



On the complexity of Putinar-Vasilescu's Positivstellensatz

Ngoc Hoang Anh Mai, Victor Magron

► To cite this version:

Ngoc Hoang Anh Mai, Victor Magron. On the complexity of Putinar-Vasilescu's Positivstellensatz. Journal of Complexity, 2022, <10.1016/j.jco.2022.101663>. <hal-03820206>

HAL Id: hal-03820206

<https://laas.hal.science/hal-03820206v1>

Submitted on 22 Jul 2024

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.



Distributed under a Creative Commons CC BY-NC 4.0 - Attribution - Non-commercial use - International License

On the complexity of Putinar–Vasilescu’s Positivstellensatz

Ngoc Hoang Anh Mai*, Victor Magron*†

January 14, 2022

Abstract

We provide a new degree bound on the weighted sum-of-squares (SOS) polynomials for Putinar–Vasilescu’s Positivstellensatz. This leads to another Positivstellensatz saying that if f is a polynomial of degree at most $2d_f$ nonnegative on a semialgebraic set having nonempty interior defined by finitely many polynomial inequalities $g_j(x) \geq 0$, $j = 1, \dots, m$ with $g_1 := L - \|x\|_2^2$ for some $L > 0$, then there exist positive constants \bar{c} and c depending on f, g_j such that for any $\varepsilon > 0$, for all $k \geq \bar{c}\varepsilon^{-c}$, f has the decomposition

$$(1 + \|x\|_2^2)^k (f + \varepsilon) = \sigma_0 + \sum_{j=1}^m \sigma_j g_j, \quad (1)$$

for some SOS polynomials σ_j being such that the degrees of $\sigma_0, \sigma_j g_j$ are at most $2(d_f + k)$. Here $\|\cdot\|_2$ denotes the ℓ_2 vector norm. As a consequence, we obtain a converging hierarchy of semidefinite relaxations for lower bounds in polynomial optimization on basic compact semialgebraic sets. The complexity of this hierarchy is $\mathcal{O}(\varepsilon^{-c})$ for prescribed accuracy $\varepsilon > 0$. In particular, if $m = L = 1$ then $c = 65$, yielding the complexity $\mathcal{O}(\varepsilon^{-65})$ for the minimization of a polynomial on the unit ball.

Keywords: positivity certificate; Putinar–Vasilescu’s Positivstellensatz; basic semialgebraic set; sum-of-squares; polynomial optimization; moment-SOS hierarchy

Contents

1	Introduction	2
1.1	Background	3
1.2	Contribution	5
1.3	Technical insights	6
2	Notation and definitions	6
3	Representation theorems and degree bounds	7
3.1	Polynomials nonnegative on general semialgebraic sets	7
3.2	Polynomials nonnegative on compact semialgebraic sets	8
3.3	Preliminary material	9
3.4	The proof of Theorem 2	12

*CNRS; LAAS; 7 avenue du Colonel Roche, F-31400 Toulouse; France.

†Université de Toulouse; LAAS; F-31400 Toulouse, France.

3.4.1	Construction of the positive weight functions	13
3.4.2	Polynomial approximations for the weight functions	19
3.4.3	Applying the global positivity certificate	20
4	Polynomial optimization	21
4.1	General case	21
4.2	Compact case	22
5	Conclusion	23

1 Introduction

For a positive $m \in \mathbb{N}$, let us consider the polynomial optimization problem (POP):

$$f^* := \inf_{x \in S} f(x), \quad (2)$$

where $f \in \mathbb{R}[x]$ and

$$S := \{x \in \mathbb{R}^n : g_j(x) \geq 0, j \in [m]\}, \quad (3)$$

for some $g_j \in \mathbb{R}[x]$, $j \in [m]$. Here $\mathbb{R}[x]$ denotes the ring of real polynomials in vector of variables $x = (x_1, \dots, x_n)$ and $[m]$ stands for the set $\{1, \dots, m\}$. Assume that f has degree at most $2d_f$ for some positive $d_f \in \mathbb{N}$. The set S is a conjunction of finitely many polynomial inequalities, and therefore is called a *basic semialgebraic set*.

Problem (2) can be written as

$$f^* = \sup_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda > 0 \text{ on } S\}. \quad (4)$$

We can replace the inequality constraint of problem (4) by an equality constraint, if one can represent positive polynomials on S . Assume that S has nonempty interior and a ball constraint is present, i.e., $g_1 = L - \|x\|_2^2$ for some $L > 0$. Our first (minor) contribution is to rely on the representation of polynomials positive on S stated by Putinar and Vasilescu [17], to obtain

$$f - \lambda = \frac{\sigma_0 + \sum_{j \in [m]} \sigma_j g_j}{(1 + \|x\|_2^2)^k}, \quad (5)$$

for some $k \in \mathbb{N}$, $\sigma_j \in \Sigma[x]$, $j \in [m]$, being such that $\deg(\sigma_0) \leq 2(k + d_f)$ and $\deg(\sigma_j g_j) \leq 2(k + d_f)$. Here $\Sigma[x]$ denotes the set of sum-of-squares (SOS) polynomials and $\deg(\cdot)$ stands for the degree of a polynomial. Such a representation of positive polynomials is called a *Positivstellensatz*.

After bounding the degrees of the SOS polynomials involved in (5), we obtain the following hierarchy of relaxations indexed by $k \in \mathbb{N}$:

$$\begin{aligned} \rho_k^{(0)} &:= \sup_{\lambda, \sigma_j} \lambda \\ \text{s.t.} \quad &\lambda \in \mathbb{R}, \sigma_j \in \Sigma[x], \\ &(1 + \|x\|_2^2)^k (f - \lambda) = \sigma_0 + \sum_{j \in [m]} \sigma_j g_j, \\ &\deg(\sigma_0) \leq 2(k + d_f), \deg(\sigma_j g_j) \leq 2(k + d_f). \end{aligned} \quad (6)$$

Problem (6) can be solved numerically using semidefinite programming [1]. It is due to the fact that every SOS polynomial σ of degree $2t$ has the form $\sigma = v_t^\top G v_t$ for some positive semidefinite matrix G (which is denoted by $G \succeq 0$, i.e., G is symmetric and each eigenvalue of G is nonnegative), with v_t being the vector of all monomials $x^\alpha := x_1^{\alpha_1} \dots x_n^{\alpha_n}$ of degree at most t . Such a matrix G is called a *Gram matrix* associated to σ . It is easy to see that for each $k \in \mathbb{N}$, $\rho_k^{(0)}$ is a lower bound of f^* , that the sequence $(\rho_k^{(0)})_{k \in \mathbb{N}}$ is monotone nondecreasing, and converges to f^* .

In the present paper, we answer the following two interesting questions:

1. How fast does $(\rho_k^{(0)})_{k \in \mathbb{N}}$ converge to f^* ? We show the convergence rate $\mathcal{O}(k^{-1/c})$ for some constant $c > 0$ depending on f and g_j .
2. Is there any explicit example to illustrate this rate of convergence? If S is the unit ball, i.e., $m = 1$ and $g_1 = 1 - \|x\|_2^2$, the sequence $(\rho_k^{(0)})_{k \in \mathbb{N}}$ converges to f^* with the rate $\mathcal{O}(k^{-1/65})$.

1.1 Background

Positivity certificates have been studied throughout history of the development of real algebraic geometry. Nowadays it becomes a powerful tool for polynomial optimization thanks to the so-called Moment-SOS hierarchy (also known as “Lasserre’s hierarchy”) [8]. The convergence rate of the Moment-SOS hierarchy to the optimal value of a POP inherently depends on the complexity of the representation of positive polynomials. Roughly speaking, obtaining a lower degree bound on the SOS polynomials involved in the positivity certificate allows one to improve the convergence rate of the corresponding Moment-SOS hierarchy. How to find such lower degree bound is an interesting question and goes hand in hand with the quest of improving the convergence analysis of the Moment-SOS hierarchy. Let us review some of the standard results on degree bounds of positivity certificates and the corresponding convergence rates of the Moment-SOS hierarchy.

Reznick provides in [19] a Positivstellensatz for positive definite forms with an explicit degree bound. Namely, if p is a positive definite form, i.e., p is homogeneous and positive except at the origin, then for all $k \in \mathbb{N}$ satisfying

$$k \geq \frac{2nd(2d-1)}{4 \log 2} \Theta(p) - \frac{n+2d}{2}, \quad (7)$$

then $\|x\|_2^{2k} p$ is a homogeneous SOS polynomial of degree $2(k+d)$, where $2d = \deg(p)$. Here for each $h \in \mathbb{R}[x]$,

$$\Theta(h) := \frac{\sup_{x \in \mathbb{S}^{n-1}} h(x)}{\inf_{x \in \mathbb{S}^{n-1}} h(x)}. \quad (8)$$

This yields a linear convergence rate of $\mathcal{O}(\varepsilon^{-1})$ for the minimization of a polynomial (see [9, Theorem 6]).

Powers and Reznick [15] improve the existing degree bound available for Pólya’s Positivstellensatz [14]. Explicitly, if p is a homogeneous polynomial of degree d positive on the simplex

$$\Delta_n = \{x \in \mathbb{R}^n : x_j \geq 0, j \in [n], \sum_{j \in [n]} x_j = 1\}, \quad (9)$$

then for all $k \in \mathbb{N}$ satisfying

$$k \geq \frac{d(d-1)\|p\|}{2 \min_{x \in \Delta_n} p(x)} - d, \quad (10)$$

$(\sum_{j \in [n]} x_j)^k p$ has positive coefficients. Here for each $h = \sum_{\alpha} h_{\alpha} x^{\alpha} \in \mathbb{R}[x]$, we note $\|h\| := \max_{\alpha} \frac{|h_{\alpha}|}{c_{\alpha}}$ with $c_{\alpha} := \frac{|\alpha|!}{\alpha_1! \dots \alpha_n!}$ for each $\alpha \in \mathbb{N}^n$. This yields a linear convergence rate of $\mathcal{O}(\varepsilon^{-1})$ for the minimization of a homogeneous polynomial on the simplex.

Applying the result of Powers and Reznick, Schweighofer [22] obtains a degree bound for Schmüdgen’s Positivstellensatz [20] claiming that given a semialgebraic $S \subset (-1, 1)^n$ defined as in (3) and a polynomial f positive on S , then there exists a real $c > 0$ depending on S such that for all $k \in \mathbb{N}$ satisfying

$$k \geq cd_f^2 \left(1 + \left(d_f^2 n^{d_f} \frac{\|f\|}{f^*} \right)^c \right), \quad (11)$$

one has $f \in \mathcal{P}_k$, where \mathcal{P}_k is the truncated preordering of order $k \in \mathbb{N}$ associated with S :

$$\mathcal{P}_k := \{\sum_{\alpha \in \{0,1\}^m} \sigma_\alpha g_1^{\alpha_1} \dots g_m^{\alpha_m} : \sigma_\alpha \in \Sigma[x], \deg(\sigma_\alpha g_1^{\alpha_1} \dots g_m^{\alpha_m}) \leq k\}. \quad (12)$$

Consequently, the corresponding SOS hierarchy of lower bounds $(\rho_k^{\text{pre}})_{k \in \mathbb{N}}$, with

$$\rho_k^{\text{pre}} := \sup_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \in \mathcal{P}_k\}, \quad k \in \mathbb{N}, \quad (13)$$

converges to f^* with the rate $\mathcal{O}(\varepsilon^{-c})$. Nevertheless, the representation of $f - \lambda$ in \mathcal{P}_k involves 2^m SOS polynomials.

Relying on the degree bound for Schmüdgen's Positivstellensatz, Nie and Schweighofer analyze in [11] the complexity of Putinar's Positivstellensatz. Namely, if $S \subset (-1, 1)^n$, $g_1 = L - \|x\|_2^2$ for some $L > 0$ and f is positive on S , then there exists a real $c > 0$ depending on S such that for all $k \in \mathbb{N}$ satisfying

$$k \geq c \exp \left(\left(d_f^2 n^{d_f} \frac{\|f\|}{f^*} \right)^c \right), \quad (14)$$

one has $f \in \mathcal{Q}_k$, where \mathcal{Q}_k stands for the truncated quadratic module of order $k \in \mathbb{N}$ associated with S :

$$\mathcal{Q}_k := \{\sigma_0 + \sum_{j \in [m]} \sigma_j g_j : \sigma_j \in \Sigma[x], \deg(\sigma_0) \leq k, \deg(\sigma_j g_j) \leq k\}. \quad (15)$$

Accordingly, the corresponding SOS hierarchy of lower bounds $(\rho_k^{\text{mod}})_{k \in \mathbb{N}}$, with

$$\rho_k^{\text{mod}} := \sup_{\lambda \in \mathbb{R}} \{\lambda : f - \lambda \in \mathcal{Q}_k\}, \quad k \in \mathbb{N}, \quad (16)$$

converges to f^* with the rate $\mathcal{O}(\exp(\varepsilon^{-c}))$. Despite of this exponential convergence rate, the representation of $f - \lambda$ in \mathcal{Q}_k involves only $m + 1$ SOS polynomials which is in deep contrast with the exponential number of SOS polynomials involved in the representation in \mathcal{P}_k .

However, the convergence rate of Schweighofer is still comparable to the later one of Nie–Schweighofer when the semialgebraic set S is defined by $m_1 \leq 1$ inequality constraints $g_j(x) \geq 0$ and m_2 equality constraints $h_i(x) = 0$. In this case, it turn out that

$$\begin{aligned} \rho_k^{\text{pre}} &= \sup_{\lambda, \sigma_j, \eta_i} \lambda \\ \text{s.t.} \quad &\lambda \in \mathbb{R}, \sigma_j \in \Sigma[x], \eta_i \in \mathbb{R}[x] \\ &f - \lambda = \sigma_0 + \sum_{j \in [m_1]} \sigma_j g_j + \sum_{i \in [m_2]} \eta_i h_i, \\ &\deg(\sigma_0) \leq k, \deg(\sigma_j g_j) \leq k, \deg(\eta_i h_i) \leq k, \end{aligned} \quad (17)$$

and the sequence $(\rho_k^{\text{pre}})_{k \in \mathbb{N}}$ converges to f^* with the rate $\mathcal{O}(\varepsilon^{-c})$. Here $[0] = \emptyset$. In particular, $c = \frac{1}{2}$ when f is homogeneous, $d_f \leq n$, $m_1 = 0$ and $m_2 = 1$ with $h_1 = 1 - \|x\|_2^2$ as shown by Fang and Fawzi [4]. Thus they obtain the quadratic convergence rate $\mathcal{O}(k^{-2})$ for the minimization of a homogeneous polynomial on the unit sphere. This improves upon the earlier linear convergence rate $\mathcal{O}(k^{-1})$ by Doherty and Wehner in [3].

