non-stable. We recall also that, despite being non-stable,  $\mathcal{T}(\mathbf{B}^2)$  is *saturated*, i.e.  $\mathcal{T}(\mathbf{B}^2) = \mathcal{P}_{\geq 0}(\mathbf{B}^2)$  (see [29] or [19, Thm. 9.4.5]). On the contrary,  $\mathcal{Q}(\mathbf{B}^2) \subsetneq \mathcal{P}_{\geq 0}(\mathbf{B}^2)$ . This is an important difference and it is exploited in Proposition 13 to prove the lower bound for the representation in  $\mathcal{Q}(\mathbf{B}^n)$ . We do not know if a quantitative version of [28, Ex. 5.1], that applies also to the preordering  $\mathcal{T}(\mathbf{B}^2)$ , would give better or worse bounds compared to the bound of Proposition 13. In general, quantitatively comparing Proposition 13 and [28, Ex. 5.1] could be the first step to understand if the lower degree bounds for representations in  $\mathcal{T}(\mathbf{g})$  and  $\mathcal{Q}(\mathbf{g})$  are significantly different.

Finally, for  $\mathbf{g}$  defining a compact semialgebraic set  $\mathcal{S}(\mathbf{g})$  of dimension  $\geq 3$ , the preordering  $\mathcal{T}(\mathbf{g})$  is non-stable [28] and it is not saturated, i.e.  $\mathcal{T}(\mathbf{g}) \subsetneq \mathcal{P}_{\geq 0}(\mathcal{S}(\mathbf{g}))$ . The same results hold true for Archimedean quadratic modules  $\mathcal{Q}(\mathbf{g})$ .

**Constraint qualification conditions.** It is a natural question to ask whether the techniques used in Proposition 13 generalize to other semialgebraic sets. In the proof, we mostly work locally around the point  $(1, 1)$ , where the zero-sets of the constraints  $1 - x^2, 1 - y^2$  intersect, and we implicitly consider the Taylor expansion for  $(1 - x^2)(1 - y^2) + \varepsilon$  at  $(1, 1)$ .

In algebraic terms, working locally around  $(1, 1)$  with Taylor expansions means working in the formal power series ring  $\mathbb{R}[[1 - x, 1 - y]]$  that contains such Taylor expansions. This ring is the completion of  $\mathbb{R}[x, y]$  at the maximal ideal  $\langle 1 - x, 1 - y \rangle$ :  $\mathbb{R}[[1 - x, 1 - y]] \cong \widehat{\mathbb{R}[x, y]_{\langle 1 - x, 1 - y \rangle}}$ .

To generalize this situation, notice that the CQC for  $\mathbf{g}$ , see Definition 6, imply that, in first order approximation, the local geometry of boundary points is similar to that of the hypercube  $B^n$ . In this case at each boundary point  $\xi \in S(\mathbf{g})$ , the active constraints  $\{g_1, \dots, g_\ell\}$  at  $\xi$  (i.e., those with  $g_i(\xi) = 0$ ) can be chosen as local coordinate functions. Algebraically, this means that  $\{g_1, \dots, g_\ell\}$  are part of a system of uniformizing parameters at  $\xi$ . In this way  $\mathbb{R}[[t_1, \dots, t_n]] \cong \widehat{\mathbb{R}[\mathbf{x}]_{\mathfrak{m}_\xi}}$  with  $t_i = g_i$  for  $i = 1, \dots, \ell$  (see e.g. [39, Cor. 2 p. 137] or [19, Thm. 12.2.2]). For these reasons, we conjecture that Proposition 13 can be extended essentially to all  $\mathbf{g}$  satisfying the CQC.