Recently Slot and Laurent [23] have provided several convergence rates for Lasserre's measure-based *upper bounds* for polynomial optimization, on specific convex domains and reference measures. This is in contrast with the present work, where we provide a new convergence rate for the Moment-SOS hierarchy of lower bounds [for the minimization of a possibly nonconvex polynomial on a possibly nonconvex semialgebraic set](#).

Our contribution is concerned with the case of basic semialgebraic sets having nonempty interiors. Basically one obtains a convergence rate similar in spirit and magnitude of Schweighofer's bound $\bar{c}\varepsilon^{-c}$, but still based on the quadratic module \mathcal{Q}_k

(not the preordering \mathcal{P}_k) thanks to the prescribed denominator $(1 + \|x\|_2^2)^k$ involved in Putinar–Vasilescu’s Positivstellensatz.

Before showing explicitly our contribution, we restate the original result of Putinar and Vasilescu (without degree bound) in the following theorem:

Theorem 1. (Putinar–Vasilescu [18, Corollary 4.3 and 4.4]) *Let θ be the quadratic polynomial $x \mapsto \theta(x) := 1 + \|x\|_2^2$, and denote by $\tilde{p} \in \mathbb{R}[x, x_{n+1}]$ the homogeneous polynomial associated with $p \in \mathbb{R}[x]$, defined by $x \mapsto \tilde{p}(x) := x_{n+1}^{\deg(p)} p(x/x_{n+1})$.*

1. *Let $f \in \mathbb{R}[x]$ such that $\tilde{f} > 0$ on $\mathbb{R}^{n+1} \setminus \{0\}$. Then $\theta^k f \in \Sigma[x]$ for some $k \in \mathbb{N}$.*
2. *Let $f, g_1, \dots, g_m \in \mathbb{R}[x]$ satisfy the following two conditions:*
 - (a) *$f = f_0 + f_1$ such that $\deg(f_0) < \deg(f_1)$ and $\tilde{f}_1 > 0$ on $\mathbb{R}^{n+1} \setminus \{0\}$;*
 - (b) *$f > 0$ on $S = \{x \in \mathbb{R}^n : g_j(x) \geq 0, j \in [m]\}$.*

Then $\theta^{2k} f \in \cup_{r=1}^{\infty} \mathcal{Q}_r$ for some $k \in \mathbb{N}$, where \mathcal{Q}_r is defined as in (15).

1.2 Contribution

The construction of the hierarchy of semidefinite relaxations (6) is based on the Positivstellensatz stated in Corollary 2. More explicitly, if S has nonempty interior such that $g_1 = L - \|x\|_2^2$ for some $L > 0$ and f is of degree at most $2d_f$ such that f is nonnegative on S , then there exist positive constants \bar{c} and c depending on f, g_j such that for all $\varepsilon > 0$, for all $k \geq \bar{c}\varepsilon^{-c}$,

$$(1 + \|x\|_2^2)^k (f + \varepsilon) = \sigma_0 + \sum_{j \in [m]} \sigma_j g_j, \quad (18)$$

for some $\sigma_j \in \Sigma[x]$ being such that $\deg(\sigma_0) \leq 2(k + d_f)$ and $\deg(\sigma_j g_j) \leq 2(k + d_f)$.

In order to prove (18), we provide a degree bound on the weighted SOS polynomials for the homogenized Putinar–Vasilescu’s Positivstellensatz [17]. This is stated in Theorem 2 as follows: If f, g_1, \dots, g_m are homogeneous polynomials of even degrees such that S has nonempty interior and f is nonnegative on S , then there exist positive constants \bar{c} and c depending on f, g_j such that for all $\varepsilon > 0$, for all $k \geq \bar{c}\varepsilon^{-c}$,

$$\|x\|_2^{2k} (f + \varepsilon \|x\|_2^{2d_f}) = \sigma_0 + \sum_{j \in [m]} \sigma_j g_j, \quad (19)$$

for some homogeneous SOS polynomials σ_j being such that $\deg(\sigma_0) = \deg(\sigma_j g_j) = 2(k + d_f)$. Here a polynomial p is homogeneous of degree $2t$ if $p(\lambda x) = \lambda^{2t} p(x)$ for all $x \in \mathbb{R}^n$ and each $\lambda \in \mathbb{R}$. Remark that the original version of Putinar–Vasilescu’s Positivstellensatz in [17] does not include any degree bound on the weighted SOS polynomials σ_j involved in the representation (19). Our proof of Theorem 2 consists of three main steps:

1. Construct iteratively some positive “weight” functions ψ_j such that $f + \varepsilon - \sum_{j \in [m]} \psi_j g_j$ is positive on $[-1, 1]^n$. The idea of this step is similar in spirit to the proof of the inductive property in [21, Proposition 3.1] and relies on the Łojasiewicz inequality.
2. Approximate $\sqrt{\psi_j}$ with the multivariate Bernstein polynomial q_j on $[-1, 1]^n$ such that the polynomial $H = f + \varepsilon - \sum_{j \in [m]} q_j^2 g_j$ is positive on the unit sphere \mathbb{S}^{n-1} .
3. Apply Reznick’s Positivstellensatz [19] to the homogenization of H .

The complexity analysis of every step is derived to get the final degree bound $\bar{c}\varepsilon^{-c}$.

Afterwards, we obtain in Corollary 1 the same degree bound for the dehomogenized Putinar–Vasilescu’s Positivstellensatz. This improves the bound $\mathcal{O}(\exp(\varepsilon^{-c}))$ obtained in our previous work [9], based on Nie–Schweighofer’s complexity result [11] for Putinar’s Positivstellensatz [16]. Corollary 1 yields the convergence rate $\mathcal{O}(\varepsilon^{-c})$ for the corresponding hierarchy of relaxations for polynomial optimization on general (not necessarily compact) basic semialgebraic sets.

1.3 Technical insights

We start to recall the two main steps in the proof of Nie and Schweighofer [11] for the degree bound of SOS polynomials involved in Putinar's Positivstellensatz:

1. Find a large enough $k \in \mathbb{N}$ such that the polynomial

$$F = f + \varepsilon - \lambda \sum_{j \in [m]} (g_j - 1)^{2k} g_j \quad (20)$$

is positive on $[-1, 1]^n$. An estimate $k \geq O(\varepsilon^{-c})$ is obtained. Here $\varepsilon > 0$ measures how close the polynomial f (assumed to be nonnegative on S) is to have a zero on S .

2. Apply Schmüdgen's Positivstellensatz to F on $[-1, 1]^n$.

Notice that Schweighofer's degree bound of Schmüdgen's Positivstellensatz is exponential in the degree of the given positive polynomial (n^{d_f} in (11)). Accordingly, Nie and Schweighofer obtain an exponential bound $n^{O(\varepsilon^{-c})}$ in the second step since $\deg(F) \sim Ck$ as $k \rightarrow \infty$ for some positive constant C .

One notable difference in our proof is that the weight $\lambda(g_j - 1)^k$ in (20) is replaced by a non-differentiable positive function ψ_j . Surprisingly, we can prove that the square root $\sqrt{\psi_j}$ is a Lipschitz continuous function. Thus each $\sqrt{\psi_j}$ can be approximated with a Bernstein polynomials q_i on $[-1, 1]$. Here, the advantage of using Bernstein polynomials is that the approximation error between $\sqrt{\psi_j}$ and q_i decreases with a rate which only depends on a Lipschitz constant of $\sqrt{\psi_j}$, and $|q_i|$ is upper bounded by the supremum of $\sqrt{\psi_j}$ on $[-1, 1]^n$.

Next, we apply Reznick's Positivstellensatz to the homogeneous polynomial \tilde{H} obtained from the homogenization of

$$H := f + \varepsilon - \sum_{j \in [m]} q_j^2 g_j, \quad (21)$$

being such that the bounds of \tilde{H} and H on the unit sphere are the same. The important point to note here is that the degree bound of Reznick's Positivstellensatz is quadratic in the degree of \tilde{H} and linear in the ratio $\Theta(\tilde{H})$ (see (7)). This is in deep contrast with Schmüdgen's Positivstellensatz, as there is no exponential dependency in these two quantities. This leads to the difference between our convergence rate $O(\varepsilon^{-c})$ and Nie-Schweighofer's rate $O(\exp(\varepsilon^{-c}))$.

One may ask whether with the same techniques from our proof, one could improve the existing degree bound for Putinar's Positivstellensatz. We have tried to apply the degree bound (10) of Pólya's Positivstellensatz to H after a change of coordinate, but unfortunately this leads to the same bound as Nie and Schweighofer. The underlying reason is that the norm $\|p\|$ in (10) depends on the coefficients of p . In our situation, p coincides with H and the coefficients of H are bounded by a value involving the coefficients of the Bernstein polynomials. The bound on the largest coefficient, even for a univariate Bernstein polynomial, seems to be exponential in the approximation order t , namely, $\binom{2t}{t} \sim \frac{4^t}{\sqrt{\pi t}}$ as $t \rightarrow \infty$. The same issue occurs when we apply the degree bound of Schmüdgen's Positivstellensatz instead of the one of Pólya's Positivstellensatz.

2 Notation and definitions

In this section, we introduce mandatory notation and definitions. With $x := (x_1, \dots, x_n)$, let $\mathbb{R}[x]$ stand for the ring of real polynomials and let $\Sigma[x] \subset \mathbb{R}[x]$ be the subset of sum-of-squares (SOS) of polynomials. Let us note $\mathbb{R}[x]_t$ and $\Sigma[x]_t$ the respective restrictions of these two sets to polynomials of degree at most t and $2t$.

Given $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, we note $|\alpha| := \alpha_1 + \dots + \alpha_n$. Denote $\mathbb{N}_t^n = \{\alpha \in \mathbb{N}^n : |\alpha| \leq t\}$ for each $t \in \mathbb{N}$. Let $(x^\alpha)_{\alpha \in \mathbb{N}^n}$ be the canonical basis of monomials

for $\mathbb{R}[x]$ (ordered according to the graded lexicographic order) and $v_t(x)$ be the vector of all monomials up to degree t , with length $\binom{n+t}{n}$. A polynomial $f \in \mathbb{R}[x]_t$ is written as $f(x) = \sum_{|\alpha| \leq t} f_\alpha x^\alpha = \bar{f}^\top v_t(x)$, where $\bar{f} := (f_\alpha)_\alpha \in \mathbb{R}^{\binom{n+t}{n}}$ is its vector of coefficients in the canonical basis. The *degree- t homogenization* \tilde{f} associated to $f \in \mathbb{R}[x]_t$ is a homogeneous polynomial of degree t in $n+1$ variables, defined by $\tilde{f}(x, x_{n+1}) := x_{n+1}^t f(x/x_{n+1})$. A *positive definite form* is a nonnegative homogeneous polynomial which is positive everywhere except at the origin.

For a given real-valued sequence $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$, let us define the *Riesz linear functional* $L_y : \mathbb{R}[x] \rightarrow \mathbb{R}$ by $f \mapsto L_y(f) := \sum_\alpha f_\alpha y_\alpha$, for all $f \in \mathbb{R}[x]$. We say that a real infinite (resp. finite) sequence $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ (resp. $(y_\alpha)_{\alpha \in \mathbb{N}_t^n}$) has a *representing measure* if there exists a finite Borel measure μ such that $y_\alpha = \int_{\mathbb{R}^n} x^\alpha d\mu(x)$ is satisfied for every $\alpha \in \mathbb{N}^n$ (resp. $\alpha \in \mathbb{N}_t^n$). In this case, $(y_\alpha)_{\alpha \in \mathbb{N}^n}$ is called the *moment sequence* of μ . Next, given $y = (y_\alpha)_{\alpha \in \mathbb{N}^n}$ and $d \in \mathbb{N}^*$, the *moment matrix* $M_d(y)$ of degree d associated to y is the real symmetric matrix of size $\binom{n+d}{d}$ defined by $M_d(y) := (y_{\alpha+\beta})_{\alpha, \beta \in \mathbb{N}_d^n}$. Let $g = \sum_\gamma g_\gamma x^\gamma \in \mathbb{R}[x]$. The localizing matrix $M_d(gy)$ of degree d associated with y and g is the real symmetric matrix of the size $\binom{n+d}{d}$ given by $M_d(gy) := (\sum_\gamma g_\gamma y_{\gamma+\alpha+\beta})_{\alpha, \beta \in \mathbb{N}_d^n}$.

3 Representation theorems and degree bounds

In this section, we derive representations of polynomials nonnegative on semialgebraic sets together with degree bounds. We extend these representations to the set of continuous functions being nonnegative on compact domains.

3.1 Polynomials nonnegative on general semialgebraic sets

We analyze the complexity of Putinar–Vasilescu’s Positivstellensatz [17] in the following theorem:

Theorem 2. (*Homogenized representation*) Let g_1, \dots, g_m be homogeneous polynomials of even degrees such that the semialgebraic set

$$S := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \dots, g_m(x) \geq 0\} \quad (22)$$

has nonempty interior. Let f be a homogeneous polynomial of degree $2d_f$ for some $d_f \in \mathbb{N}$ such that f is nonnegative on S . Then there exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that for all $\varepsilon > 0$, for all $k \in \mathbb{N}$ satisfying

$$k \geq \bar{c}\varepsilon^{-c}, \quad (23)$$

there exist homogeneous SOS polynomials $\sigma_0, \dots, \sigma_m$ such that

$$\deg(\sigma_0) = \deg(\sigma_1 g_1) = \dots = \deg(\sigma_m g_m) = 2(k + d_f) \quad (24)$$

and

$$\|x\|_2^{2k} (f + \varepsilon \|x\|_2^{2d_f}) = \sigma_0 + \sigma_1 g_1 + \dots + \sigma_m g_m. \quad (25)$$

In particular, if $m = 1$ and $g_1 = x_n^2 - \|x'\|_2^2$ with $x' := (x_1, \dots, x_{n-1})$, then $c = 65$. Moreover, each SOS polynomial σ_i involved in (25) can be chosen as the (single) square of a homogeneous polynomial, for $i \in [m]$.

The proof of Theorem 2 is postponed to Subsection 3.4. In Theorem 2, all polynomials are assumed to have even degrees. Next, we provide two corollaries where each polynomial can have odd degree.

The following corollary is a direct consequence of Theorem 2.

Corollary 1. (*Dehomogenized representation*) Let g_1, \dots, g_m be polynomials such that the semialgebraic set

$$S := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \dots, g_m(x) \geq 0\} \quad (26)$$

has nonempty interior. Let f be a polynomial nonnegative on S . Denote $d_f := \lfloor \deg(f)/2 \rfloor + 1$. Then there exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that for all $\varepsilon > 0$, for all $k \in \mathbb{N}$ satisfying

$$k \geq \bar{c}\varepsilon^{-c}, \quad (27)$$

there exist SOS polynomials $\sigma_0, \dots, \sigma_m$ such that

$$\deg(\sigma_0) \leq 2(k + d_f) \quad \text{and} \quad \deg(\sigma_j g_j) \leq 2(k + d_f), \quad j = 1, \dots, m, \quad (28)$$

and

$$\theta^k(f + \varepsilon \theta^{d_f}) = \sigma_0 + \sigma_1 g_1 + \dots + \sigma_m g_m, \quad (29)$$

where $\theta := 1 + \|x\|_2^2$. Moreover, if $m = 1$ and $g_1 = 1 - \|x\|_2^2$, then $c = 65$.

Proof. The proof of Corollary 1 is similar to the proof of [9, Theorems 4 and 5]. We recall the basic ingredients. Let \tilde{S} be a homogenized version of S , defined by

$$\tilde{S} := \{(x, x_{n+1}) \in \mathbb{R}^{n+1} : \tilde{g}_j(x, x_{n+1}) \geq 0, j \in [m]\}, \quad (30)$$

with $\tilde{g}_j(x, x_{n+1}) := x_{n+1}^{2d_{g_j}} g_j(x/x_{n+1})$ being the degree- $2d_{g_j}$ homogenization of g_j and $d_{g_j} := \lceil \deg(g_j)/2 \rceil$, for $j \in [m]$. Then the proof consists of three steps:

1. Prove that the degree- $2d_f$ homogenization of f , denoted by \tilde{f} , is nonnegative on \tilde{S} .
2. Use Theorem 2 to obtain a representation of \tilde{f} together with the degree bound on SOS polynomials.
3. Obtain a representation of f by evaluating the representation of \tilde{f} at $x_{n+1} = 1$.

To apply Theorem 2, we need to show that if S has nonempty interior, then \tilde{S} has nonempty interior. This statement holds since when a belongs to the interior of S , one has $\tilde{g}_j(a, 1) = g_j(a) > 0$, implying that $(a, 1)$ belongs to the interior of \tilde{S} . \square

Note that the ice cream constraint $x_{n+1}^2 - \|x\|_2^2$ is the degree-2 homogenization associated to the ball constraint $1 - \|x\|_2^2$.

3.2 Polynomials nonnegative on compact semialgebraic sets

The following corollary is deduced from Corollary 1.

Corollary 2. Let g_1, \dots, g_m be polynomials such that $g_1 := L - \|x\|_2^2$ for some $L > 0$ and the semialgebraic set

$$S := \{x \in \mathbb{R}^n : g_1(x) \geq 0, \dots, g_m(x) \geq 0\} \quad (31)$$

has nonempty interior. Let f be a polynomial nonnegative on S . Denote $d_f := \lfloor \deg(f)/2 \rfloor + 1$. Then there exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that for all $\varepsilon > 0$, for all $k \in \mathbb{N}$ satisfying

$$k \geq \bar{c}\varepsilon^{-c}, \quad (32)$$

there exist SOS polynomials $\sigma_0, \dots, \sigma_m$ such that

$$\deg(\sigma_0) \leq 2(k + d_f) \quad \text{and} \quad \deg(\sigma_j g_j) \leq 2(k + d_f), \quad j = 1, \dots, m, \quad (33)$$

and

$$(1 + \|x\|_2^2)^k(f + \varepsilon) = \sigma_0 + \sigma_1 g_1 + \dots + \sigma_m g_m. \quad (34)$$

Moreover, if $m = 1$ and $L = 1$, then $c = 65$.

Proof. By using [Corollary 1](#), there exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that for all $\varepsilon > 0$, for all $k \in \mathbb{N}$ satisfying $k \geq \bar{c}\varepsilon^{-c}$, there exist SOS polynomials $\sigma_0, \dots, \sigma_m$ such that

$$\deg(\sigma_0) \leq 2(k + d_f) \quad \text{and} \quad \deg(\sigma_j g_j) \leq 2(k + d_f), \quad j = 1, \dots, m. \quad (35)$$

and

$$\theta^k(f + \varepsilon \theta^{d_f}) = \sigma_0 + \sigma_1 g_1 + \dots + \sigma_m g_m, \quad (36)$$

where $\theta := 1 + \|x\|_2^2$. In addition,

$$(L + 1)^{d_f} - \theta^{d_f} = (L - \|x\|_2^2) \sum_{j=0}^{d_f-1} (L + 1)^{d_f-1-j} \theta^j = s_1 g_1, \quad (37)$$

where $s_1 = \sum_{j=0}^{d_f-1} (L + 1)^{d_f-1-j} \theta^j$ is an SOS polynomial of degree $2(d_f - 1)$. From this,

$$\begin{aligned} \theta^k[f + \varepsilon(L + 1)^{d_f}] &= \theta^k(f + \varepsilon \theta^{d_f}) + \varepsilon \theta^k[(L + 1)^{d_f} - \theta^{d_f}] \\ &= \sigma_0 + (\varepsilon s_1 \theta^k + \sigma_1) g_1 + \sum_{j=2}^m \sigma_j g_j, \end{aligned} \quad (38)$$

which yields the desired result. \square

Remark 1. In [Corollary 2](#), if $m = 1$ and $L > 0$, we still obtain $c = 65$. Indeed, up to a scaling, one can always assume $L = 1$.

Remark 2. We can apply the technique used in the proof of [Corollary 2](#), which consists of replacing the perturbation $\varepsilon \theta^{d_f}$ by ε , to represent polynomials nonnegative on \mathbb{R}^n . Let us consider an arbitrary large positive constant L and a polynomial f of degree $2d_f$ which is nonnegative on \mathbb{R}^n . Then, thanks to [\[9, Theorem 3.2\]](#), for any $\varepsilon > 0$, for all $k \in \mathbb{N}$ such that $k \geq \mathcal{O}(\varepsilon^{-1})$, $\theta^k(f + \varepsilon \theta^{d_f})$ is an SOS polynomial, so that $\theta^k(f + \varepsilon) = \sigma_0 + \sigma_1(L - \|x\|_2^2)$ for some SOS polynomials σ_i , $i = 0, 1$. This is the so-called “big ball trick”. This representation yields a linear convergence rate $\mathcal{O}(\varepsilon^{-1})$ for the minimization of polynomials on \mathbb{R}^n .

3.3 Preliminary material

This subsection presents some important lemmas that we use to prove the main results.

Given $\Omega \subset \mathbb{R}^n$, the distance of $a \in \mathbb{R}^n$ to Ω is denoted by $\text{dist}(a, \Omega)$. Denote by $B(a, r)$ (resp. $B^\circ(a, r)$) the closed (resp. open) ball centered at $a \in \mathbb{R}^n$ with radius $r > 0$.

We recall the [Lojasiewicz inequality](#) in the following lemma:

Lemma 1. Let $g : U \rightarrow \mathbb{R}$ be an analytic function on an open set $U \subset \mathbb{R}^n$ and $Z := \{x \in K : g(x) = 0\}$ for some compact set $K \subset U$. Then there exists $\alpha > 0$ and $C > 0$ such that

$$d(x, Z)^\alpha \leq C|g(x)|, \quad \forall x \in K.$$

As a consequence of [Lemma 1](#), we obtain the following result:

Lemma 2. ([\[2, Corollary 2.6.7\]](#)) Let $r > 0$ and the semialgebraic set $S := \{x \in \mathbb{R}^n : g_j(x) \geq 0, j \in [m]\}$, where g_1, \dots, g_m are polynomials. Then there exist positive constants α and C such that, for all x in $B(0, r)$,

$$\text{dist}(x, S)^\alpha \leq -C \min\{g_1(x), \dots, g_m(x), 0\}. \quad (39)$$

Given an open set $U \subset \mathbb{R}^n$ and a differentiable function $\varphi : U \rightarrow \mathbb{R}$, denote by $\nabla\varphi(x) = [\partial_{x_1}\varphi(x), \dots, \partial_{x_n}\varphi(x)]$ the gradient of φ at $x \in U$. Given $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, let $x' := (x_1, \dots, x_{n-1})$.

As mentioned in [13, Theorem 3.8], one can prove that $\alpha \in \{1, 2\}$ in Lemma 1 if g is a quadratic polynomial. The following lemma states an instance of this result:

Lemma 3. (*Łojasiewicz inequality with ice cream constraint*) Let $g := x_n^2 - \|x'\|_2^2$ and $Z := \{x \in \mathbb{R}^n : g(x) = 0\}$. Then for all $x \in \mathbb{R}^n$,

$$\text{dist}(x, Z)^2 \leq \frac{|g(x)|}{2}. \quad (40)$$

Proof. If $x \in Z$, both sides of (40) are zeros. Let $x \in \mathbb{R}^n \setminus Z$ be fixed. Then $d(x, Z)^2 = \min_y \{\|x - y\|_2^2 : g(y) = 0\}$. Assume that $(y, \mu) \in \mathbb{R}^n \times \mathbb{R}$ satisfies the Karush–Kuhn–Tucker conditions:

$$\begin{cases} \nabla_y \|x - y\|_2^2 = \mu \nabla_y g(y), \\ g(y) = 0. \end{cases} \quad (41)$$

The first condition of (41) implies that $2(x - y) = \mu \begin{bmatrix} -2y' \\ 2y_n \end{bmatrix}$, so $x' - y' = -\mu y'$ and $x_n - y_n = \mu y_n$. Assume that $\mu \notin \{1, -1\}$. Then $y' = \frac{x'}{1-\mu}$ and $y_n = \frac{x_n}{1+\mu}$. Since $g(y) = y_n^2 - \|y'\|_2^2 = 0$, $y_n = \pm \|y'\|_2$.

Let us consider the first case $y_n = \|y'\|_2$. Then $\frac{x_n}{1+\mu} = \frac{\|x'\|_2}{1-\mu}$. It implies that $\mu = \frac{x_n - \|x'\|_2}{x_n + \|x'\|_2}$. Note that $x_n \neq -\|x'\|_2$ since $g(x) \neq 0$. From this, $y' = \frac{(x_n + \|x'\|_2)x'}{2\|x'\|_2}$ and $y_n = \frac{x_n + \|x'\|_2}{2}$. Thus, $\|x - y\|_2^2 = \frac{(x_n - \|x'\|_2)^2}{2}$.

Similarly, if we consider the case $y_n = -\|y'\|_2$, then $\|x - y\|_2^2 = \frac{(x_n + \|x'\|_2)^2}{2}$.

Let us consider the case of $\mu \in \{1, -1\}$. Assume that $\mu = 1$. Then $x' = 0$ and $y_n = \frac{x_n}{2}$. From this and the fact that $0 = g(y) = y_n^2 - \|y'\|_2^2$, we obtain $\|y'\|_2^2 = \frac{x_n^2}{4}$. It implies that $\|x - y\|_2^2 = \|y'\|_2^2 + (x_n - y_n)^2 = \frac{x_n^2}{4} + \frac{x_n^2}{4} = \frac{x_n^2}{2} = \frac{(x_n - \|x'\|_2)^2}{2}$. Thus, $\|x - y\|_2^2 = \frac{(x_n - \|x'\|_2)^2}{2}$.

Similarly, if we consider the case $\mu = -1$, then $\|x - y\|_2^2 = \frac{(x_n + \|x'\|_2)^2}{2}$.

Thus,

$$\begin{aligned} d(x, Z)^2 &\leq \frac{1}{2} \min\{(x_n - \|x'\|_2)^2, (x_n + \|x'\|_2)^2\} \\ &\leq \frac{1}{2} \sqrt{(x_n - \|x'\|_2)^2 (x_n + \|x'\|_2)^2} \\ &= \frac{1}{2} |x_n^2 - \|x'\|_2^2| = \frac{|g(x)|}{2}, \end{aligned} \quad (42)$$

yielding (40). \square

A real-valued function $f : U \rightarrow \mathbb{R}$ for some $U \subset \mathbb{R}^n$ is called L -Lipschitz (or Lipschitz) continuous on $K \subset U$ if there exists a real $L > 0$ such that $|f(x) - f(y)| \leq L\|x - y\|_2$, for all $x, y \in K$. In this case, L is called the Lipschitz constant of f on K . Given an open set $U \subset \mathbb{R}^n$, a function $f : U \rightarrow \mathbb{R}$ is called locally Lipschitz continuous on $K \subset U$ if for every $x \in K$ there exists a neighborhood $W \subset U$ of x such that f is Lipschitz continuous on $W \cap K$.