**Separating the convex cones  $\mathcal{Q}(B^n)_r$ ,  $\mathcal{T}(B^n)_r$  and  $\mathcal{P}_{\geq 0}(B^n)_r$ .** In Proposition 29, we use the linear functional

$$L : f \mapsto \left. \frac{d^2 f((1, 1) + t(1, -1))}{dt^2} \right|_{t=0}$$

to show that  $(1 - x^2)(1 - y^2) \notin \mathcal{Q}(B^2)$ ; namely, we show that  $L(p) \geq 0$  for any  $p \in \mathcal{Q}(B^2)$ , whereas  $L((1 - x^2)(1 - y^2)) < 0$ . The functional  $L$  thus separates the quadratic module  $\mathcal{Q}(B^2)$  from a polynomial in the preordering  $\mathcal{T}(B^2)$ . The same idea is exploited in the proof of Proposition 13, but in a quantitative way. This allows us to deduce bounds on  $\varepsilon > 0$ , depending on  $r$ , in such a way that  $(1 - x^2)(1 - y^2) + \varepsilon \notin \mathcal{Q}(B^2)_r$ .

This idea also works if we replace the direction  $(1, -1)$  with any other direction pointing in the second or fourth quadrant, or if we replace the base point  $(1, 1)$  with another vertex of  $B^2$ . We may therefore define *families* of linear functionals that separate polynomials in the convex cones  $\mathcal{T}(B^2)_r \subset \mathcal{P}_{\geq 0}(B^2)_r$  from the convex cone  $\mathcal{Q}(B^2)_r$ .

We conjecture that it is possible to exploit these linear functionals to give bounds on the ratio of volumes of compact sections of  $\mathcal{Q}(B^2)_r$  and  $\mathcal{T}(B^2)_r$  (or  $\mathcal{Q}(B^2)_r$  and  $\mathcal{P}_{\geq 0}(B^2)_r$ ). More generally, a similar technique could be exploited for  $B^n$ ,  $n \geq 2$ . In the spirit of [5], the limit for  $n \rightarrow \infty$  of the ratio of volumes of these sections could also be investigated.

**Improving the upper degree bound.** In the proof of Theorem 3, we use an effective Schmüdgen's Positivstellensatz on a scaled hypercube  $[-\eta, \eta]^n$ ; namely Corollary 22. This corollary is responsible for the term of order  $O(\sqrt{f_{\max}/f_{\min}})$  in our result. Corollary 22 could be replaced with any other effective Schmüdgen's Positivstellensatz on  $[-\eta, \eta]^n$  with sufficiently good rate of convergence, and this

could lead to improvements of the final result. In particular, the dependence on  $n, d$  of the constant  $C(n, d)$  appearing in Corollary 22 is quite bad (see [15, Eq. (18)]), especially compared to the constant  $d^2(\log n)$  we introduce in our proof of Theorem 3. Combing the proof of Theorem 3 with a better effective Schmüdgen's Positivstellensatz on  $[-\eta, \eta]^n$  would lead to an effective Putinar's Positivstellensatz on  $B^n$  that is asymptotically interesting also for  $n, d \rightarrow \infty$ .

**Logarithmic degree bounds.** In their recent work, Bach & Rudi [2] give an alternative proof of Theorem 8, working from the perspective of trigonometric polynomials. Furthermore, they show bounds with *logarithmic* dependence in  $f_{\max}/f_{\min}$  on the required degree  $r$  in a Schmüdgen-type representation for a class of positive polynomials on  $B^n$  satisfying a strict local optimality condition.

It would be interesting to see if such assumptions might lead to better degree bounds for the quadratic module as well. We remark that the polynomials  $\prod_{i=1}^n (1 - x_i^2) + \varepsilon$  featured in Proposition 13 ( $n = 2$ ) and in Conjecture 9 ( $n \geq 2$ ) do not satisfy this condition. In fact, Proposition 13 shows that it is not possible to achieve logarithmic degree bounds for representations of general polynomials in the quadratic module. It is an open question whether a similar lower bound for general polynomials (not satisfying the strict local optimality condition) holds for the preordering as well.