The following lemma is similar in spirit to [12, Section 2.4, Lemma 2]:

Lemma 4. Given an open set $U \subset \mathbb{R}^n$, if the function $f : U \rightarrow \mathbb{R}$ is locally Lipschitz on a compact set $K \subset U$, then f is Lipschitz on K .

Proof. Since f is locally Lipschitz on K , for each $x \in K$ there is some $r_x > 0$ and $L_x > 0$ such that $B(x, r_x) \subset U$ and f is L_x -Lipschitz on $B(x, r_x) \cap K$. Then the sets $B(x, \frac{1}{2}r_x)$, $x \in K$ form an open cover of K . Due to the compactness of K , there exists a finite subsequence of $B(x, \frac{1}{2}r_x)$, $x \in K$ covering K . For convenience, denote these

by $B(x_k, \frac{1}{2}r_k)$ and $L_k := L_{x_k}$, $k \in [l]$. Let $M := \sup_{x \in K} |f(x)|$, $r := \frac{1}{2} \min_{k \in [l]} r_k$, $L_0 := \frac{2M}{r}$ and $L := \max\{L_0, L_k : k \in [l]\}$. Then L is a Lipschitz constant of f on K . To see this, pick $x, y \in K$. If $\|x - y\|_2 \geq r$ then we see that $\frac{|f(x) - f(y)|}{\|x - y\|_2} \leq \frac{2M}{r} = L_0 \leq L$. If $\|x - y\|_2 < r$, then for some x_k we have $x \in B(x_k, \frac{1}{2}r_k)$. Then $y \in B(x_k, r_k)$ and so $|f(x) - f(y)| \leq L_k \|x - y\|_2 \leq L \|x - y\|_2$. \square

Lemma 5. (Kirschbraun's theorem [7]) *If U is a subset of \mathbb{R}^n and $f : U \rightarrow \mathbb{R}$ is a Lipschitz continuous function, then there is a Lipschitz continuous function $F : \mathbb{R}^n \rightarrow \mathbb{R}$ that extends f and has the same Lipschitz constant as f . Moreover the extension is provided by*

$$F(x) := \inf_{u \in U} \{f(u) + L_f \|x - u\|_2\}, \quad (43)$$

where L_f is the Lipschitz constant of f on U .

Given $\Omega \subset \mathbb{R}^n$, denote by $C(\Omega)$ the set of continuous functions from Ω to \mathbb{R} . We recall basic properties of the multivariate Bernstein polynomials described, e.g., in [5, 6].

Definition 1. (Multivariate Bernstein polynomials) *Let $d \in \mathbb{N}^n$ and $f \in C([0, 1]^n)$. The polynomials*

$$B_{f,d}(x) := \sum_{k_1=0}^{d_1} \cdots \sum_{k_n=0}^{d_n} f\left(\frac{k_1}{d_1}, \dots, \frac{k_n}{d_n}\right) \prod_{j=1}^n \left[\binom{d_j}{k_j} x_j^{k_j} (1 - x_j)^{d_j - k_j} \right] \quad (44)$$

are called the multivariate Bernstein polynomials of f .

Note that $\deg(B_{f,d}) = \sum_{j \in [n]} d_j$ and the binomial identity implies

$$\sup_{x \in [0, 1]^n} |B_{f,d}(x)| \leq \sup_{x \in [0, 1]^n} |f(x)|. \quad (45)$$

Lemma 6. (Error bound) *If $f \in C([0, 1]^n)$ is L -Lipschitz, namely $|f(x) - f(y)| \leq L \|x - y\|_2$ on $[0, 1]^n$, then for all $d \in \mathbb{N}^n$, the inequality*

$$|B_{f,d}(x) - f(x)| \leq \frac{L}{2} \left(\sum_{j=1}^n \frac{1}{d_j} \right)^{\frac{1}{2}} \quad (46)$$

holds for all $x \in [0, 1]^n$.

Proof. Let $x \in [0, 1]^n$ be fixed. To simplify the notation, we define $b_j := \binom{d_j}{k_j} x_j^{k_j} (1 - x_j)^{d_j - k_j}$, for $j \in [n]$, and $k = (\frac{k_1}{d_1}, \dots, \frac{k_n}{d_n})$. Then, we have:

$$\begin{aligned} |B_{f,d}(x) - f(x)|^2 &\leq \left(\sum |f(k) - f(x)| b_1 \dots b_n \right)^2 \\ &\leq L^2 \left(\sum \|k - x\|_2 b_1 \dots b_n \right)^2 \\ &\leq L^2 \left(\sum \|k - x\|_2^2 b_1 \dots b_n \right) \left(\sum b_1 \dots b_n \right) \\ &= L^2 \sum_r \left(\sum_r \left(\frac{k_r}{d_r} - x_r \right)^2 b_1 \dots b_n \right) \\ &= L^2 \sum_r \left(\sum_r \left(\frac{k_r}{d_r} - x_r \right)^2 b_1 \dots b_n \right) \\ &= L^2 \sum_r \frac{x_r (1 - x_r)}{d_r} \\ &\leq L^2 \sum_r \frac{1}{4d_r}. \end{aligned} \quad (47)$$

For the first inequality, we have used the multinomial identity $1 = \prod_j [x_j + (1 - x_j)]^{d_j} = \sum b_1 \dots b_n$ and the triangle inequality. For the second one, we have used the fact that f is L -Lipschitz. For the third one, we use that for all $a, b \in \mathbb{R}^n$ and all $c \in \mathbb{R}_+^n$, one has $(\sum a_k b_k c_k)^2 \leq (\sum a_k^2 c_k) (\sum b_k^2 c_k)$. The last equality comes from the identities $\sum_{k_r=0}^{d_r} (k_r - d_r x_r)^2 b_r = d_r x_r (1 - x_r)$ and $\prod_{j \neq r} (\sum_{k_j=0}^{d_j} b_j) = 1$. The last inequality is obtained by noticing that we have $x_r (1 - x_r) \leq 1/4$. \square

Let $e := (1, \dots, 1) \in \mathbb{R}^n$. As a consequence of Lemma 6, we obtain the following result after a change of coordinates.

Lemma 7. If $f \in C([0, 1]^n)$ is L -Lipschitz, namely $|f(x) - f(y)| \leq L\|x - y\|_2$ on $[-1, 1]^n$, then for all $k \in \mathbb{N}^{\geq 1}$, the inequality

$$\left| B_{y \rightarrow f(2y-e), ke} \left(\frac{x+e}{2} \right) - f(x) \right| \leq L \left(\frac{n}{k} \right)^{\frac{1}{2}} \quad (48)$$

holds for all $x \in [-1, 1]^n$. Moreover, we have

$$\sup_{x \in [-1, 1]^n} |B_{y \rightarrow f(2y-e), ke} \left(\frac{x+e}{2} \right)| \leq \sup_{x \in [-1, 1]^n} |f(x)|. \quad (49)$$

Proof. Define $g : [0, 1]^n \rightarrow \mathbb{R}$ by $g(x) := f(2x - e)$. Let us compute a Lipschitz constant of g . With $x, y \in [0, 1]^n$, by the Lipschitz continuity of f , we have

$$\begin{aligned} |g(x) - g(y)| &= |f(2x - e) - f(2y - e)| \\ &\leq L\|2x - e - 2y + e\|_2 \\ &= 2L\|x - y\|_2. \end{aligned} \quad (50)$$

Then $2L$ is a Lipschitz constant of g . Let $k \in \mathbb{N}^{\geq 1}$. Using Lemma 6, we get that for all $x \in [0, 1]^n$,

$$|B_{g, ke}(x) - g(x)| \leq \frac{2L}{2} \left(\sum_{j=1}^n \frac{1}{k} \right)^{\frac{1}{2}} = L \left(\frac{n}{k} \right)^{\frac{1}{2}}. \quad (51)$$

Let $y \in [-1, 1]^n$. Then $\frac{y+e}{2} \in [0, 1]^n$ implies that

$$|B_{g, ke}(\frac{y+e}{2}) - f(y)| = |B_{g, ke}(\frac{y+e}{2}) - g(\frac{y+e}{2})| \leq L \left(\frac{n}{k} \right)^{\frac{1}{2}}. \quad (52)$$

yielding (48).

In addition, from (45),

$$\begin{aligned} \sup_{y \in [-1, 1]^n} |B_{g, ke}(\frac{y+e}{2})| &= \sup_{x \in [0, 1]^n} |B_{g, ke}(x)| \\ &\leq \sup_{x \in [0, 1]^n} |g(x)| \\ &= \sup_{y \in [-1, 1]^n} |g(\frac{y+e}{2})| = \sup_{y \in [-1, 1]^n} |f(y)|, \end{aligned} \quad (53)$$

which yields (49). \square

For each $h \in \mathbb{R}[x]$, let

$$\Theta(h) := \frac{\sup_{x \in \mathbb{S}^{n-1}} h(x)}{\inf_{x \in \mathbb{S}^{n-1}} h(x)}. \quad (54)$$

For later use recall the following theorem.

Lemma 8. (Reznick [19, Theorem 3.12]) Suppose that $p \in \mathbb{R}[x]$ is a positive definite form of degree $2d$, for some $d \in \mathbb{N}$. Then for all $k \in \mathbb{N}$ satisfying

$$k \geq \frac{2nd(2d-1)}{4 \log 2} \Theta(p) - \frac{n+2d}{2}, \quad (55)$$

$\|x\|_2^{2k} p$ is a homogeneous SOS polynomial of degree $2(k+d)$. *More precisely, $\|x\|_2^{2k} p$ is a sum of powers of linear forms.*

3.4 The proof of Theorem 2

Recall that $[l] := \{1, \dots, l\}$ for $l \in \mathbb{N}^{\geq 1}$. Given real value functions p, q , we use the notation $\{p * q\} = \{x \in \mathbb{R}^n : p(x) * q(x)\}$, where $*$ $\in \{=, \geq, \leq, >, <\}$. Given a real value function p on $\Omega \subset \mathbb{R}^n$, note $\|p\|_\Omega := \sup_{x \in \Omega} |p(x)|$. With $\Omega \subset \mathbb{R}^n$, denote by $\text{int}(\Omega)$ the interior of Ω .

Given $U, V \subseteq \mathbb{R}^n$ and $r \in \mathbb{R}$, note $U + V = \{u + v : u \in U, v \in V\}$ and $rU = \{ru : u \in U\}$. Given a function $f : U \rightarrow \mathbb{R}$ and $A \subset U \subset \mathbb{R}^n$ such that $A = -A$, f is called even on A if $f(-x) = f(x)$ for all $x \in A$. Denote by \mathbb{S}^{n-1} the unit sphere of \mathbb{R}^n .

To begin the proof, let us fix $\varepsilon > 0$. By assumption, $\deg(f) = 2d_f$, $\deg(g_j) = 2d_{g_j}$ for some $d, d_{g_j} \in \mathbb{N}$, for $j \in [m]$.

3.4.1 Construction of the positive weight functions

For $j \in [m]$, define

$$S_j := \{x \in \mathbb{R}^n : g_i(x) \geq 0, i \in [j]\}. \quad (56)$$

Obviously, we have $S_m = S$. Note $S_0 := \mathbb{R}^n$ and $f_m := f$.

We will prove that there exist functions $\bar{\varphi}_m : \mathbb{R}^n \rightarrow \mathbb{R}$ such that the following conditions hold:

1. $\bar{\varphi}_m$ is positive, even and bounded from above by $C_{\bar{\varphi}_m} = \bar{r}_m \varepsilon^{-r_m}$ on $B(0, \sqrt{n}+m)$ for some positive constants \bar{r}_m and r_m independent of ε .
2. $\bar{\varphi}_m$ is Lipschitz with Lipschitz constant $L_{\bar{\varphi}_m} = \bar{t}_j \varepsilon^{-t_m}$ for some positive constants \bar{t}_m and t_m independent of ε .
3. $f_{m-1} := f_m + \frac{\varepsilon}{2} - \bar{\varphi}_m^2 g_m$ satisfies:
 - (a) $f_{m-1} \geq 0$ on $S_{m-1} \cap B(0, \sqrt{n} + m - 1)$;
 - (b) $f_{m-1} \leq C_{f_{m-1}}$ on $B(0, \sqrt{n} + m)$, where $C_{f_{m-1}} = \bar{c}_{m-1} \varepsilon^{-c_{m-1}}$ for some positive constants \bar{c}_{m-1} and c_{m-1} independent of ε ;
 - (c) f_{m-1} is Lipschitz on $B(0, \sqrt{n} + m)$ with Lipschitz constant $L_{f_{m-1}} = \bar{l}_{m-1} \varepsilon^{-l_{m-1}}$ for some positive constants \bar{l}_{m-1} and l_{m-1} independent of ε .

Let

$$M_m := \inf_{x \in S_m \cap B(0, \sqrt{n}+m)} \frac{f(x) + \frac{\varepsilon}{2}}{g_m(x)}. \quad (57)$$

The constant M_m is a positive real number. Let $C_{g_m} = \|g_m\|_{B(0, \sqrt{n}+m)}$. We claim that $\frac{\varepsilon}{2C_{g_m}} < M_m < \infty$. Indeed, if z is a feasible solution of (57), $z \in S$ yielding $f(z) \geq 0$ so that

$$\frac{f(z) + \frac{\varepsilon}{2}}{g_m(z)} \geq \frac{\varepsilon}{2g_m(z)} \geq \frac{\varepsilon}{2C_{g_m}}. \quad (58)$$

From this, we have $M_m > \frac{\varepsilon}{2C_{g_m}}$. On the other hand, there exists $a \in \mathbb{R}^n$ such that $g_j(a) > 0$ for $j \in [m]$ since S has nonempty interior. For $j \in [m]$, since g_j is homogeneous, $a = 0$ yields $g_j(a) = 0$. It implies that $a \neq 0$. With $\bar{a} = \frac{a}{\|a\|_2} \in B(0, 1) \subset B(0, \sqrt{n} + m)$, we obtain $g_j(\bar{a}) > 0$ for $j \in [m]$ since

$$g_j(\bar{a}) = g_j\left(\frac{a}{\|a\|_2}\right) = \frac{g_j(a)}{\|a\|_2^{2d_{g_j}}} > 0, \forall j \in [m]. \quad (59)$$

Thus, \bar{a} is a feasible solution of (57) which yields

$$\frac{\varepsilon}{2C_{g_m}} \leq M_m \leq \frac{f(\bar{a}) + \frac{\varepsilon}{2}}{g_m(\bar{a})} \leq \frac{C_f + \frac{\varepsilon}{2}}{g_m(\bar{a})} < \infty, \quad (60)$$

where $C_f := \|f\|_{B(0, \sqrt{n}+m)}$.

Let $\psi_m : \mathbb{R}^n \rightarrow \mathbb{R}$ be the function defined by

$$\psi_m(x) := \begin{cases} \max\{M_m, \frac{f(x) + \frac{\varepsilon}{2}}{g_m(x)}\} & \text{if } g_m(x) < 0, \\ M_m & \text{otherwise.} \end{cases} \quad (61)$$

The function $f + \frac{\varepsilon}{2} - \psi_m g_m$ is nonnegative on $S_{m-1} \cap B(0, \sqrt{n} + m)$. Namely, we claim that

$$f + \frac{\varepsilon}{2} - \psi_m g_m \geq 0 \text{ on } S_{m-1} \cap B(0, \sqrt{n} + m). \quad (62)$$

Let $y \in S_{m-1} \cap B(0, \sqrt{n} + m)$. If $g_m(y) < 0$, then

$$\begin{aligned} f(y) + \frac{\varepsilon}{2} - \psi_m(y) g_m(y) &= f(y) + \frac{\varepsilon}{2} - g_m(y) \max\{M_m, \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)}\} \\ &\geq f(y) + \frac{\varepsilon}{2} - g_m(y) \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} = 0. \end{aligned} \quad (63)$$

Otherwise, $g_m(y) \geq 0$ gives

$$\begin{aligned} f(y) + \frac{\varepsilon}{2} - \psi_m(y) g_m(y) &= f(y) + \frac{\varepsilon}{2} - g_m(y) M_m \\ &\begin{cases} \geq f(y) + \frac{\varepsilon}{2} - g_m(y) \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} = 0 & \text{if } g_m(y) > 0, \\ = f(y) + \frac{\varepsilon}{2} \geq 0 & \text{if } g_m(y) = 0, \end{cases} \end{aligned} \quad (64)$$

since $y \in S$ is a feasible solution of (57).

The function ψ_m is positive, even on $B(0, \sqrt{n} + m)$ and continuous on $S_{m-1} \cap B(0, \sqrt{n} + m)$. It is easy to see that ψ_m is bounded from below by M_m and continuous on $B(0, \sqrt{n} + m) \setminus \{g_m = 0\}$ since the max function $(t_1, t_2) \mapsto \max\{t_1, t_2\}$ is continuous.

We claim that ψ_m is continuous on $S_{m-1} \cap B(0, \sqrt{n} + m) \cap \{g_m = 0\}$. Indeed, let us consider a sequence $(y_l)_l \subset S_{m-1} \cap B(0, \sqrt{n} + m) \cap \{g_m < 0\}$ such that $y_l \rightarrow \bar{y} \in S_{m-1} \cap B(0, \sqrt{n} + m) \cap \{g_m = 0\}$. Then $g_m(y_l) \rightarrow 0^-$ and $f(y_l) \rightarrow f(\bar{y}) \geq 0$ (since $\bar{y} \in S$) yielding that $\frac{f(y_l) + \frac{\varepsilon}{2}}{g_m(y_l)} \rightarrow -\infty$. It implies that $\max\{M_m, \frac{f(y_l) + \frac{\varepsilon}{2}}{g_m(y_l)}\} \rightarrow M_m$. Thus, $\psi_m = M_m$ on a sufficiently small neighborhood of any point in $S_{m-1} \cap B(0, \sqrt{n} + m) \cap \{g_m = 0\}$. On the other hand, ψ_m is even, i.e., $\psi_m(x) = \psi_m(-x)$ due to the fact that f, g_1, \dots, g_m are even and $B(0, \sqrt{n} + m) = -B(0, \sqrt{n} + m)$.

The upper bound of ψ_m depends on ε . It follows from (60) that $\psi_m = M_m$ on $B(0, \sqrt{n} + m) \cap \{g_m \geq 0\}$ and so is bounded from above by $\frac{f(\bar{a}) + \frac{\varepsilon}{2}}{g_m(\bar{a})}$.

Let us compute an upper bound of ψ_m on $S_{m-1} \cap B(0, \sqrt{n} + m) \cap \{g_m < 0\}$. Let $y \in S_{m-1} \cap B(0, \sqrt{n} + m)$ be such that $g_m(y) < 0$ and $\frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} > M_m$. Then $\psi_m(y) = \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)}$. By using the Łojasiewicz inequality (see Lemma 2), there exist $C_m > 0$ and $\alpha_m > 0$ depending on g_1, \dots, g_m such that for all $x \in S_{m-1} \cap B(0, \sqrt{n} + m) \cap \{g_m < 0\}$,

$$\text{dist}(x, S)^{\alpha_m} \leq -C_m \min\{g_1(x), \dots, g_m(x), 0\} = -C_m g_m(x). \quad (65)$$

Let $\delta_m = \frac{1}{C_m} (\frac{\varepsilon}{2L_f})^{\alpha_m}$, where L_f is a Lipschitz constant of f on $B(0, \sqrt{n} + m)$. Consider the following two cases:

- Case 1: $g_m(y) \leq -\delta_m < 0$. Then

$$\psi_m(y) = \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} = \frac{-f(y) - \frac{\varepsilon}{2}}{-g_m(y)} \leq \frac{C_f}{-g_m(y)} \leq \frac{C_f}{\delta_m} \leq C_m C_f \left(\frac{2L_f}{\varepsilon} \right)^{\alpha_m}. \quad (66)$$

- Case 2: $-\delta_m \leq g_m(y) < 0$. Let $z \in S$ such that $\text{dist}(y, S) = \|y - z\|_2$. Then (65) turns to $-f(y) \leq \frac{\varepsilon}{2}$ according to

$$\begin{aligned} -f(y) &\leq -f(z) + L_f \|y - z\|_2 \leq L_f \text{dist}(y, S) \\ &\leq L_f (-C_m g_m(y))^{\frac{1}{\alpha_m}} \leq L_f (C_m \delta_m)^{\frac{1}{\alpha_m}} = \frac{\varepsilon}{2}. \end{aligned} \quad (67)$$

From this, we obtain

$$M_m < \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} = \frac{-f(y) - \frac{\varepsilon}{2}}{-g_m(y)} \leq \frac{\frac{\varepsilon}{2} - \frac{\varepsilon}{2}}{-g_m(y)} = 0 < M_m. \quad (68)$$

The contradiction indicates that this case does not occur.

Thus, the bound is given as follows

$$\sup_{x \in S_{m-1} \cap B(0, \sqrt{n}+m)} \psi_m(x) \leq \max \left\{ \frac{f(\bar{a}) + \frac{\varepsilon}{2}}{g_m(\bar{a})}, C_m C_f \left(\frac{2L_f}{\varepsilon} \right)^{\alpha_m} \right\} =: C_{\psi_m}. \quad (69)$$

Moreover, we obtain the inclusion

$$S_{m-1} \cap B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq 0\} \subset \{g_m \leq -\delta_m\}, \quad (70)$$

where $\xi_m(x) = \frac{f(x) + \frac{\varepsilon}{2}}{g_m(x)}$. Let φ_m be the square root of ψ_m , i.e., $\varphi_m(x) := \sqrt{\psi_m(x)}$. Then φ_m is well-defined on $B(0, \sqrt{n}+m)$ since ψ_m is positive. Moreover, φ_m is finitely bounded from above on $S_{m-1} \cap B(0, \sqrt{n}+m)$ by $C_{\varphi_m} := \sqrt{C_{\psi_m}}$ and φ_m is continuous on $S_{m-1} \cap B(0, \sqrt{n}+m)$ since ξ_m is continuous on $S_{m-1} \cap B(0, \sqrt{n}+m)$.

The function φ_m is Lipschitz continuous on $S_{m-1} \cap B(0, \sqrt{n}+m-1)$. Keep in mind that ψ_m is defined by the constant function M_m and the function ξ_m . Since φ_m takes the constant value $\sqrt{M_m}$ on $B(0, \sqrt{n}+m) \setminus (\{\xi_m \geq M_m\} \cap \{g_m \leq 0\})$, φ_m is Lipschitz continuous on $B(0, \sqrt{n}+m) \setminus (\{\varphi_m \geq M_m\} \cap \{g_m \leq 0\})$ with zero Lipschitz constant.

On the other hand, $\varphi_m = \sqrt{\xi_m}$ on $B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq 0\}$. As a consequence of (70), we have

$$\begin{aligned} & S_{m-1} \cap B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq 0\} \\ &= S_{m-1} \cap B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}. \end{aligned} \quad (71)$$

It implies that

$$\varphi_m(x) = \begin{cases} \sqrt{\xi_m(x)} & \text{if } x \in S_{m-1} \cap B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}, \\ \sqrt{M_m} & \text{if } x \in (S_{m-1} \cap B(0, \sqrt{n}+m)) \setminus (\{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}). \end{cases} \quad (72)$$

The second equality is due to the fact that $\varphi_m = \sqrt{M_m}$ on $(S_{m-1} \cap B(0, \sqrt{n}+m)) \setminus (\{\xi_m \geq M_m\} \cap \{g_m \leq 0\})$ and

$$\begin{aligned} & (S_{m-1} \cap B(0, \sqrt{n}+m)) \setminus (\{\xi_m \geq M_m\} \cap \{g_m \leq 0\}) \\ &= (S_{m-1} \cap B(0, \sqrt{n}+m)) \setminus [S_{m-1} \cap B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq 0\}] \\ &= (S_{m-1} \cap B(0, \sqrt{n}+m)) \setminus [S_{m-1} \cap B(0, \sqrt{n}+m) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}] \\ &= (S_{m-1} \cap B(0, \sqrt{n}+m)) \setminus (\{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}). \end{aligned} \quad (73)$$

Let L_{g_m} be a Lipschitz constant of g_m on $B(0, \sqrt{n}+m)$. Set

$$w_m := \min \left\{ 1, \frac{\delta_m}{2L_{g_m}}, \frac{\varepsilon \delta_m^2}{8C_{g_m}[L_f C_{g_m} + (C_f + \frac{\varepsilon}{2})L_{g_m}]} \right\}. \quad (74)$$

and

$$W_m := (B(0, \sqrt{n}+m-1) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}) + w_m B(0, 1). \quad (75)$$

Then $B(0, \sqrt{n}+m-1) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\} \subset W_m$. Next, we prove that

$$W_m \subset B(0, \sqrt{n}+m) \cap \{\xi_m \geq \frac{M_m}{2}\} \cap \{g_m \leq -\frac{\delta_m}{2}\}. \quad (76)$$

Let $y \in W_m$. Then $y = z + w_m u$ for some $z \in S_{m-1} \cap B(0, \sqrt{n}+m-1) \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}$ and for some $u \in B(0, 1)$. Combining $\|z\|_2 \leq \sqrt{n}+m-1$, $0 < w_m < 1$ and $\|u\|_2 \leq 1$, one has $\|y\|_2 \leq \|z\|_2 + w_m \|u\|_2 \leq \sqrt{n}+m$, yielding $y \in B(0, \sqrt{n}+m)$. Since $g_m(z) \leq -\delta_m$, we have

$$g_m(y) \leq g_m(z) + L_{g_m} \|y-z\|_2 \leq -\delta_m + L_{g_m} w_m \|u\|_2 \leq -\delta_m + L_{g_m} \frac{\delta_m}{2L_{g_m}} \leq -\frac{\delta_m}{2}. \quad (77)$$

Thus $y \in \{g_m \leq -\frac{\delta_m}{2}\}$. This in turn implies

$$\begin{aligned}
& |\xi_m(y) - \xi_m(z)| \\
&= \left| \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} - \frac{f(z) + \frac{\varepsilon}{2}}{g_m(z)} \right| \\
&= \frac{|(f(y) + \frac{\varepsilon}{2})g_m(z) - (f(z) + \frac{\varepsilon}{2})g_m(y)|}{|g_m(y)g_m(z)|} \\
&\leq \frac{2}{\delta_m^2} |(f(y) + \frac{\varepsilon}{2} - f(z) - \frac{\varepsilon}{2})g_m(z) + (f(z) + \frac{\varepsilon}{2})(g_m(z) - g_m(y))| \quad (78) \\
&\leq \frac{2}{\delta_m^2} [|f(y) - f(z)||g_m(z)| + (|f(z)| + \frac{\varepsilon}{2})|g_m(z) - g_m(y)|] \\
&\leq \frac{2}{\delta_m^2} [L_f \|y - z\|_2 C_{g_m} + (C_f + \frac{\varepsilon}{2}) L_{g_m} \|z - y\|_2] \\
&\leq \frac{2}{\delta_m^2} [L_f C_{g_m} + (C_f + \frac{\varepsilon}{2}) L_{g_m}] w_m \|u\|_2 \leq \frac{\varepsilon}{4C_{g_m}} \leq \frac{M_m}{2}.
\end{aligned}$$

Since $\xi_m(z) \geq M_m$, we obtain $\xi_m(y) \geq \xi_m(z) - |\xi_m(y) - \xi_m(z)| \geq M_m - \frac{M_m}{2} = \frac{M_m}{2}$, yielding $y \in \{\xi_m \geq \frac{M_m}{2}\}$, which concludes the proof of (76) and ensures that $\sqrt{\xi_m}$ is well-defined on W_m .