**Acknowledgments.** The authors wish to thank Claus Scheiderer for the detailed discussion about Archimedean and stability properties for preorderings and quadratic modules. We further wish to thank Monique Laurent for her helpful comments and suggestions on an early version of this work, and Bernard Mourrain for useful discussions.

The first author was partially funded by the European Union's Horizon 2020 research and innovation programme POEMA, under the Marie Skłodowska-Curie Actions, grant agreement 813211.

## REFERENCES

- [1] E. Artin, *Über die Zerlegung definiter Funktionen in Quadrate*, Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg **5** (1927), no. 1, 100–115.
- [2] F. Bach and A. Rudi, *Exponential convergence of sum-of-squares hierarchies for trigonometric polynomials*, 2022, <https://arxiv.org/abs/2211.04889>.
- [3] L. Baldi and B. Mourrain, *On the Effective Putinar's Positivstellensatz and Moment Approximation*, Mathematical Programming, Series A (2022), <https://doi.org/10.1007/s10107-022-01877-6>.
- [4] L. Baldi, B. Mourrain, and A. Parusinski, *On Lojasiewicz inequalities and the effective Putinar's Positivstellensatz*, 2022, <https://arxiv.org/abs/2212.09551>.
- [5] G. Blekherman, *There are significantly more nonnegative polynomials than sums of squares*, Israel Journal of Mathematics **153** (2006), no. 1, 355–380.
- [6] E. de Klerk and M. Laurent, *Error bounds for some semidefinite programming approaches to polynomial minimization on the hypercube*, SIAM J. Optim. **20** (2010), no. 6, 3104–3120.
- [7] K. Fang and H. Fawzi, *The sum-of-squares hierarchy on the sphere and applications in quantum information theory*, Math. Program. **190** (2021), 331–360.
- [8] L. A. Harris, *A proof of Markov's theorem for polynomials on Banach spaces*, Journal of Mathematical Analysis and Applications **368** (2010), no. 1, 374–381.
- [9] J. L. Krivine, *Anneaux préordonnés*, Journal d'Analyse Mathématique **12** (1964), no. 1, 307–326.
- [10] S. Kuhlmann and M. Marshall, *Positivity, Sums of Squares and the Multi-Dimensional Moment Problem*, Transactions of the American Mathematical Society **354** (2002), no. 11, 4285–4301, Publisher: American Mathematical Society.