Let us prove that $\sqrt{\xi_m}$ is Lipschitz on W_m . Let $y, z \in W_m$ such that $y \neq z$. Then

$$\begin{aligned}
& \frac{|\sqrt{\xi_m(y)} - \sqrt{\xi_m(z)}|}{\|y - z\|_2} \\
&= \frac{|\xi_m(y) - \xi_m(z)|}{\|y - z\|_2 (\sqrt{\xi_m(y)} + \sqrt{\xi_m(z)})} \\
&\leq \frac{\left| \frac{f(y) + \frac{\varepsilon}{2}}{g_m(y)} - \frac{f(z) + \frac{\varepsilon}{2}}{g_m(z)} \right|}{2\sqrt{\frac{M_m}{2}} \|y - z\|_2} \\
&\leq \frac{|(f(y) + \frac{\varepsilon}{2})g_m(z) - (f(z) + \frac{\varepsilon}{2})g_m(y)|}{2g_m(y)g_m(z)\sqrt{\frac{\varepsilon}{4C_{g_m}}} \|y - z\|_2} \\
&\leq \frac{2|(f(y) + \frac{\varepsilon}{2})g_m(z) - (f(z) + \frac{\varepsilon}{2})g_m(y)|}{\delta_m^2 \sqrt{\frac{\varepsilon}{4C_{g_m}}} \|y - z\|_2} \quad (79) \\
&= \frac{2|(f(y) + \frac{\varepsilon}{2} - f(z) - \frac{\varepsilon}{2})g_m(z) + (f(z) + \frac{\varepsilon}{2})(g_m(z) - g_m(y))|}{\delta_m^2 \sqrt{\frac{\varepsilon}{4C_{g_m}}} \|y - z\|_2} \\
&\leq \frac{2[|f(y) - f(z)||g_m(z)| + (|f(z)| + \frac{\varepsilon}{2})|g_m(z) - g_m(y)|]}{\delta_m^2 \sqrt{\frac{\varepsilon}{4C_{g_m}}} \|y - z\|_2} \\
&\leq \frac{2[L_f \|y - z\|_2 C_{g_m} + (C_f + \frac{\varepsilon}{2}) L_{g_m} \|z - y\|_2]}{\delta_m^2 \sqrt{\frac{\varepsilon}{4C_{g_m}}} \|y - z\|_2} \\
&\leq \frac{2[L_f C_{g_m} + (C_f + \frac{\varepsilon}{2}) L_{g_m}]}{\delta_m^2 \sqrt{\frac{\varepsilon}{4C_{g_m}}}} =: L_{\sqrt{\xi_m}},
\end{aligned}$$

Thus, $L_{\sqrt{\xi_m}}$ is a Lipschitz constant of $\sqrt{\xi_m}$ on W_m .

Set $K := S_{m-1} \cap B(0, \sqrt{n} + m - 1)$, $K_1 := K \cap \{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\}$ and $K_2 := K \setminus (\{\xi_m \geq M_m\} \cap \{g_m \leq -\delta_m\})$. Note that $K = K_1 \cup K_2$ and $K_1 \cap K_2 = \emptyset$. From (72), $\varphi_m = \sqrt{\xi_m}$ on K_1 and $\varphi_m = \sqrt{M_m}$ on K_2 .

To conclude that φ_m is Lipschitz on K according to Lemma 4 (see Figure 1), it is sufficient to prove that φ_m is locally Lipschitz on K .

Explicitly, we will show that for all $z \in K$, φ_m is Lipschitz on $B(z, \frac{w_m}{2}) \cap K$ with Lipschitz constant $L_{\sqrt{\xi_m}}$. Let $z \in K$. Let $u, v \in B(z, \frac{w_m}{2}) \cap K$ and consider the following cases:

- Case 1: $u, v \in K_1$. Then $u, v \in W_m$ by definition of W_m . Moreover, $\varphi_m(u) = \sqrt{\xi_m(u)}$ and $\varphi_m(v) = \sqrt{\xi_m(v)}$. In this case, by the Lipschitz continuity of $\sqrt{\xi_m}$ on W_m ,

$$|\varphi_m(u) - \varphi_m(v)| = |\sqrt{\xi_m(u)} - \sqrt{\xi_m(v)}| \leq L_{\sqrt{\xi_m}} \|u - v\|_2. \quad (80)$$

- Case 2: $u, v \in K_2$. In this case, $\varphi_m(u) = \varphi_m(v) = \sqrt{M_m}$, so that

$$|\varphi_m(u) - \varphi_m(v)| = 0 \leq L_{\sqrt{\xi_m}} \|u - v\|_2. \quad (81)$$

- Case 3: $u \in K_1$ and $v \in K_2$. We claim that $B(z, \frac{w_m}{2}) \subset W_m$. Let $q \in B(z, \frac{w_m}{2})$. Then $\|q - u\|_2 \leq \|q - z\|_2 + \|z - u\|_2 \leq w_m$ yielding $q \in u + w_m B(0, 1) \subset K_1 + w_m B(0, 1) \subset W_m$. Then $u, v \in B(z, \frac{w_m}{2}) \subset W_m$. Moreover, $\varphi_m(u) = \sqrt{\xi_m(u)}$

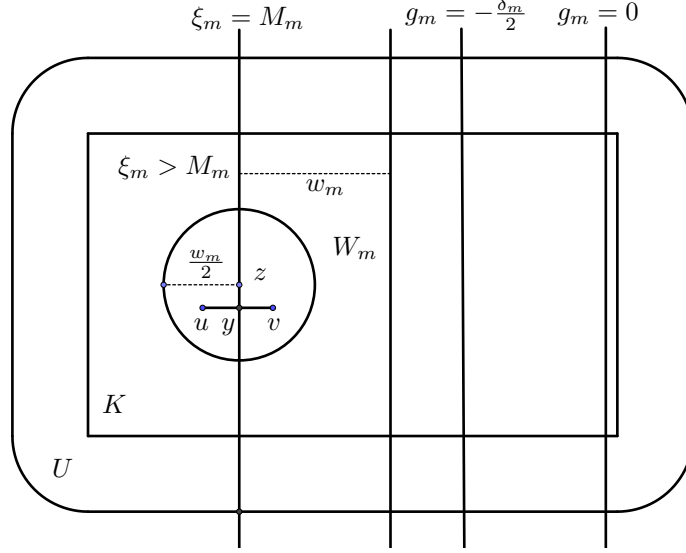


Figure 1: Illustration for the proof of the Lipschitz continuity of φ_m on K (rectangle). Here $K = S_{m-1} \cap B(0, \sqrt{n} + m - 1)$ and $U = K + \frac{w_m}{2} B^\circ(0, 1)$ with the notation of Lemma 4.

and $\varphi_m(v) = \sqrt{M_m}$. According to the continuity of ξ_m on $B(z, \frac{w_m}{2}) \subset W_m$ and the convexity of $B(z, \frac{w_m}{2})$, there exists $y \in B(z, \frac{w_m}{2}) \cap \{\xi_m = M_m\} \cap \{tu + (1-t)v : t \in [0, 1]\}$. Then with $y = \lambda u + (1-\lambda)v$ for some $\lambda \in [0, 1]$, we have

$$\begin{aligned}
 |\varphi_m(u) - \varphi_m(v)| &\leq |\varphi_m(u) - \varphi_m(y)| + |\varphi_m(y) - \varphi_m(v)| \\
 &\leq |\sqrt{\xi_m(u)} - \sqrt{\xi_m(y)}| + |\sqrt{M_m} - \sqrt{M_m}| \\
 &\leq L_{\sqrt{\xi_m}} \|u - y\|_2 \\
 &\leq L_{\sqrt{\xi_m}} \|u - \lambda u - (1-\lambda)v\|_2 \\
 &\leq L_{\sqrt{\xi_m}} (1-\lambda) \|u - v\|_2 \leq L_{\sqrt{\xi_m}} \|u - v\|_2.
 \end{aligned} \tag{82}$$

From the proof of Lemma 4, the Lipschitz constant of φ_m on K is given by

$$L_{\bar{\varphi}_m} := \max \left\{ \frac{4C_{\varphi_m}}{w_m}, L_{\sqrt{\xi_m}} \right\}, \tag{83}$$

Here we have covered K by a finite sequence of balls with radii $\frac{w_m}{2}$ and centers lying on K .

The function φ_m has a Lipschitz continuous extension $\bar{\varphi}_m$. Let $\bar{\varphi}_m : \mathbb{R}^n \rightarrow \mathbb{R}$ be the function defined by

$$\bar{\varphi}_m(x) := \inf_y \{ \varphi_m(y) + L_{\bar{\varphi}_m} \|x - y\|_2 : y \in S_{m-1} \cap B(0, \sqrt{n} + m - 1) \}. \tag{84}$$

By Kirszbraun's theorem (stated in Lemma 5), $\bar{\varphi}_m$ is Lipschitz continuous with Lipschitz constant $L_{\bar{\varphi}_m}$ and $\bar{\varphi}_m = \varphi_m$ on $S_{m-1} \cap B(0, \sqrt{n} + m - 1)$.

The function $\bar{\varphi}_m$ is even, positive and has a finite upper bound on $B(0, \sqrt{n} + m)$ depending on ε . Let us prove that $\bar{\varphi}_m$ is even. Consider

$$\bar{\varphi}_m(-x) = \inf_y \{ \varphi_m(y) + L_{\bar{\varphi}_m} \| -x - y \|_2 : y \in S_{m-1} \cap B(0, \sqrt{n} + m - 1) \}. \quad (85)$$

Let y be any feasible solution of (85). Since g_1, \dots, g_{m-1} are even, $S_{m-1} \cap B(0, \sqrt{n} + m)$ is symmetric, i.e., $S_{m-1} \cap B(0, \sqrt{n} + m) = -S_{m-1} \cap B(0, \sqrt{n} + m)$, it turns out that $-y$ is a feasible solution of (85). Thus,

$$\begin{aligned} \bar{\varphi}_m(-x) &= \inf_{-y} \{ \varphi_m(-y) + L_{\bar{\varphi}_m} \| -x + y \|_2 : -y \in S_{m-1} \cap B(0, \sqrt{n} + m - 1) \} \\ &= \inf_y \{ \varphi_m(y) + L_{\bar{\varphi}_m} \| y - x \|_2 : y \in S_{m-1} \cap B(0, \sqrt{n} + m - 1) \} = \bar{\varphi}_m(x), \end{aligned} \quad (86)$$

where the latter inequality is due to the fact that φ_m is even (since ξ_m, g_m are even). From this, $\bar{\varphi}_m$ is even. It is not hard to show that $\bar{\varphi}_m \geq \sqrt{M_m}$ since $\varphi_m \geq \sqrt{M_m}$.

Let us estimate the upper bound of $\bar{\varphi}_m$ on $B(0, \sqrt{n} + m)$. Let $x \in B(0, \sqrt{n} + m)$ and $y \in S_{m-1} \cap B(0, \sqrt{n} + m - 1)$. From (84), we get

$$\bar{\varphi}_m(x) \leq \varphi_m(y) + L_{\bar{\varphi}_m} \|x - y\|_2 \leq C_{\varphi_m} + (2(\sqrt{n} + m) - 1)L_{\bar{\varphi}_m} =: C_{\bar{\varphi}_m}. \quad (87)$$

Thus,

$$\sup_{x \in B(0, \sqrt{n} + m)} \bar{\varphi}_m(x) \leq C_{\bar{\varphi}_m}. \quad (88)$$

Set $f_{m-1} := f + \frac{\varepsilon}{2} - \bar{\varphi}_m^2 g_m$.

From (62) and since $\bar{\varphi}_m = \sqrt{\psi_m}$ on $S_{m-1} \cap B(0, \sqrt{n} + m - 1)$, $f_{m-1} \geq 0$ on $S_{m-1} \cap B(0, \sqrt{n} + m - 1)$. Since $\bar{\varphi}_m$ is Lipschitz continuous, f_{m-1} is Lipschitz continuous on $B(0, \sqrt{n} + m)$.

A bound and a Lipschitz constant of f_{m-1} on $B(0, \sqrt{n} + m)$ both depend on ε . Let us compute an upper bound of $|f_{m-1}|$ on $B(0, \sqrt{n} + m)$. Let $y \in B(0, \sqrt{n} + m)$. Then

$$|f_{m-1}(y)| \leq |f(y)| + \frac{\varepsilon}{2} + \bar{\varphi}_m(y)^2 |g_m(y)| \leq C_f + \frac{\varepsilon}{2} + C_{g_m} C_{\bar{\varphi}_m}^2 =: C_{f_{m-1}}. \quad (89)$$

Thus,

$$\|f_{m-1}\|_{B(0, \sqrt{n} + m)} \leq C_{f_{m-1}}. \quad (90)$$

We now estimate the Lipschitz constant of f_{m-1} on $B(0, \sqrt{n} + m)$. Let $y, z \in B(0, \sqrt{n} + m)$ such that $y \neq z$. Then

$$\begin{aligned} & \frac{|f_{m-1}(y) - f_{m-1}(z)|}{\|y - z\|_2} \\ & \leq \frac{|f(y) - f(z)| + |\bar{\varphi}_m(y)^2 g_m(y) - \bar{\varphi}_m(z)^2 g_m(z)|}{\|y - z\|_2} \\ & \leq L_f + \frac{|\bar{\varphi}_m(y)^2 g_m(y) - \bar{\varphi}_m(z)^2 g_m(z)|}{\|y - z\|_2} \\ & \quad + \frac{|\bar{\varphi}_m(z)^2 g_m(y) - \bar{\varphi}_m(z)^2 g_m(z)|}{\|y - z\|_2} \\ & = L_f + \frac{|g_m(y)| |\bar{\varphi}_m(y) + \bar{\varphi}_m(z)| |\bar{\varphi}_m(y) - \bar{\varphi}_m(z)| + \bar{\varphi}_m(z)^2 |g_m(y) - g_m(z)|}{\|y - z\|_2} \\ & \leq L_f + \frac{2C_{g_m} C_{\bar{\varphi}_m} L_{\bar{\varphi}_m} \|y - z\|_2 + C_{\bar{\varphi}_m}^2 L_{g_m} \|y - z\|_2}{\|y - z\|_2} \\ & = L_f + 2C_{g_m} L_{\bar{\varphi}_m} C_{\bar{\varphi}_m} + L_{g_m} C_{\bar{\varphi}_m}^2 =: L_{f_{m-1}}. \end{aligned} \quad (91)$$

Then, $L_{f_{m-1}}$ is a Lipschitz constant of f_{m-1} on $B(0, \sqrt{n} + m)$.

Notice that $C_{\bar{\varphi}_m}, L_{\bar{\varphi}_m}, C_{f_{m-1}}, L_{f_{m-1}}$ are obtained by composing finitely many times the following operators: “+”, “−”, “×”, “÷”, “| · |”, “ $(x_1, x_2) \mapsto \max\{x_1, x_2\}$ ”, “ $(x_1, x_2) \mapsto \min\{x_1, x_2\}$ ”, “ $(\cdot)^{\alpha_m}$ ” and “ $\sqrt{\cdot}$ ”, where all arguments possibly depend on ε . Without loss of generality we can assume $C_{\bar{\varphi}_m} = \bar{r}_m \varepsilon^{-r_m}$, $L_{\bar{\varphi}_m} = \bar{t}_m \varepsilon^{-t_m}$, $C_{f_{m-1}} = \bar{c}_{m-1} \varepsilon^{-c_{m-1}}$, $L_{f_{m-1}} = \bar{l}_{m-1} \varepsilon^{-l_{m-1}}$ for some $\bar{r}_m, r_m, \bar{t}_m, t_m, \bar{c}_{m-1}, c_{m-1}, \bar{l}_{m-1}, l_{m-1}$ large enough and independent of ε .

Backward induction. Repeating the above process (after replacing f_j by f_{j-1}) several times, we obtain functions $\bar{\varphi}_j : \mathbb{R}^n \rightarrow \mathbb{R}$, $j = m, m-1, \dots, 1$, such that,

1. $\bar{\varphi}_j$ is positive, even and bounded from above by $C_{\bar{\varphi}_j} = \bar{r}_j \varepsilon^{-r_j}$ on $B(0, \sqrt{n} + j)$ for some positive constants \bar{r}_j and r_j independent of ε .
2. $\bar{\varphi}_j$ is Lipschitz with Lipschitz constant $L_{\bar{\varphi}_j} = \bar{t}_j \varepsilon^{-t_j}$ for some positive constants \bar{t}_j and t_j independent of ε .
3. $f_{j-1} := f_j + \frac{\varepsilon}{2^{m-j+1}} - \bar{\varphi}_j^2 g_j$ satisfies:
 - (a) $f_{j-1} \geq 0$ on $S_{j-1} \cap B(0, \sqrt{n} + j - 1)$;
 - (b) $f_{j-1} \leq C_{f_{j-1}}$ on $B(0, \sqrt{n} + j)$, where $C_{f_{j-1}} = \bar{c}_{j-1} \varepsilon^{-c_{j-1}}$ for some positive constants \bar{c}_{j-1} and c_{j-1} independent of ε ;
 - (c) f_{j-1} is Lipschitz on $B(0, \sqrt{n} + j)$ with Lipschitz constant $L_{f_{j-1}} = \bar{l}_{j-1} \varepsilon^{-l_{j-1}}$ for some positive constants \bar{l}_{j-1} and l_{j-1} independent of ε .

Then

$$\begin{aligned}
f_0 &= f_1 + \frac{\varepsilon}{2^m} - \bar{\varphi}_1^2 g_1 \\
&= \left(f_2 + \frac{\varepsilon}{2^{m-1}} - \bar{\varphi}_2^2 g_2 \right) + \frac{\varepsilon}{2^m} - \bar{\varphi}_1^2 g_1 \\
&= f_2 + \left(\frac{\varepsilon}{2^{m-1}} + \frac{\varepsilon}{2^m} \right) - \bar{\varphi}_2^2 g_2 - \bar{\varphi}_1^2 g_1 \\
&= \dots = f_m + \varepsilon \sum_{i=1}^m \frac{1}{2^i} - \sum_{i=1}^m \bar{\varphi}_i^2 g_i \\
&= f + \frac{\varepsilon}{2} \frac{1 - \frac{1}{2^m}}{1 - \frac{1}{2}} - \sum_{i=1}^m \bar{\varphi}_i^2 g_i \\
&= f + \varepsilon \left(1 - \frac{1}{2^m} \right) - \sum_{i=1}^m \bar{\varphi}_i^2 g_i.
\end{aligned} \tag{92}$$

From this and since $f_0 \geq 0$ on $S_0 \cap B(0, \sqrt{n}) = B(0, \sqrt{n}) \supset [-1, 1]^n$, we obtain

$$f + \varepsilon - \sum_{i=1}^m \bar{\varphi}_i^2 g_i \geq \frac{\varepsilon}{2^m} \text{ on } [-1, 1]^n. \tag{93}$$

3.4.2 Polynomial approximations for the weight functions

Approximating with Bernstein polynomials. For each $i \in [m]$, we now approximate $\bar{\varphi}_i$ on $[-1, 1]^n$ with the following Bernstein polynomials:

$$B_i^{(d)}(x) = B_{y \rightarrow \bar{\varphi}_i(2y-e), de} \left(\frac{x+e}{2} \right), \quad d \in \mathbb{N}, \tag{94}$$

with $e = (1, \dots, 1) \in \mathbb{R}^n$. By using Lemma 7, for all $x \in [-1, 1]^n$, for $i \in [m]$,

$$|B_i^{(d)}(x) - \bar{\varphi}_i(x)| \leq L_{\bar{\varphi}_i} \left(\frac{n}{d} \right)^{\frac{1}{2}}, \quad d \in \mathbb{N}, \tag{95}$$

and the following inequality holds for all $x \in [-1, 1]^n$, for $i \in [m]$:

$$|B_i^{(d)}(x)| \leq \sup_{x \in [-1, 1]^n} |\bar{\varphi}_i(x)| \leq C_{\bar{\varphi}_i}. \tag{96}$$

For $i \in [m]$, let

$$d_i := 2u_i \quad \text{with} \quad u_i = \left\lceil \frac{2C_{g_i}^2 C_{\bar{\varphi}_i}^2 n L_{\bar{\varphi}_i}^2 (m+1)^2 2^{2m}}{\varepsilon^2} \right\rceil, \tag{97}$$

where $C_{g_i} := \|g_i\|_{B(0, \sqrt{n}+i)}$, for $i \in [m]$. Then for all $x \in [-1, 1]^n$,

$$\begin{aligned}
|B_i^{(d_i)}(x) - \bar{\varphi}_i(x)| &\leq L_{\bar{\varphi}_i} \left(\frac{n}{d_i} \right)^{\frac{1}{2}} \\
&\leq L_{\bar{\varphi}_i} \left(\frac{n}{\frac{4C_{g_i}^2 C_{\bar{\varphi}_i}^2 n L_{\bar{\varphi}_i}^2 (m+1)^2 2^{2m}}{\varepsilon^2}} \right)^{\frac{1}{2}} \\
&= \frac{\varepsilon}{2C_{g_i} C_{\bar{\varphi}_i} (m+1) 2^m}.
\end{aligned} \tag{98}$$

Converting to homogeneous approximations. For $i \in [m]$, we write $B_i^{(d_i)} = \sum_{j=0}^{nd_i} h_i^{(j)}$ such that $h_i^{(j)}$ is a homogeneous polynomial with $\deg(h_i^{(j)}) = j$. Set $p_i := \frac{1}{2}[B_i^{(d_i)}(x) + B_i^{(d_i)}(-x)]$, for $i \in [m]$. Then $p_i = \sum_{t=0}^{nu_i} h_i^{(2t)}$, for $i \in [m]$, since $h_i^{(j)}(x) = h_i^{(j)}(-x)$ if j is even and $h_i^{(j)}(x) = -h_i^{(j)}(-x)$ otherwise. Since $\bar{\varphi}_i$ is even, $\bar{\varphi}_i(x) = \frac{1}{2}[\bar{\varphi}_i(x) + \bar{\varphi}_i(-x)]$. It implies that for $x \in [-1, 1]^n$, for $i \in [m]$,

$$\begin{aligned} |p_i(x) - \bar{\varphi}_i(x)| &= \left| \frac{1}{2}[B_i^{(d_i)}(x) + B_i^{(d_i)}(-x)] - \frac{1}{2}[\bar{\varphi}_i(x) + \bar{\varphi}_i(-x)] \right| \\ &\leq \frac{1}{2}|B_i^{(d_i)}(x) - \bar{\varphi}_i(x)| + \frac{1}{2}|B_i^{(d_i)}(-x) - \bar{\varphi}_i(-x)| \\ &\leq \frac{\varepsilon}{4C_{g_i}C_{\bar{\varphi}_i}(m+1)2^m} + \frac{\varepsilon}{4C_{g_i}C_{\bar{\varphi}_i}(m+1)2^m} = \frac{\varepsilon}{2C_{g_i}C_{\bar{\varphi}_i}(m+1)2^m}. \end{aligned} \quad (99)$$

and

$$|p_i(x)| \leq \frac{1}{2}(|B_i^{(d_i)}(x)| + |B_i^{(d_i)}(-x)|) \leq \frac{1}{2}(C_{\bar{\varphi}_i} + C_{\bar{\varphi}_i}) = C_{\bar{\varphi}_i}. \quad (100)$$

Set $q_i := \sum_{t=0}^{nu_i} h_i^{(2t)} \|x\|_2^{2(nu_i-t)}$. Then q_i is a homogeneous polynomial of degree $2nu_i$ and $q_i = p_i$ on \mathbb{S}^{n-1} , for $i \in [m]$. Thus for $i \in [m]$, $|q_i(x) - \bar{\varphi}_i(x)| \leq \frac{\varepsilon}{2C_{g_i}C_{\bar{\varphi}_i}(m+1)2^m}$ and $|q_i(x)| \leq C_{\bar{\varphi}_i}$, for all $x \in \mathbb{S}^{n-1}$. From these and (93), for all $x \in \mathbb{S}^{n-1}$,

$$\begin{aligned} &f(x) + \varepsilon - \sum_{i=1}^m q_i(x)^2 g_i(x) \\ &= f(x) + \varepsilon - \sum_{i=1}^m \bar{\varphi}_i(x)^2 g_i(x) + \sum_{i=1}^m g_i(x)[\bar{\varphi}_i(x)^2 - q_i(x)^2] \\ &\geq \frac{\varepsilon}{2^m} - \sum_{i=1}^m |g_i(x)| |\bar{\varphi}_i(x) + q_i(x)| |\bar{\varphi}_i(x) - q_i(x)| \\ &\geq \frac{\varepsilon}{2^m} - \sum_{i=1}^m C_{g_i}(x) (|\bar{\varphi}_i(x)| + |q_i(x)|) \frac{\varepsilon}{2C_{g_i}C_{\bar{\varphi}_i}(m+1)2^m} \\ &\geq \frac{\varepsilon}{2^m} - \sum_{i=1}^m 2C_{g_i}C_{\bar{\varphi}_i} \frac{\varepsilon}{2C_{g_i}C_{\bar{\varphi}_i}(m+1)2^m} \\ &= \frac{\varepsilon}{2^m} - \frac{m\varepsilon}{(m+1)2^m} = \frac{\varepsilon}{(m+1)2^m}. \end{aligned} \quad (101)$$

Moreover, for all $x \in \mathbb{S}^{n-1}$,

$$f(x) + \varepsilon - \sum_{i=1}^m q_i(x)^2 g_i(x) \leq C_f + \varepsilon + \sum_{i=1}^m C_{\bar{\varphi}_i}^2 C_{g_i} =: C_F. \quad (102)$$

3.4.3 Applying the global positivity certificate

Set $D := \max_{i \in [m]} \{2nu_i + d_{g_i}\}$ and

$$F = \|x\|_2^{2(D-d_f)} (f + \varepsilon \|x\|_2^{2d_f}) - \sum_{i=1}^m g_i q_i^2 \|x\|_2^{2(D-2nu_i-d_{g_i})}. \quad (103)$$

Then F is a homogeneous polynomial of degree $2D$ and for all $x \in \mathbb{S}^{n-1}$,

$$C_F \geq F(x) = f(x) + \varepsilon - \sum_{i=1}^m q_i(x)^2 g_i(x) \geq \frac{\varepsilon}{(m+1)2^m}. \quad (104)$$

It implies that F is a positive definite form of degree $2D$ with $\inf_{x \in \mathbb{S}^{n-1}} F(x) \geq \frac{\varepsilon}{(m+1)2^m}$ and $\sup_{x \in \mathbb{S}^{n-1}} F(x) \leq C_F$. There is no loss of generality in assuming $C_F = b\varepsilon^{-b}$ for some large enough $b > 0$ independent of ε . Similarly assume that $D \geq d\varepsilon^{-d}$ for some large enough $d > 0$ independent of ε . From this,

$$\Theta(F) \leq \frac{b\varepsilon^{-b}}{\frac{\varepsilon}{(m+1)2^m}} = b(m+1)2^m \varepsilon^{-b-1}. \quad (105)$$

Set

$$\bar{K} := \frac{2nd\varepsilon^{-d}(2d\varepsilon^{-d}-1)}{4\log 2} b(m+1)2^m \varepsilon^{-b-1}. \quad (106)$$

Then

$$\bar{K} \geq \frac{2nD(2D-1)}{4\log 2} \Theta(F) - \frac{n+2D}{2}. \quad (107)$$

Clearly there exist positive constants \bar{c} and c independent of ε such that $\bar{c}\varepsilon^{-c} \geq \bar{K}$. Let $K \in \mathbb{N}$ and $K \geq \bar{c}\varepsilon^{-c} \geq \bar{K}$. According to Lemma 8, there exists a homogeneous SOS polynomial s_0 of degree $2(D+K)$ such that $\|x\|_2^{2K} F = s_0$. It implies that

$$\begin{aligned} \|x\|_2^{2(D-d_f+K)} (f + \varepsilon \|x\|_2^{2d_f}) &= s_0 + \sum_{i=1}^m g_i q_i^2 \|x\|_2^{2(D-2nu_i-d_{g_i}+K)} \\ &= s_0 + \sum_{i=1}^m g_i s_i, \end{aligned} \quad (108)$$

where $s_i := q_i^2 \|x\|_2^{2(D-2nu_i-d_{g_i}+K)}$ is a homogeneous SOS polynomial such that $\deg(g_i s_i) = 2(K+D)$, for $i \in [m]$. Set $k = D - d_f + K$. Then $\|x\|_2^{2k} (f + \varepsilon \|x\|_2^{2d_f}) = s_0 + \sum_{i=1}^m g_i s_i$ with $\deg(s_0) = \deg(g_i s_i) = 2(k + d_f)$, for $i \in [m]$.