- [11] A. Kurpisz, S. Leppänen, and M. Mastroilli, *On the hardest problem formulations for the 0/1 Lasserre hierarchy*, *Mathematics of Operations Research* **42** (2016), 145–143.
- [12] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, *SIAM J. Optim.* **11** (2001), no. 3, 796–817.
- [13] ———, *Moments, positive polynomials and their applications*, Imperial College Press, 2009.
- [14] M. Laurent, *Sums of Squares, Moment Matrices and Optimization Over Polynomials*, 157–270, Springer New York, New York, NY, 2009, pp. 157–270.
- [15] M. Laurent and L. Slot, *An effective version of Schmüdgen's Positivstellensatz for the hypercube*, *Opt. Lett.* (2022), <https://doi.org/10.1007/s11590-022-01922-5>.
- [16] V. Magron, *Error bounds for polynomial optimization over the hypercube using Putinar type representations*, *Opt. Lett* **9** (2015), 887–895.
- [17] W. Markoff and J. Grossmann, *Über Polynome, die in einem gegebenen Intervalle möglichst wenig von Null abweichen*, *Math. Ann.* **77** (1916), no. 2, 213–258.
- [18] A. A. Markov, *Ob odnom voprosе di mendeleeva*, *Zapiski Imperatorskoi Akademii Nauk SP6* **62** (1890), 1–24.
- [19] M. Marshall, *Positive Polynomials and Sums of Squares*, American Mathematical Society, 2008.
- [20] J. Nie and M. Schweighofer, *On the complexity of Putinar's Positivstellensatz*, *J. Complex.* **23** (2007), no. 1, 135–150.
- [21] M. Park and S. Hong, *Handelman rank of zero-diagonal quadratic programs over a hypercube and its applications*, *J. Glob. Optim.* **56** (2013), 727–736.
- [22] P. Parrilo, *Structured semidefinite programs and semialgebraic geometry methods in robustness and optimization*, 2000, Ph.D. Thesis, California Institute of Technology.
- [23] V. Powers and C. Scheiderer, *The moment problem for non-compact semialgebraic sets*, *Advances in Geometry* **1** (2001), no. 1, 71–88.
- [24] A. Prestel and C. Delzell, *Positive polynomials: from Hilbert's 17th problem to real algebra*, Springer, Berlin, Heidelberg, 2001.
- [25] M. Putinar, *Positive polynomials on compact semi-algebraic sets*, *Indiana Univ. Math. J.* **42** (1993), 969–984.
- [26] C. Scheiderer, *Sums of squares on real algebraic curves*, *Mathematische Zeitschrift* **245** (2003), no. 4, 725–760.
- [27] ———, *Distinguished representations of non-negative polynomials*, *Journal of Algebra* **289** (2005), no. 2, 558–573.
- [28] ———, *Non-existence of degree bounds for weighted sums of squares representations*, *Journal of Complexity* **21** (2005), no. 6, 823–844.
- [29] ———, *Sums of squares on real algebraic surfaces*, *Manuscripta Mathematica* **119** (2006), no. 4, 395–410.
- [30] ———, *Semidefinite Representation for Convex Hulls of Real Algebraic Curves*, *SIAM Journal on Applied Algebra and Geometry* **2** (2018), no. 1, 1–25, Publisher: Society for Industrial and Applied Mathematics.
- [31] K. Schmüdgen, *The K-moment problem for compact semi-algebraic sets.*, *Mathematische Annalen* **289** (1991), no. 2, 203–206.
- [32] M. Schweighofer, *On the complexity of Schmüdgen's Positivstellensatz*, *J. Complex.* **20** (2004), no. 4, 529–543.
- [33] A. Shadrin, *Twelve proofs of the Markov inequality*, *Approximation theory – a volume dedicated to B. Bojanov* (2004), 233–298.
- [34] V. Skalyga, *Bounds for the derivatives of polynomials on centrally symmetric convex bodies*, *Izv. Math.* **69** (2005), 607–621.
- [35] L. Slot, *Sum-of-squares hierarchies for polynomial optimization and the Christoffel-Darboux kernel*, *SIAM J. Opt.* **32** (2022), no. 4, 2612–2635.
- [36] G. Stengle, *A nullstellensatz and a positivstellensatz in semialgebraic geometry*, *Mathematische Annalen* **207** (1974), no. 2, 87–97.
- [37] ———, *Complexity estimates for the Schmüdgen Positivstellensatz*, *J. Complex.* **12** (1996), 167–174.
- [38] G. Szegő, *Orthogonal polynomials*, American Mathematical Society, 1975.
- [39] O. Zariski and P. Samuel, *Commutative Algebra*, Graduate Texts in Mathematics, vol. 29, Springer, Berlin, Heidelberg, 1960.



## APPENDIX A. EXPLICIT EXPRESSIONS FOR DEGREE SHIFTS

In this appendix we discuss (13):

$$1 - x^2 = \frac{(q-1) - qx^2 + x^{2q}}{q} + \frac{1 - x^{2q}}{q}$$

that is a key ingredient for the proof of the upper bound. Hereafter we provide some related explicit formulae and their consequences in the proof of Theorem 27.

Despite its simplicity, it is difficult to derive this kind of expressions. Indeed, this is a representation for  $1 - x^2 \in \mathcal{Q}(1 - x^{2q})$  (see (27) below for an explicit sum-of-squares expression of  $(q-1) - qx^2 + x^{2q}$ ). Obtaining *exact* representations for polynomials in quadratic modules is challenging, even in the univariate case, and to the authors' best knowledge there is currently no software available to solve the problem in general.

We therefore discuss in more detail how (13) was obtained, and provide explicit expressions for Lemma 25 and Lemma 26.

Consider the equation:

$$1 - x = \frac{1}{2} ((1-x)^2 + 1 - x^2) \in \mathcal{Q}(1 - x^2)_2$$

If we substitute  $x = x^2$  we obtain

$$1 - x^2 = \frac{1}{2} ((1-x^2)^2 + 1 - x^4) \in \mathcal{Q}(1 - x^4)_4$$

More generally, substituting  $x = x^{2^{m-1}}$ , we have:

$$1 - x^{2^{m-1}} = \frac{1}{2} ((1 - x^{2^{m-1}})^2 + 1 - x^{2^m}) \in \mathcal{Q}(1 - x^{2^m})_{2^m}$$

It is then possible to obtain the explicit formula:

$$1 - x^2 = \sum_{i=1}^{m-1} \left( \frac{1}{2^i} (1 - x^{2^i})^2 \right) + \frac{1}{2^{m-1}} (1 - x^{2^m}) \in \mathcal{Q}(1 - x^{2^m})_{2^m} \quad (25)$$

which is equivalent to (13) with  $2q = 2^m$ .

Using (25), we can deduce also the explicit expression for  $\eta^2 - x_k^2 \in \mathcal{Q}(n - \|\mathbf{x}\|_{2^m}^2)_{2^m}$  in Lemma 26 :

$$\eta^2 - x_k^2 = \eta^2 \sum_{i=1}^{m-1} \left( \frac{1}{2^i} \left( 1 - \left( \frac{x_k}{\eta^2} \right)^{2^i} \right)^2 \right) + \frac{\eta^2}{2^{m-1}n} \left( \sum_{j \neq k} x_j^{2^m} + n - \|\mathbf{x}\|_{2^m}^2 \right) \quad (26)$$

We have therefore seen that for  $2q = 2^m$  the necessary representations can be derived easily. It is then possible to make an educated guess to avoid the power of 2, writing  $1 - x^2 = f_q + \frac{1-x^{2q}}{q}$ . We then obtain the polynomial  $f_q = \frac{(q-1) - qx^2 + x^{2q}}{q}$ , that is nonnegative and thus a sums of squares. An explicit way to note this is by writing:

$$\begin{cases} f_1 = 0 \\ f_{q+1} = \frac{q}{q+1} x^2 f_q + \frac{q}{q+1} (1 - x^2)^2 \end{cases}$$

or more directly:

$$f_q = \sum_{i=1}^{q-1} \frac{q-i}{q} x^{2(i-1)} (1 - x^2)^2 \quad (27)$$

Therefore the final explicit expression for  $1 - x^2 \in \mathcal{Q}(1 - x^{2q})_{2q}$  in Lemma 25 is:

$$1 - x^2 = \left( \sum_{i=1}^{q-1} \frac{q-i}{q} x^{2(i-1)} (1-x^2)^2 \right) + \frac{1-x^{2q}}{q} \in \mathcal{Q}(1-x^{2q})_{2q}$$

We can deduce also an explicit expression for Lemma 26, i.e. for  $\eta^2 - x_i^2 \in \mathcal{Q}(n - \|\mathbf{x}\|_{2q}^{2q})_{2q}$ :

$$\eta^2 - x_i^2 = \eta^2 f_q \left( \frac{x_i}{\eta} \right) + \frac{\eta^2}{qn} \left( \sum_{j \neq i} x_j^{2q} + n - \|\mathbf{x}\|_{2q}^{2q} \right) \quad (28)$$

with  $f_q$  as in (27). The equations (28) and (27) give also explicit expressions for the inclusion

$$\mathcal{T}(\eta^2 - x_1^2, \dots, \eta^2 - x_n^2)_k \subseteq \mathcal{Q}(1 - x_1^2, \dots, 1 - x_n^2)_{k+n(2q-2)}$$

in Theorem 27.