The case of the ice cream constraint. Assume that $m = 1$ and $g_1 = x_n^2 - \|x'\|_2^2$ with $x' := (x_1, \dots, x_{n-1})$. We shall show that $c = 65$. Using Lemma 3, we can take $\alpha_m = 2$ in (65). We then obtain the following asymptotic equivalences as $\varepsilon \rightarrow 0^+$:

$$\begin{aligned} \delta_m &\sim R_1 \varepsilon^2 \Rightarrow C_{\psi_m} \sim R_2 \varepsilon^{-2} \Rightarrow C_{\varphi_m} \sim R_3 \varepsilon^{-1} \Rightarrow w_m \sim R_4 \varepsilon^5 \Rightarrow L_{\sqrt{\varepsilon_m}} \sim R_5 \varepsilon^{-\frac{9}{2}} \\ &\Rightarrow L_{\bar{\varphi}_m} \sim R_6 \varepsilon^{-6} \Rightarrow C_{\bar{\varphi}_m} \sim R_7 \varepsilon^{-6} \Rightarrow u_m \sim R_8 \varepsilon^{-26} \Rightarrow d_m \sim R_9 \varepsilon^{-26} \\ &\Rightarrow C_F \sim R_{10} \varepsilon^{-12} \Rightarrow D \sim R_{11} \varepsilon^{-26} \Rightarrow b = 12 \Rightarrow d = 26 \Rightarrow \bar{K} \sim R_{12} \varepsilon^{-65} \\ &\Rightarrow c = 65. \end{aligned} \quad (109)$$

for some $R_j > 0$ independent of ε , $j \in [12]$. This completes the proof of Theorem 2.

4 Polynomial optimization

This section is concerned with some applications to polynomial optimization. Consider the following POP:

$$f^* := \inf_{x \in S} f(x), \quad (110)$$

where $f \in \mathbb{R}[x]$ and

$$S = \{x \in \mathbb{R}^n : g_j(x) \geq 0, j \in [m]\}, \quad (111)$$

for some $g_j \in \mathbb{R}[x]$, $j \in [m]$. Recall that $\theta = 1 + \|x\|_2^2$.

4.1 General case

In this subsection, we improve the convergence rate of the Moment-SOS hierarchy described in [9, Theorem 4.3], based on Putinar–Vasilescu’s Positivstellensatz [17].

Theorem 3. *Let f, g_1, \dots, g_m be polynomials such that f^* defined as in (110) and S defined as in (111) satisfy that S has nonempty interior and $f^* > -\infty$. Let $\varepsilon > 0$ and denote $g_0 := 1$. Let $d_f := \lfloor \deg(f)/2 \rfloor + 1$ and $d_{g_j} := \lceil \deg(g_j)/2 \rceil$, $j \in \{0\} \cup [m]$. Consider the hierarchy of semidefinite programs indexed by $k \in \mathbb{N}$:*

$$\begin{aligned} \tau_k^{(\varepsilon)} &:= \inf_{s.t.} L_y(\theta^k (f + \varepsilon \theta^{d_f})) \\ &\quad y = (y_\alpha)_{\alpha \in \mathbb{N}_{2(d_f+k)}^n} \subset \mathbb{R}, \\ &\quad M_{k+d_f}(y) \succeq 0, \\ &\quad M_{k+d_f-d_{g_j}}(g_j y) \succeq 0, j \in [m], \\ &\quad L_y(\theta^k) = 1. \end{aligned} \quad (112)$$

For every $k \in \mathbb{N}$, the dual of (112) reads as:

$$\rho_k^{(\varepsilon)} := \sup_{\lambda \in \mathbb{R}} \{\lambda : \theta^k (f - \lambda + \varepsilon \theta^{d_f}) \in \mathcal{Q}_{k+d_f}\}, \quad (113)$$

where \mathcal{Q}_r is defined as in (15). The following statements hold:

1. For all $k \in \mathbb{N}$,

$$\rho_k^{(\varepsilon)} \leq \rho_{k+1}^{(\varepsilon)} \leq f^*. \quad (114)$$

2. Assume that problem (110) has an optimal solution x^* . Then there exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that for all $k \geq \bar{c}\varepsilon^{-c}$,

$$0 \leq \rho_k^{(\varepsilon)} - f^* \leq \varepsilon \theta(x^*)^{d_f}. \quad (115)$$

3. Strong duality holds for all orders k of the primal-dual problems (112)-(113).

The proof of Theorem 3 is exactly the same as the proof of [9, Theorem 7]. The second statement relies on Corollary 1. The third statement is due to the Slater condition [24, Theorem 3.1] since S has nonempty interior (see in detail [9, Proposition 2 and Remark 3]).

4.2 Compact case

In this subsection, we consider the case when S is compact by assuming that a ball constraint is present. We can then remove the perturbation term $\varepsilon \theta^{d_f}$ in the hierarchy based on Putinar–Vasilescu’s Positivstellensatz, described in the previous subsection.

Theorem 4. *Let f, g_1, \dots, g_m be polynomials such that f^* defined as in (110) and S defined as in (111) satisfy that S has nonempty interior and $f^* > -\infty$. Denote $g_0 := 1$. Let $d_f := \lfloor \deg(f)/2 \rfloor + 1$ and $d_{g_j} := \lceil \deg(g_j)/2 \rceil$, $j \in \{0\} \cup [m]$. Assume that $g_1 = L - \|x\|_2^2$ for some $L > 0$. Consider the hierarchy of semidefinite programs indexed by $k \in \mathbb{N}$:*

$$\begin{aligned} \tau_k^{(0)} &:= \inf && L_y(\theta^k f) \\ &s.t. && y = (y_\alpha)_{\alpha \in \mathbb{N}_{2(d_f+k)}^n} \subset \mathbb{R}, \\ &&& M_{k+d_f}(y) \succeq 0, \\ &&& M_{k+d_f-d_{g_j}}(g_j y) \succeq 0, \quad j \in [m], \\ &&& L_y(\theta^k) = 1. \end{aligned} \quad (116)$$

For every $k \in \mathbb{N}$, the dual of (116) reads as:

$$\rho_k^{(0)} := \sup_{\lambda \in \mathbb{R}} \{ \lambda : \theta^k (f - \lambda) \in \mathcal{Q}_{k+d_f} \}, \quad (117)$$

where \mathcal{Q}_r is defined as in (15). The following statements hold:

1. For all $k \in \mathbb{N}$,

$$\rho_k^{(0)} \leq \rho_{k+1}^{(0)} \leq f^*. \quad (118)$$

2. There exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that

$$0 \leq f^* - \rho_k^{(0)} \leq \left(\frac{\bar{c}}{k} \right)^{\frac{1}{c}} \quad (119)$$

3. Strong duality holds for all orders k of the primal-dual problems (116)-(117).

Proof. The first and third statements of Theorem 4 can be proved similarly to the ones of Theorem 3. Let us prove the second statement. By using Corollary 2, there exist positive constants \bar{c} and c depending on f, g_1, \dots, g_m such that for any $\varepsilon > 0$, for all $k \geq \bar{c}\varepsilon^{-c}$,

$$\theta^k (f - f^* + \varepsilon) \in \mathcal{Q}_{k+d_f}. \quad (120)$$

Let $K \in \mathbb{N}$. Set $\epsilon = \left(\frac{\bar{c}}{K} \right)^{\frac{1}{c}}$. Then $\epsilon > 0$ and $K = \bar{c}\epsilon^{-c}$, so that

$$\theta^K (f - f^* + \epsilon) \in \mathcal{Q}_{K+d_f}. \quad (121)$$

It implies that $f^* - \epsilon$ is a feasible solution of (117) with relaxation order K , yielding

$$0 \leq f^* - \rho_K^{(0)} \leq f^* - (f^* - \epsilon) = \epsilon = \left(\frac{\bar{c}}{K}\right)^{\frac{1}{c}}. \quad (122)$$

Hence the desired result follows. \square

Remark 3. *The authors' bounds in Theorems 3 and 4 are only for worst cases. In fact, for generic cases of polynomials, the Moment-SOS hierarchy based on Putinar's Positivstellensatz (in [8]) has finite convergence [10].*

5 Conclusion

We have provided a new degree bound on the sum-of-squares (SOS) polynomials involved in Putinar–Vasilescu's Positivstellensatz. The Moment-SOS hierarchy of semidefinite relaxations based on this Positivstellensatz provide a sequence of lower bounds on the minimum of a polynomial on a basic compact semialgebraic set. Moreover, this sequence of lower bounds converges to the minimum with the rate $\mathcal{O}(\varepsilon^{-c})$, at prescribed accuracy $\varepsilon > 0$. We emphasize that our result yields better convergence rates than the previous rate $\mathcal{O}(\exp(\varepsilon^{-c}))$ one by Nie and Schweighofer [11] for Putinar's Positivstellensatz, at the price of making stronger assumptions, namely that the set of constraints has nonempty interior and involves a ball constraint, and denominators are used.

A topic of further investigation is the analysis of the convergence rate of the Moment-SOS hierarchy for lower bounds in some special cases of basic (compact) semialgebraic sets. A fairly straightforward idea is to find the explicit constant α in the Łojasiewicz inequality stated in Lemma 2. We could then proceed analogously to the proof of the rate $\mathcal{O}(\varepsilon^{-65})$ for the minimization of a polynomial on the unit ball.

Acknowledgements. The first author was supported by the MESRI funding from EDMITT. The second author was supported by the Tremplin ERC Stg Grant ANR-18-ERC2-0004-01 (T-COPS project) and by the FMJH Program PGMO (EPICS project) and EDF, Thales, Orange et Criteo. This work has benefited from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Actions, grant agreement 813211 (POEMA) as well as from the AI Interdisciplinary Institute ANITI funding, through the French “Investing for the Future PIA3” program under the Grant agreement n°ANR-19-PI3A-0004.

References

- [1] A. Ben-Tal and A. Nemirovski. *Lectures on modern convex optimization: analysis, algorithms, and engineering applications*, volume 2. Siam, 2001.
- [2] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36. Springer Science & Business Media, 2013.
- [3] A. C. Doherty and S. Wehner. Convergence of SDP hierarchies for polynomial optimization on the hypersphere. *arXiv preprint arXiv:1210.5048*, 2012.
- [4] K. Fang and H. Fawzi. The sum-of-squares hierarchy on the sphere and applications in quantum information theory. *Mathematical Programming*, pages 1–30, 2020.
- [5] C. Heitzinger. *Simulation and inverse modeling of semiconductor manufacturing processes*. na, 2002.

- [6] T. Hildebrandt and I. Schoenberg. On linear functional operations and the moment problem for a finite interval in one or several dimensions. *Annals of Mathematics*, pages 317–328, 1933.
- [7] M. Kirszbraun. Über die zusammenziehende und Lipschitzsche Transformationen. *Fundamenta Mathematicae*, 22(1):77–108, 1934.
- [8] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on optimization*, 11(3):796–817, 2001.
- [9] N. H. A. Mai, J.-B. Lasserre, and V. Magron. Positivity certificates and polynomial optimization on non-compact semialgebraic sets. *Mathematical Programming*, pages 1–43, 2021.
- [10] J. Nie. Optimality conditions and finite convergence of Lasserre’s hierarchy. *Mathematical programming*, 146(1-2):97–121, 2014.
- [11] J. Nie and M. Schweighofer. On the complexity of Putinar’s Positivstellensatz. *Journal of Complexity*, 23(1):135–150, 2007.
- [12] L. Perko. *Differential equations and dynamical systems*, volume 7. Springer Science & Business Media, 2013.
- [13] T. S. Pham and H. H. Vui. *Genericity in polynomial optimization*, volume 3. World Scientific, 2016.
- [14] G. Pólya. Über Positive Darstellung von Polynomen. *Vierteljschr. Naturforsch. Ges. Zürich*, 73:141–145, 1928.
- [15] V. Powers and B. Reznick. A new bound for Pólya’s Theorem with applications to polynomials positive on polyhedra. *Journal of pure and applied algebra*, 164(1-2):221–229, 2001.
- [16] M. Putinar. Positive polynomials on compact semi-algebraic sets. *Indiana University Mathematics Journal*, 42(3):969–984, 1993.
- [17] M. Putinar and F.-H. Vasilescu. Positive polynomials on semi-algebraic sets. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, 328(7):585–589, 1999.
- [18] M. Putinar and F.-H. Vasilescu. Solving moment problems by dimensional extension. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, 328(6):495–499, 1999.
- [19] B. Reznick. Uniform denominators in Hilbert’s seventeenth problem. *Mathematische Zeitschrift*, 220(1):75–97, 1995.
- [20] K. Schmüdgen. The K-moment problem for compact semi-algebraic sets. *Mathematische Annalen*, 289(1):203–206, 1991.
- [21] C. Schulze. Schmüdgen’s theorem and results of positivity. *arXiv preprint arXiv:1411.4446*, 2014.
- [22] M. Schweighofer. On the complexity of Schmüdgen’s Positivstellensatz. *Journal of Complexity*, 20(4):529–543, 2004.
- [23] L. Slot and M. Laurent. Improved convergence analysis of Lasserre’s measure-based upper bounds for polynomial minimization on compact sets. *Mathematical Programming*, pages 1–41, 2020.
- [24] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM review*, 38(1):49–95, 1996.